

INTERNATIONAL DOCTORAL
SCHOOL OF THE USC

Aida
Bao Caamaño

PhD Thesis

Epigenomic characterization of circulating tumour cells in advanced colorectal cancer

Santiago de Compostela, 2022

Doctoral Programme in Molecular Medicine

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

DOCTORAL THESIS

EPIGENOMIC CHARACTERIZATION OF CIRCULATING TUMOUR CELLS IN ADVANCED COLORECTAL CANCER

Aida Bao Caamaño

INTERNATIONAL DOCTORAL SCHOOL OF
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

PROGRAMA DE DOUTORAMENTO EN MEDICINA MOLECULAR

SANTIAGO DE COMPOSTELA
2022

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

TESE DE DOUTORAMENTO

EPIGENOMIC CHARACTERIZATION OF CIRCULATING TUMOUR CELLS IN ADVANCED COLORECTAL CANCER

Aida Bao Caamaño

ESCOLA DE DOUTORAMENTO INTERNACIONAL DA UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

PROGRAMA DE DOUTORAMENTO EN MEDICINA MOLECULAR

SANTIAGO DE COMPOSTELA

2022

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

DECLARACIÓN DA AUTORA DA TESE

Dna. **Aida Bao Caamaño**

Título da tese: **EPIGENOMIC CHARACTERIZATION OF CIRCULATING TUMOUR CELLS IN ADVANCED COLORECTAL CANCER**

Presento a miña tese, seguindo o procedemento axeitado ao Regulamento, e declaro que:

1) A tese abarca os resultados da elaboración do meu traballo.
2) De ser o caso, na tese faise referencia ás colaboracións que tivo este traballo.
3) Confirmo que a tese non incorre en ningún tipo de plaxio doutros autores nin de traballos presentados por min para a obtención doutros títulos.
4) A tese é a versión definitiva presentada para a súa defensa e coincide a versión impresa coa presentada en formato electrónico.

E comprométome a presentar o Compromiso Documental de Supervisión no caso de que o orixinal non estea na Escola.

En **Santiago de Compostela**, **21 de decembro de 2022**

Sinatura electrónica

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

AUTORIZACIÓN DO
DIRECTOR / TITOR DA TESE

EPIGENOMIC CHARACTERIZATION OF CIRCULATING TUMOUR CELLS IN ADVANCED COLORECTAL CANCER

D. Ángel Diaz Lagares
D. Rafael López López

INFORMAN:

Que a presente tese, correspóndese co traballo realizado por Dna. Aida Bao Caamaño, baixo a miña dirección/titorización, e autorizo a súa presentación, considerando que reúne os requisitos esixidos no Regulamento de Estudos de Doutoramento da USC, e que como director desta non incorre nas causas de abstención establecidas na Lei 40/2015.

En Santiago de Compostela, 21 de decembro de 2022.

Asdo. Ángel Díaz Lagares
(Director)

Asdo. Rafael López López
(Director e titor)

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CONFLICT OF INTEREST

The author, Aida Bao Caamaño, declares no conflict of interest in relation to the Doctoral thesis named:

EPIGENOMIC CHARACTERIZATION OF CIRCULATING TUMOUR CELLS IN ADVANCED COLORECTAL CANCER

Santiago de Compostela, 21st of December 2022.

Aida Bao Caamaño

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

DECLARATION OF THE AUTHOR

Aida Bao Caamaño, author of this thesis, was financially supported by a predoctoral contract *PFIS* from *Instituto de Salud Carlos III* (*ISCIII*) (FI19/00240) co-funded by *Fondo Social Europeo (FSE)* and was also funded by a predoctoral contract from *Xunta de Galicia* (IN606A-2019/001) and a fellowship from *Deputación de A Coruña (Xunta de Galicia).*

The work in this thesis was funded (PI18/00307) by the *Instituto de Salud Carlos III* (*ISCIII*) and the European Regional Development Fund (FEDER).

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Nothing is imposible to a determined woman

-**Louisa May Alcott**

O verdadeiro heroísmo consiste en transformar
os desexos en realidades e as ideas en feitos

-Castelao

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

A miña abueliña Carmen

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

INDEX

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

INDEX

ABBREVIATIONS AND ACRONYMS **23**
RESUMO *IN EXTENSO* **34**
RESUMEN *IN EXTENSO* **47**
SUMMARY *IN EXTENSO* **60**
INTRODUCTION **73**
1. Colorectal cancer **75**
1.1 Epidemiology 75
1.2 Development 76
1.3 Classification 78
1.4 Staging 80
1.5 Treatment 80
1.5.1 Treatment response and resistance 82
2. The metastatic cascade **86**
2.1 Local invasion 86
2.2 Intravasation, dissemination and extravasation 87
2.3 Colonization 89
3. Liquid biopsy **91**
3.1 Concept and characteristics 91
3.2 Circulating tumour cells 93
3.2.1 CTC isolation methods 94
3.3 Relevance of the study of CTCs 97
3.4 CTC cell lines as models for the study of CTCs 100
3.4.1 Colorectal CTC cell lines 100
4. Epigenetics **102**
4.1 Epigenetic mechanisms 103
4.1.1 DNA Methylation 103
4.1.2 Other epigenetic mechanisms 105

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.2 Interplay between epigenetic mechanisms 109
4.3 DNA Methylation and cancer 110
4.3.1 Relevance of the study of DNA methylation in cancer ...111
5. CTCs, DNA methylation and CRC 113
HYPOTHESIS 115
OBJECTIVES 119

CHAPTER I Epigenomic characterization of the metastasis-competent cell line CTC-MCC-41 respect to primary and metastatic CRC cell lines **.....123**
1. Abstract 125
2. Introduction 126
3. Materials and methods 127
3.1 Cell culture 127
3.2 Cell line treatment with AZA 129
3.3 Nucleic acids extraction and quantification 130
3.4 DNA methylation analysis 131
3.5 Expression analysis 137
3.6 Statistical analysis 138
4. Results 139
5. Discussion 157

CHAPTER II Epigenomic evaluation of colorectal CTC cell lines in association with therapy response **.....163**
1. Abstract 165
2. Introduction 166
3. Materials and methods 167
3.1 Cell culture 167
3.2 Nucleic acids extraction and quantification 168
3.3 DNA methylation analysis 169
3.4 Expression analysis 171
3.5 Statistical analysis 173

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4. Results **173**

5. Discussion **187**

CHAPTER III Epigenomic analysis of CTCs from mCRC patients in relation to therapy response and the progression of the disease **193**

1. Abstract **195**

2. Introduction **196**

3. Materials and methods **198**

3.1 Patient cohort 198

3.2 Enumeration and enrichment of CTCs 200

3.3 Isolation of CTCs 202

3.4 DNA methylation analysis of CTCs 203

3.5 Statistical analysis 204

4. Results **205**

5. Discussion **229**

OVERALL DISCUSSION **235**

CONCLUSIONS **239**

SUPPLEMENTARY MATERIAL **243**

REFERENCES **265**

AGRADECIMIENTOS **295**

ANNEXES **301**

Annex I. Ethics committee approval 303

Annex II. Images use 305

Annex III. Publications 307

Annex IV. Other publications 309

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

ABBREVIATIONS AND ACRONYMS

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

ABBREVIATIONS AND ACRONYMS

	A
ABCB1	ATP Binding Cassette Subfamily B Member 1
ACTA2	Actin alpha2
AJCC	American Joint Committee on Cancer
AKT	Protein kinase B
AMER1	APC Membrane Recruitment Protein 1
AP2	Assembly protein complex 2
AP2M1	Adaptor Related Protein Complex 2 Subunit Mu 1
APC	Adenomatous polyposis coli gene
APC	Allophycocyanin
ARID1A	AT-Rich Interaction Domain 1A
ARRB1	Arrestin beta 1
ATP	Adenosine triphosphate
AZA	5-aza-2′-deoxycytidine
	B
B2M	Beta-2-microglobulin
BCL11A	BAF Chromatin Remodeling Complex Subunit BCL11A
BDNF	Brain Derived Neurotrophic Factor
bFGF	Basic human fibroblast growth factor
BMIQ	Beta-mixture quantile
BMP3	Bone Morphogenetic Protein 3
BMP7	Bone Morphogenetic Protein 7
bp	Base pair
BRAF	B-Raf proto-oncogene, serine/threonine kinase
Btn	Biotin
	C
CASK	Calcium/Calmodulin Dependent Serine Protein Kinase
CCND2	Cyclin D2

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

CCS	Colon Cancer Subtype
CDH3	Cadherin 3
CD44v6	Cluster of differentiation 44 variant 6
CD45	Cluster of differentiation 45
CD47	Cluster of differentiation 47
CD68+	Cluster of Differentiation 68
CD8+	Cluster of differentiation 8
cDNA	Complementary DNA
CEA	Carcinoembriogenic antigen
cfDNA	Cell-free DNA
CpG	Cytosine-phosphate-guanine dinucleotids
CpGIs	CpG islands
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
CK	Citokeratin
CMS	Consensus molecular subtypes
COL9A3	Collagen Type IX Alpha 3 Chain
CRC	Colorectal cancer
CRCSC	CRC Subtyping Consortium
CTC	Circulating tumour cell
CTNNB1	Catenin Beta 1
CXCL1	C-X-C Motif Chemokine Ligand 1
CXCR2	C-X-C Motif Chemokine Receptor 2

DKK1	Dickkopf WNT Signaling Pathway Inhibitor 1
DLL1	Delta Like Canonical Notch Ligand 1
DMCpGs	Differentially methylated CpGs
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferases
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double strand DNA

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

E

ECM	Extracellular matrix
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchimal transition
EpCAM	Epithelial Cell Adhesion Molecule
ERCC1	Excision Repair cross-complementation group 1
ExonBnd	Exon boundaries

F

FBS	Fetal bovine serum
FDR	False Discovery Rate
FGF	Fibroblast growth factor
FKBP11	FKBP Prolyl Isomerase 11
FN1	Fibronectin 1
FOLFIRI	5-Fluorouracilo + irinotecan
FOLFOX	5-Fluorouracilo + oxaliplatino
FOLFOXIRI	5-fluorouracilo + irinotecan + leucovorina
FZD	Frizzled Receptor
FZD5	Frizzled Class Receptor 5

G

GAL	Galanin And GMAP Prepropeptide
GALNT6	Polypeptide N-Acetylgalactosaminyltransferase 6
GATA2	GATA Binding Protein 2
GJB6	Gap Junction Protein Beta 6
GNG4	G Protein Subunit Gamma 4

H

h	Hours
hEGF	Human epidermal growth factor
HER2	Human epidermal growth factor receptor 2
hg38	Homo sapiens (human) genome assembly GRCh38

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

HGF	Hepatocyte growth factor
HOXB9	Homeobox B9
HyperM	Hypermethylated
HypoM	Hypomethylated
	I
IGF2	Insulin Like Growth Factor 2
IKBKB	Inhibitor of Nuclear Factor Kappa B Kinase Subunit Beta
IL-6	Interleukin 6
ITS-G	Insulin-transferrin-selenium
	J
JAK	Janus kinase
	K
kb	Kilobase
KRAS	Kirsten rat sarcoma virus
	L
LEF1/*LEF1*	Lymphoid Enhancer Binding Factor 1
LINE	Long interspread transposable elements
lncRNAs	Long non-coding RNAs
LPXN	Leupaxin
LYN	LYN proto-oncogene, Src family tyrosine kinase
	M
mAbs	Monoclonal antibodies
MACS	Magnetic cell sorting
MAPK	Mitogen-activated protein kinase
mbar	Milibar
mCRC	Metastatic colorectal cancer
MDR	Multidrug resistance
MGMT	Cyclin Dependent Kinase Inhibitor 2A
min	Minutes
mL	Mililitres
MLH1	MutL Homolog 1

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

mmol	Milimoles
MMPs	Matrix metalloproteinases
mol	Moles
mRNA	Messenger ribonucleic acid
MSH2	MutS Homolog 2
MSH6	MutS Homolog 6
MSI	Microsatellite instability
mtDNA	Mitochondrial DNA
mTOR	Mammalian target of rapamycin
mTORC1	Mammalian target of rapamycin complex 1
Myc	MYC Proto-Oncogene, BHLH Transcription Factor

N

ncRNAs	Non-coding RNAs
NDRG4	NDRG Family Member 4
NER	Nucleotide excision repair
ng	Nanograms
NGS	Next generation sequencing
nm	Nanometer
NME2	Nucleoside diphosphate kinase B
Notch	Notch receptor 1
NR4A1	Nuclear Receptor Subfamily 4 Group A Member 1
NRAS	NRAS proto-oncogene, GTPase
NRF2	Nuclear factor erythroid 2-related factor 2
nt	Nucleotides

O

OMP-18R5	Vantictumab
OS	Overall survival

P

p14	Cyclin Dependent Kinase Inhibitor 2A
p53	Tumour protein p53
PCA	Principal component analysis

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

PCDHA	Protocadherin alpha cluster
PCR	Polymerase chain reaction
PD1	Programmed cell death protein 1
PDGF	Platelet derived growth factor
PDGFC	Platelet Derived Growth Factor C
PD-L1	Programmed cell death ligand 1
PE	Phycoerythrin
PFS	Progression free survival
PI3K	Phosphatidylinositol 3-kinases
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
PMS1	PMS1 Homolog 1, Mismatch Repair System Component
PMS2	PMS1 Homolog 2, Mismatch Repair System Component
PTEN	Phosphatase And Tensin Homolog
PTGS2	Prostaglandin-Endoperoxide Synthase 2

Q

qMSP	Quantitative methylation-specific PCR
qPCR	Quantitative polimerase chain reaction

R

RAF	Rapidly accelerated fibrosarcoma
RAS	Rat sarcoma virus
RASSF1A	Ras Association Domain Family Member 1
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
RT-qPCR	Quantitative reverse transcription PCR

S

SDS	Sodium dodecyl sulfate
sec	Seconds
SEMA6A	Semaphorin 6A
SEPT9	Septin 9
SFRP2	Secreted Frizzled Related Protein 2

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

SGI-110	Guadecitabine
SINE	Short interspread transposable elements
SMARCA1	SWI/SNF Related, Matrix Associated, Actin Dependent Regulator of Chromatin, Subfamily A, Member 1
SNAIL	Zinc finger protein SNAI1
SNPs	Single nucleotide polymorphysms
SOX9	SRY-Box Transcription Factor 9
STAT	Signal transducer and activator of transcription

T

TET	Ten-eleven translocation
TFAP2E	Transcription Factor AP-2 Epsilon
TFBS	Transcription factor binding sites
TGF	Transcription growth factor
TGFB2	Transforming growth factor Beta 2
TIC	Tumour initiating cells
TNM	Tumour, Node and Metastasis
TP53	Tumour protein p53 gene
tRNA	Transfer ribonucleic acid
TSSs	Transcription start sites
TWIST	Twist-related protein 1

U

UTR	Untranslated region

V

VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
VIM/*VIM*	Vimentin

W

WBSCR22	BUD23 rRNA methyltransferase and ribosome maturation factor
WGBS	Whole-genome bisulfite sequencing
Wnt	Wingless-related integration site

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

	X
XDP	Xeroderma pigmentosum group D
XRCC1	X-Ray Repair Cross Complementing 1
	Y
Yrs	Years
	Z
ZEB	Zinc finger E-box-binding homeobox
	Others
5-FU	Fluorouracil
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
µg	Micrograms
µL	Microlitres
µm	Micrometer
µM	Micromolar

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

RESUMO *IN EXTENSO*

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

RESUMO *IN EXTENSO*

O cancro colorrectal (CCR) é o terceiro cancro máis comúnmente diagnosticado e a segunda causa principal de mortes relacionadas co cancro no mundo. Ademáis, estímase que tanto as taxas de incidencia coma de mortalidade seguirán aumentando nos próximos anos. A maioría dos CCR xeralmente desenvólvense a partir de pólipos, que son protuberancias na luz do colon debido a crecementos localizados ou agregacións de células anormais dentro da mucosa do colon. Porén, só unha pequena porcentaxe de pólipos adquire características malignas e a progresión cara ao cancro resulta de diferentes cambios xenéticos e epixenéticos que se acumulan ao longo do tempo e que poden tardar ata unha década en completarse. O CCR pódese clasificar segundo criterios moi diversos, como os factores que o orixinan, a localización do tumor, a histoloxía, as vías moleculares, o estado mutacional, a orixe celular ou a expresión xénica. Non obstante, o CCR é unha enfermidade moi heteroxénea, o que dificulta a súa clasificación dun xeito clínicamente relevante.

A terapia sistémica para o CCR avanzado e metastásico baséase no uso de quimioterapia, terapias dirixidas e inmunoterapia, incluíndo varios fármacos en combinación ou coma axentes únicos. Para a maioría dos pacientes con CCR metastásico (CCRm), o tratamento consiste na administración inicial dunha fluoropirimidina (*e.g.* 5-fluorouracilo) xeralmente en combinación con oxaliplatino (*e.g.* FOLFOX) ou irinotecán (*e.g.* FOLFIRI) para mellorar as taxas de resposta, que se poden considerar opcións estándar de tratamento de primeira liña. É bastante común engadir terapias dirixidas anti-EGFR (*e.g.* cetuximab) ou anti-VEGF (*e.g.* bevacizumab) a estes réximes, que demostraron nalgúns casos mellorar os resultados. Porén, a pesar dos avances terapéuticos no tratamento do CCR, a resistencia

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

aos medicamentos, xa sexa primaria ou adquirida, é un gran obstáculo para mellorar as taxas de supervivencia dos pacientes con enfermidade avanzada.

Neste sentido, a enfermidade metastásica segue sendo un reto e é responsable de máis do 90% das mortes por cancro. A metástase, que é a última etapa da progresión do cancro, é unha sucesión de procesos biolóxicos polos que as células do tumor primario se propagan e colonizan un órgano distante. Unha vez que as células tumorais están no lume dos vasos sanguíneos ou linfáticos, pódense diseminar pola circulación coñecéndose como células tumorais circulantes (CTCs). Estas CTCs, derivadas dun tumor primario ou metastásico, son as responsables do desenvolvemento das metástases. O seu estudo foi o que acuñou inicialmente o termo de biopsia líquida, termo ao que hoxe se refire tamén á análise doutros compoñentes derivados de tumores coma ácidos nucleicos circulantes libres, plaquetas educadas por tumores ou vesículas extracelulares coma os exosomas, illadas principalmente do sangue, pero tamén de diferentes fluídos biolóxicos.

A biopsia líquida é unha ferramenta con gran potencial para mellorar o manexo clínico dos pacientes, xa que supera as limitacións da biopsia de tecido, que é o estándar para a caracterización tumoral. A biopsia líquida é unha técnica mínimamente invasiva que permite cuestionar de xeito reiterado e en tempo real a evolución dinámica dos tumores, reflectindo a heteroxeneidade molecular das diferentes lesións tumorais coexistentes nun paciente. Así, podería axudar no diagnóstico precoz, predición do prognóstico, avaliación e seguimento da resposta á terapia e á estratificación do paciente en función da detección de dianas terapéuticas ou mecanismos de resistencia. Non obstante, a pesar das súas numerosas vantaxes, a detección de CTCs no sangue periférico dos pacientes con cancro supón un reto técnico, xa que son células extremadamente pouco frecuentes no sangue e moi heteroxéneas. As tecnoloxías para o illamento de CTCs deben ter unha alta sensibilidade e especificidade analítica e, polo xeral, baséanse

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

en características que diferencian as CTCs das células sanguíneas. Poden basearse en características biolóxicas, como viabilidade ou expresión de proteínas na superficie celular, ou en características físicas como tamaño, deformabilidade e cargas eléctricas, ou unha combinación de ambas.

O estudo destas CTCs podería axudar a comprender mellor a bioloxía do proceso metastásico e, tanto a enumeración como a caracterización molecular, poderían implementarse na clínica como unha valiosa ferramenta con significado pronóstico, para o seguimento en tempo real do tratamento ou a identificación de biomarcadores, dianas e mecanismos de resistencia. Actualmente, o único uso validado clínicamente das CTCs baséase na súa enumeración no contexto da enfermidade metastásica. Neste sentido, o CellSearch é a única plataforma aprobada pola *Food and Drug Administration* (FDA) dos Estados Unidos validada clinicamente para a detección de CTCs en diferentes tipos de cancro, incluído o CRC. Non obstante, máis aló da súa enumeración, a caracterización molecular das CTCs tamén ten proporcionado información valiosa non só para comprender o proceso metastásico senón tamén para a súa transferencia á práctica clínica. Porén, o estudo da epixenética en CTCs en pacientes con CCR segue sendo moi limitado.

A epixenética defínese como o estudo dos cambios na función dos xenes que son herdables mitóticamente e/ou meióticamente e que non implican cambios na secuencia do ADN. Entre as súas características, é de gran relevancia que as marcas epixenéticas sexan reversibles e poidan ser moduladas por factores ambientais. A marca epixenética máis estudada no xenoma humano é a metilación do ADN. Esta marca consiste na modificación covalente do ADN mediante a incorporación dun grupo metilo (CH_3) ao carbono 5′ das citosinas en dinucleótidos de citosina-fosfato-guanina (CpG) para producir 5-metilcitosina (5mC). É esencial para a regulación da expresión xénica no desenvolvemento e diferenciación celular normal. A

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

metilación das citocinas ocorre case principalmente no contexto dos dinucleótidos CpG, que tenden a agruparse en rexións chamadas illas CpG (CpGI). Xeralmente, a metilación de CpGI no promotor dun xene está inversamente asociada coa súa transcrición, asociando a hipermetilación co silenciamento xenético e a hipometilación coa expresión xénica. Ademais das CpGI, a metilación do ADN tamén se describiu noutras rexións que, segundo a súa distancia dos CpGI, denomínanse CpGI *shores*, CpGI *shelves* e *open sea*. As alteracións nos patróns de metilación do ADN están relacionadas con un gran número de enfermidades humanas, incluíndo o cancro. En xeral, ocorren dous tipos de cambios nas células tumorais con respecto ás células normais, que son a hipometilación xeneralizada do xenoma e a hipermetilación de sitios específicos en CpGI dos promotores dos xenes. A metilación do ADN é un biomarcador prometedor debido á súa aparición precoz e frecuente no cancro, á súa detección relativamente fácil mediante técnicas ben establecidas, á súa estabilidade, á súa presenza nunha ampla gama de fluídos corporais e á súa especificidade tisular.

En vista de todo o anterior, o obxectivo desta tese foi a caracterización epixenómica do perfil de metilación do ADN das CTCs en CCRm, co fin de achegar nova información sobre a bioloxía destas células e o desenvolvemento das metástases; así como revelar novas dianas terapéuticas, mecanismos de resistencia e biomarcadores potenciais que poderían beneficiar aos pacientes con CCRm na práctica clínica.

Para iso, no primeiro capítulo analizouse o metiloma do ADN da liña celular de CTC CTC-MCC-41 en comparación coas liñas celulares de cancro colorrectal primario (HT29, Caco2, HCT116, RKO) e metastásico (SW620 e COLO205). As liñas celulares de cancro úsanse amplamente para estudar os mecanismos moleculares da bioloxía celular tumoral e para identificar novos biomarcadores, dianas terapéuticas e mecanismos de resistencia. En concreto, esta liña CTC-MCC-41 foi establecida por Cayrefourcq L. *et al.* dun

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

paciente con CCRm, sendo a primeira liña celular de CTCs de cancro colorrectal. Esta liña celular de CTCs, así coma o resto das liñas celulares primarias e metastásicas, foron postas en cultivo e recollidas para a súa extracción de ADN, que foi tratado posteriormente con bisulfito de sodio, unha técnica moi utilizada para o estudo da metilación do ADN que convirte as citosinas non metiladas en uracilo. O ADN bisulfitado das liñas hibridouse en Infinium MethylationEPIC BeadChips, un *array* que permite interrogar o estado de metilación de máis de 850.000 sitios CpG en todo o xenoma humano cunha resolución dun só nucleótido. Os resultados obtidos no *array* revelaron un perfil de metilación diferente entre CTC-MCC-41 e HT29. Tanto as CpG hipometiladas coma as hipermetiladas de CTC-MCC-41 distribuíronse de xeito similar dependendo do contexto de CpG e da rexión do xene e, nunha porcentaxe moi pequena, tamén se distribuíron nos sitios de unión do factor de transcrición (TFBS) ou nas rexións *enhancer*. Ao centrar a análise nos CpGI e *shore* dos promotores dos xenes, que son rexións nas que a desregulación epixenética foi descrita como unha característica moi relevante no cancro, tamén se identificou un perfil de metilación capaz de diferenciar claramente CTC-MCC-41 das células de tumor primario HT29. Ademáis, a análise das vías de Panther realizada cos xenes diferencialmente metilados destas CpG nos promotores, revelou un enriquecemento de diversos procesos biolóxicos no desenvolvemento e metástase do cancro, tales como: regulación da transcrición, adhesión e diferenciación celular, apoptose, procesos metabólicos e reparación do ADN, entre outros; e diferentes vías como: a sinalización de Wnt, a vía da cadherina, a inflamación, a anxioxénese e a sinalización das integrinas.

Tamén se avaliou o perfil de metilación da liña CTC-MCC-41 con respecto á liña celular de cancro colorrectal metastásico COLO205. Neste caso, tamén se identificou un perfil de metilación que diferencia claramente ambas liñas celulares. A maioría das CpG diferencialmente metiladas

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

atopadas en CTC-MCC-41 con respecto a COLO205 estaban hipometilados, e tamén mostraron unha distribución similar segundo o contexto CpG e a rexión xenética, e unha pequena porcentaxe destes localizáronse en rexións TFBS e *enhancer*. Do mesmo xeito, cando a análise se centrou nas CpGI e *shore* dos promotores dos xenes, CTC-MCC-41 tamén mostrou un perfil de metilación diferente en comparación coa liña tumoral COLO205. A análise de ontoloxías xenéticas revelou o enriquecemento de procesos biolóxicos clave, como a regulación da transcrición, a adhesión e diferenciación celular, a apoptose, o ciclo celular e a proliferación, os procesos metabólicos e a reparación do ADN. Ademáis, a análise da vía Panther revelou un enriquecemento de vías importantes relacionadas co desenvolvemento do cancro e a metástase, como a sinalización Wnt, a vía da cadherina, a inflamación, a anxioxénese e a sinalización de integrinas, entre outras. En liña con isto, para confirmar a existencia dunha sinatura de metilación específica de CTCs no CCR, incluíronse na análise outras liñas celulares de CCR que representan tumores primarios (Caco2, HCT116, RKO) e metastásicos distantes (SW620). Así, obtivemos un perfil de metilación específico que permitiu clasificar as células de CCR segundo o seu tipo celular: CTC, liña celular tumoral primaria ou metastásica.

Os datos de expresión da liña celular CTC-MCC-41 en comparación coa liña celular tumoral primaria HT29, foran publicados previamente nun traballo de Cayrefourq *et al*. Estes datos utilizáronse para determinar se os cambios na metilación do ADN en CpGI e shore dos promotores dos xenes poderían regular a expresión xénica nesta liña celular de CTCs. Neste sentido, centramos a nosa análise naqueles xenes cuxa expresión xa fora validada (expresión disminuída en CTC-MCC-41: *TGFB2, DKK1, GJB6, PTGS2, PDGFC, SMARCA1* e *GATA2*; sobreexpresados en CTC-MCC-41: *BMP7, BCL11A, SEMA6A, FN1, ABCB1, CCND2* e *GAL*). Os valores de metilación obtidos no *array* destes xenes validáronse mediante

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

pirosecuenciación ou PCR cuantitativa específica de metilación (qMSP), obtendo os valores de metilación esperados en todos os xenes, a excepción de *GATA2*. Ademáis, para determinar a correlación entre a expresión e a metilación, a liña celular CTC-MCC-41 foi tratada con 5-Aza-2′-desoxicitidina, un epifármaco que inhibe a metilación. Despois do tratamento, recolléronse as células para a extracción do ARN e, tras a retrotranscripción, analizouse a expresión dos xenes de interese mediante PCR cuantitativa con transcripción inversa (RT-qPCR). Estes resultados mostraron que a expresión dos xenes silenciados epixenéticamente *TGFB2, DKK1, GJB6* e *PTGS2* foi completamente restaurada, evidenciando que o programa epixenético dos xenes hipermetilados nas CTCs colorrectais foi revertido, favorecendo a súa expresión. Estes resultados permitiron, polo tanto, evidenciar que a liña CTC estudada ten un perfil de metilación tanto a nivel global coma nas CpG de promotores diferente ao das liñas celulares tumorais primarias e metastásicas. Ademais, os xenes metilados de forma aberrante que se identificaron son potenciais biomarcadores e dianas terapéuticas para o CCRm.

Seguindo cas análises con liñas celulares derivadas de CTCs de pacientes con CCR, no segundo capítulo cuestionámonos se o perfil de metilación do ADN das CTCs podería sufrir cambios asociados coa resistencia á terapia e a progresión da enfermidade. Para iso, estudamos o metiloma dunha colección única de nove liñas celulares humanas establecidas por Cayrefourcq *et al.* e Soler *et al*., a partir de CTCs dun paciente con CCRm obtidas en diferentes momentos durante o tratamento e a progresión da enfermidade: A liña CTC-MCC-41, que se obtivo antes do inicio do tratamento; CTC-MCC-41.4, obtida despois do tratamento de primeira liña con FOLFIRI combinado con bevacizumab e despois do tratamento de segunda liña con FOLFOX e bevacizumab; e as liñas CTC-MCC-5A, CTC-MCC-5B, CTC-MCC-5C, CTC-MCC-5D, CTC-MCC-5E, CTC-

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

MCC-5F e CTC-MCC-5G (CTC-MCC-5 [A-G]), que foron obtidas unha semana antes da morte do paciente. Todas estas liñas de CTCs puxéronse en cultivo e realizouse a extracción do ADN que, despois de ser tratado con bisulfito de sodio, hibridouse no Infinium Methylation EPIC BeadChips. Os resultados do *array* revelaron un perfil de metilación diferente da liña celular CTC-MCC-41 de pre-tratamento con respecto a todas as liñas celulares de CTC resistentes ao tratamento. As CpG diferencialmente metiladas entre esta liña de pre-tratamento e as liñas celulares resistentes mostraron unha ampla distribución ao longo de todos os cromosomas do xenoma, coa liña de pre-tratamento mostrando unha metilación maior que as células resistentes. Ademáis, tanto as CpG hipermetiladas coma as hipometiladas de CTC-MCC-41 mostraron unha distribución similar segundo a rexión xenética e o contexto CpG, e só unha pequena porcentaxe destes CpG estaban en rexións TFBS e *enhancer*. Así mesmo, o estudo do enriquecemento das vías KEGG e a análise das vías Panther para os xenes diferencialmente metilados entre CTC-MCC-41 e as liñas de CTCs resistentes, revelou un predominio de vías relevantes relacionadas coa sinalización PI3K/AKT e MAPK, o metabolismo, a sinalización Wnt, a sinalización de cadherinas e integrinas, inflamación e apoptose, entre outros.

As diferencias no perfil de metilación entre as células pretratamento CTC-MCC-41 e o resto de células resistentes (CTC-MCC-41.4 e CTC-MCC-41.5 [A-G]), concorda cos datos de expresión previamente publicados para estas liñas nun traballo de Cayrefourcq *et al.*, nos que ambos grupos mostran un programa de expresión diferente. En base a isto, e coa idea de identificar biomarcadores asociados á resistencia á terapia, avaliouse a relación entre a metilación das illas e rexións *shore* dos promotores entre estes dous grupos de CTCs (pre-tratamento e resistentes). Deste xeito, identificamos seis xenes candidatos que presentaban maior metilación e menor expresión en CTC-MCC-41 que nas liñas resistentes ao tratamento: *AP2M1, FKBP11, GALNT6,*

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

BDNF, COL9A3 e *NR4A1*. A continuación, validouse o estado de metilación dos seis xenes seleccionados mediante pirosecuenciación nas nove liñas celulares, de xeito que en todos os xenes analizados puidemos confirmar a esperada presenza de hipermetilación en CTC-MCC-41 con respecto ás liñas celulares de CTCs resistentes. Entre os seis xenes seleccionados, *A2PM1* foi o xene que mostrou as maiores diferenzas de metilación entre a liña celular de pre-tratamento e as liñas de CTCs resistentes, polo que exploramos con máis detalle a asociación entre a súa metilación e os seus niveis transcripcionais, que foron analizados mediante RT-qPCR en todas as liñas celulares. Esta análise revelou unha asociación negativa entre os niveis de metilación do promotor *A2PM1* e os seus niveis de transcrición. Polo tanto, foi posible confirmar que a liña de pretratamento CTC-MCC-41 presenta un perfil de metilación global diferente ás liñas de CTCs resistentes e que en *AP2M1* existe unha correlación entre a metilación e a expresión en todas as liñas de CTCs analizadas. Tanto *AP2M1* como o resto dos candidatos (*FKBP11, GALNT6, BDNF, COL9A3* e *NR4A1*) representan potenciais biomarcadores de resistencia ao tratamento das CTC e dianas terapéuticas para o CCRm.

Finalmente, no terceiro capítulo, illamos CTCs de pacientes con CCRm para estudar o seu metiloma en relación coa resposta á terapia baseada en FOLFOX en monoterapia ou en combinación co anti-VEGF bevacizumab. Para iso, en primeiro lugar, comparamos dous métodos de enriquecemento de CTCs mediante a análise de mostras de sangue periférico dunha parte da nosa cohorte de pacientes con CCRm. Por unha banda, empregamos un sistema baseado en propiedades biolóxicas, o sistema CellSearch, que illa CTCs en función da expresión do marcador epitelial EpCAM. Por outra banda, utilizamos un sistema baseado en propiedades físicas, o sistema Parsortix, que é un sistema microfluídico que illa as CTCs en función do seu tamaño e deformabilidade. A pesar da alta

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

correlación entre ambos métodos, o sistema CellSearch foi o que detectou un maior número de CTCs, polo que resultou ser o sistema empregado en todos os pacientes da nosa cohorte. En consonancia con isto, os nosos resultados indicaron que a presenza de ≥3 CTCs/7,5 mL no sangue periférico antes do tratamento está asociada a un maior número de metástases, á presenza de metástases nos ganglios linfáticos e unha peor resposta ao tratamento. Ademáis, observouse un aumento significativo nos recontos de CTCs co aumento dos niveis sanguíneos de CEA e cfDNA. A análise univarante e multivariante tamén mostrou que a presenza de ≥3 CTCs/7,5 mL de sangue periférico é un factor pronóstico independente para a supervivencia libre de progresión e a supervivencia global. En liña con isto, as curvas de supervivencia de Kaplan-Meier mostraron que os pacientes con CCR avanzado cun número ≥3 CTCs/7,5 ml de sangue periférico tiñan unha supervivencia libre de progresión e unha supervivencia global significativamente menores que os pacientes que tiñan <3 CTCs/7,5 ml de sangue periférico.

As mostras de CTCs enriquecidas mediante o sistema CellSearch presentan, ademáis das células de interese, un gran número de linfocitos. É importante para unha análise posterior máis eficaz e precisa, poder illar estas CTCs libres doutros tipos celulares. Por este motivo, en colaboración co grupo liderado pola doutora Alix-Panabières, o *Laboratoire Cellules Circulantes Rares Humaines* (LCCRH), as mostras enriquecidas co sistema CellSearch foron sometidas a un segundo proceso de illamento co sistema DEPArray. Esta tecnoloxía, baseada nas propiedades dielectroforéticas das células, permitiu recuperar as CTCs de forma pura para, posteriormente, ser tratadas para illar o ADN e modificalo con bisulfito de sodio. Posteriormente, este ADN bisulfitado utilizouse para preparar librerías para a análise de metiloma completo mediante secuenciación masiva. Cos resultados obtidos tras a secuenciación realizáronse tres comparacións para caracterizar o

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

metiloma destas CTCs en relación coa resposta e resistencia ao tratamento. Por unha banda, comparouse o perfil de metilación das CTCs antes do inicio da terapia dun paciente respondedor co dun non respondedor ao tratamento con FOLFOX en monoterapia. Por outra banda, o perfil de metilación das CTCs antes do inicio da terapia dun paciente que respondiu foi comparado co dun que non respondiu ao tratamento con FOLFOX e bevacizumab. E, finalmente, nun mesmo paciente, avaliouse o metiloma das CTCs antes do inicio da terapia combinada con FOLFOX e bevacizumab con respecto ao metiloma das CTCs obtido despois do inicio da terapia e no momento da progresión. Estes resultados revelaron que o perfil de metilación do ADN das CTC a basal podería ser útil para a selección de pacientes que poderían beneficiarse do tratamento con FOLFOX en monoterapia ou en combinación con bevacizumab. Ademáis, a presión selectiva do tratamento con FOLFOX e bevacizumab indicou cambios no perfil de metilación das CTCs asociados ao desenvolvemento da resistencia á terapia e, tamén se detectaron cambios relevantes na metilación do ADN mitocondrial asociados á resposta e á resistencia ao tratamento.

Os resultados obtidos nesta tese doutoral axudan, polo tanto, a revelar nova información da bioloxía das CTCs, así coma a revelar xenes aberrantemente metilados no seu ADN relacionados coa formación de metástases, e identificar novas dianas terapéuticas, mecanismos de resistencia e biomarcadores potenciais que poderían beneficiar aos pacientes de CCRm na práctica clínica.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

RESUMEN *IN EXTENSO*

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

RESUMEN *IN EXTENSO*

El cáncer colorrectal (CCR) es el tercer cáncer más comúnmente diagnosticado y la segunda causa principal de muertes relacionadas con el cáncer en el mundo. Además, se estima que tanto las tasas de incidencia como de mortalidad seguirán en aumento los próximos años. La mayoría de los CCRs se desarrollan generalmente a partir de pólipos, que son protuberancias en la luz del colon debido a crecimientos localizados o agregaciones de células anormales dentro de la mucosa del colon. Sin embargo, solo un pequeño porcentaje de pólipos adquieren características malignas y la progresión a cáncer resulta de diferentes cambios genéticos y epigenéticos que se van acumulando con el tiempo y puede tardar incluso hasta una década en completarse. La clasificación del CCR se puede realizar atendiendo a criterios muy diversos como pueden ser los factores que lo originan, la localización del tumor, la histología, las vías moleculares, el estado mutacional, el origen celular o la expresión génica. Sin embargo, el CCR es una enfermedad muy heterogénea, lo que dificulta su clasificación de forma relevante para la clínica.

La terapia sistémica para el CCR avanzado y metastásico se basa en el uso de quimioterapia, terapias dirigidas e inmunoterapia, incluyendo diversos fármacos ya sea en combinación o como agentes únicos. Para la mayoría de los pacientes con CCR metastásico (CCRm), el tratamiento consiste en la administración inicial de una fluoropirimidina (*e.g.* 5-Fluorouracilo) en combinación generalmente con oxaliplatino (*e.g.* FOLFOX) o irinotecán (*e.g.* FOLFIRI) para mejorar las tasas de respuesta, que pueden considerarse opciones estándar para el tratamiento de primera línea. Es bastante frecuente la adición a estos regímenes de terapias dirigidas anti-EGFR (*e.g.* cetuximab) o anti-VEGF (*e.g.* bevacizumab), que han demostrado en algunos casos mejorar los resultados. Sin embargo, a pesar de los

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

avances terapéuticos en el tratamiento del CCR, la resistencia a los fármacos, ya sea primaria o adquirida, es un obstáculo importante en la mejora de las tasas de supervivencia de los pacientes con enfermedad avanzada.

En este sentido, la enfermedad metastásica sigue siendo un desafío y es responsable de más del 90% de las muertes por cáncer. La metástasis, que es la última etapa de la progresión del cáncer, es una sucesión de procesos biológicos mediante los cuales las células del tumor primario se diseminan y colonizan un órgano distante. Una vez que las células tumorales están en la luz de los vasos sanguíneos o linfáticos, pueden diseminarse a través de la circulación conociéndose como células tumorales circulantes (CTCs). Estas CTCs, derivadas de un tumor primario o metastásico, son las responsables del desarrollo de las metástasis. Su estudio fue lo que inicialmente acuñó el término biopsia líquida, término que hoy en día también está referido al análisis de otros componentes derivados de tumores como son los ácidos nucleicos circulantes libres, plaquetas educadas por tumores o vesículas extracelulares como los exosomas, aisladas principalmente de la sangre, aunque también de diferentes fluidos biológicos.

La biopsia líquida es una herramienta con gran potencial para la mejora del manejo clínico de los pacientes, ya que supera las limitaciones de la biopsia de tejido, que es el estándar para la caracterización tumoral. La biopsia líquida es una técnica mínimamente invasiva que permite cuestionar de forma repetida y en tiempo real la evolución dinámica de los tumores, reflejando la heterogeneidad molecular de las diferentes lesiones tumorales coexistentes en un paciente. Así, podría ayudar con el diagnóstico temprano, la predicción del pronóstico, la evaluación y el seguimiento de la respuesta a la terapia y la estratificación de los pacientes en función de la detección de dianas terapéuticas o mecanismos de resistencia. Sin embargo, a pesar de sus numerosas ventajas, la detección de CTCs en sangre periférica de pacientes con cáncer representa un desafío técnico, ya que son células

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

extremadamente poco frecuentes en la sangre y muy heterogéneas. Las tecnologías para el aislamiento de CTCs deben de tener una alta sensibilidad y especificidad analíticas, y generalmente se basan en características que diferencian a las CTC de las células sanguíneas. Pueden basarse en características biológicas, como la viabilidad o la expresión de proteínas en la superficie celular, o bien en características físicas como el tamaño, la deformabilidad y las cargas eléctricas, o una combinación de ambas.

El estudio de estas CTCs podría ayudar a comprender mejor la biología del proceso metastásico y, tanto la enumeración como la caracterización molecular, podría implementarse en la clínica como una herramienta valiosa con significado pronóstico, para el seguimiento en tiempo real del tratamiento o la identificación de biomarcadores, dianas y mecanismos de resistencia. Actualmente, el único uso clínicamente validado de las CTCs se basa en su enumeración en el contexto de la enfermedad metastásica. En este sentido, el CellSearch es la única plataforma aprobada por la *Food and Drug Administration* (FDA) de Estados Unidos validada clínicamente para la detección de CTCs en diferentes tipos de cáncer, incluido el CCR. No obstante, más allá de su enumeración, la caracterización molecular de las CTCs también ha proporcionado información valiosa no solo para comprender el proceso metastásico sino también para su traslado a la práctica clínica. Sin embargo, el estudio de la epigenética en CTCs en pacientes con CCR sigue siendo muy limitado.

La epigenética se define como el estudio de los cambios en la función de los genes que son mitóticamente y/o meioticamente hereditarios y que no implican cambios en la secuencia del ADN. Entre sus características es de gran relevancia que las marcas epigenéticas son reversibles y pueden ser moduladas por factores ambientales. La marca epigenética más ampliamente estudiada en el genoma humano es la metilación del ADN. Esta marca consiste en la modificación covalente del ADN mediante la incorporación de

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

un grupo metilo (CH_3) al carbono 5′ de las citosinas en los dinucleótidos de citosina-fosfato-guanina (CpG) para producir 5-metilcitosina (5mC). Es esencial para la regulación de la expresión génica en el desarrollo y diferenciación celular normal. La metilación de las citosinas ocurre casi principalmente en el contexto de los dinucleótidos CpG, que tienden a agruparse en regiones llamadas islas CpG (CpGI). Generalmente, la metilación de CpGI en el promotor de un gen está inversamente asociada con su transcripción, asociando la hipermetilación con el silenciamiento génico y la hipometilación con la expresión génica. Además de en las CpGIs, la metilación del ADN también se ha descrito en otras regiones que, de acuerdo con su distancia de los CpGI, se denominan CpGI *shores*, CpGI *shelves* y *open sea*. Las alteraciones en los patrones de metilación del ADN se han relacionado con un amplio número de enfermedades humanas, incluyendo el cáncer. De forma general, se producen dos tipos de cambios en las células tumorales con respecto a las células normales que son, una hipometilación generalizada del genoma y una hipermetilación de sitios específicos en CpGIs de los promotores de los genes. La metilación del ADN es un biomarcador prometedor debido a su aparición temprana y frecuente en cáncer, su detección relativamente fácil mediante técnicas bien establecidas, su estabilidad, la presencia en una amplia gama de fluidos corporales y su especificidad de tejido.

A la luz de todo lo expuesto anteriormente, el objetivo de esta tesis fue la caracterización epigenómica del perfil de metilación del ADN de las CTCs en CCRm, con el fin de aportar nueva información sobre la biología de estas células y el desarrollo de metástasis; así como revelar nuevas dianas terapéuticas, mecanismos de resistencia y biomarcadores potenciales que podrían beneficiar a los pacientes con CCRm en la práctica clínica.

Para ello, en el primer capítulo analizamos el metiloma del ADN de la línea celular de CTC CTC-MCC-41 en comparación con líneas celulares de

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

cáncer colorrectal primario (HT29, Caco2, HCT116, RKO) y metastásico (SW620 y COLO205). Las líneas celulares de cáncer se utilizan ampliamente para estudiar los mecanismos moleculares de la biología de las células tumorales y para identificar nuevos biomarcadores, dianas terapéuticas y mecanismos de resistencia. En concreto, esta línea CTC-MCC-41 fue establecida por Cayrefourcq L. *et al*. a partir de un paciente con CCRm, siendo la primera línea celular de CTCs de cáncer colorrectal. Esta línea celular de CTCs, así como el resto de líneas celulares primarias y metastásicas, se pusieron en cultivo y se recogieron para la extracción de ADN, que fue posteriormente tratado con bisulfito de sodio, una técnica ampliamente usada para el estudio de la metilación del ADN que convierte las citosinas no metiladas en uracilo. El ADN bisulfitado de las líneas se hibridó en un Infinium MethylationEPIC BeadChips, un *array* que permite interrogar el estado de metilación de más de 850.000 sitios CpG en todo el genoma humano con una resolución de un solo nucleótido. Los resultados obtenidos en el *array* revelaron un perfil de metilación diferente entre CTC-MCC-41 y HT29. Tanto las CpG hipometiladas como las hipermetiladas de CTC-MCC-41 se distribuyeron de manera similar según el contexto de CpG y la región génica y, en un porcentaje muy pequeño, también se distribuyeron en los sitios de unión del factor de transcripción (TFBS) o regiones *enhancer*. Al centrar el análisis en CpGI y *shore* de los promotores de los genes, que son regiones en las que la desregulación epigenética se ha descrito como una característica muy relevante en cáncer, también se identificó un perfil de metilación capaz de diferenciar claramente CTC-MCC-41 de las células de tumor primario HT29. Además, el análisis de vías Panther realizado con los genes diferencialmente metilados de estas CpG en los promotores, reveló un enriquecimiento de diversos procesos biológicos en el desarrollo del cáncer y la metástasis como son la regulación de la transcripción, adhesión y diferenciación celular, apoptosis, procesos metabólicos y reparación del

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ADN, entre otros; y diferentes vías como son la señalización de Wnt, la vía de la cadherina, inflamación, angiogénesis y la señalización de las integrinas.

El perfil de metilación de la línea CTC-MCC-41, se evaluó también con respecto a la línea celular de cáncer colorrectal metastásico COLO205. En este caso, también se identificó un perfil de metilación que diferencia claramente ambas líneas celulares. La mayoría de las CpGs diferencialmente metiladas encontradas en CTC-MCC-41 con respecto a COLO205 estaban hipometiladas, y mostraban también una distribución similar según el contexto CpG y la región génica, y un pequeño porcentaje de estas se localizó en TFBS y regiones *enhancer*. Del mismo modo, cuando el análisis se centró en las CpGI y *shore* de los promotores de los genes, CTC-MCC-41 también mostró un perfil de metilación diferente con respecto a la línea tumoral COLO205. El análisis mediante ontología genética reveló el enriquecimiento de procesos biológicos clave, como la regulación de la transcripción, la adhesión celular y diferenciación, apoptosis, ciclo celular y proliferación, procesos metabólicos y de reparación del ADN. Además, el análisis de vías Panther reveló un enriquecimiento de vías importantes relacionadas con el desarrollo del cáncer y la metástasis, como la señalización de Wnt, la vía de la cadherina, inflamación, angiogénesis y señalización de integrinas, entre otras. En línea con esto, para confirmar la existencia de una firma de metilación específica de las CTCs en el CCR, se incluyó en el análisis otras líneas celulares de CCR que representan tumores primarios (Caco2, HCT116, RKO) y metastásicos distantes (SW620). Así, obtuvimos un perfil de metilación específico que permitió clasificar las células de CCR según su tipo celular: CTC, línea celular de tumor primario o metastásico.

Los datos de expresión de la línea celular CTC-MCC-41 de CTCs en comparación con la línea celular de tumor primario HT29, habían sido publicados previamente en un trabajo de Cayrefourq *et al.* Estos datos se utilizaron para determinar si los cambios de metilación del ADN en CpGI y

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

shore de los promotores de los genes podían regular la expresión génica de esta línea celular de CTCs. En este sentido, centramos nuestro análisis en aquellos genes cuya expresión ya había sido validada (expresión disminuida en CTC-MCC-41: *TGFB2, DKK1, GJB6, PTGS2, PDGFC, SMARCA1* y *GATA2*; *sobreexpresados* en CTC-MCC-41: *BMP7, BCL11A, SEMA6A, FN1, ABCB1, CCND2* y *GAL*). Los valores de metilación obtenidos en el *array* para estos genes fueron validados mediante pirosecuenciación o PCR cuantitativa específica de metilación (qMSP), obteniendo en todos los genes, a excepción de *GATA2*, los valores de metilación esperados. Además, para determinar la correlación entre expresión y metilación, la línea celular CTC-MCC-41 se trató con 5-Aza-2′-desoxicitidina, un epifármaco que inhibe la metilación. Después del tratamiento, las células se recogieron para la extracción de ARN, y tras la retrotranscripción, se analizó la expresión de los genes de interés mediante PCR cuantitativa con transcripción inversa (RT-qPCR). Estos resultados mostraron que la expresión de los genes silenciados epigenéticamente *TGFB2, DKK1, GJB6* y *PTGS2* se restauró por completo, evidenciando que el programa epigenético de los genes hipermetilados en las CTC colorrectales se revirtió favoreciendo su expresión. Estos resultados permitieron, por tanto, poner de relieve que la línea de CTCs estudiada tiene un perfil de metilación tanto global como en las CpG de promotores diferente a las líneas celulares de tumor primario y metastásico. Además, los genes metilados de forma aberrante que se identificaron constituyen potenciales biomarcadores y dianas terapéuticas del CCRm.

Siguiendo los análisis con líneas celulares derivadas de CTCs de pacientes con CCR, en el segundo capítulo nos cuestionamos si el perfil de metilación del ADN de las CTCs podría sufrir cambios aosciados con la resistencia a la terapia y la progresión de la enfermedad. Con este fin, estudiamos el metiloma de una colección única de nueve líneas celulares humanas establecidas por Cayrefourcq *et al.* y Soler *et al*., a partir de CTCs

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

de un paciente con mCRC obtenidas a diferentes momentos durante el tratamiento y la progresión de la enfermedad: La línea CTC-MCC-41, que se obtuvo antes del inicio de tratamiento; la línea CTC-MCC-41.4, obtenida tras el tratamiento de primera línea con FOLFIRI combinado con bevacizumab y tras el tratamiento de segunda línea con FOLFOX y bevacizumab; y las líneas CTC-MCC-5A, CTC-MCC-5B, CTC-MCC-5C, CTC-MCC-5D, CTC-MCC-5E, CTC-MCC-5F e CTC- MCC-5G (CTC-MCC-5 [A-G]), que se obtuvieron una semana antes del fallecimiento del paciente. Todas estas líneas de CTCs se pusieron en cultivo y se realizó la extracción del ADN que, tras ser tratado con bisulfito de sodio, se hibridó en Infinium MethylationEPIC BeadChips. Los resultados del *array* revelaron un perfil de metilación diferente de la línea celular pretratamiento CTC-MCC-41 con respecto a todas las líneas celulares de CTCs resistentes al tratamiento. Las CpG diferencialmente metiladas entre esta línea pretratamiento y las líneas celulares resistentes mostraron una amplia distribución a lo largo de todos los cromosomas del genoma, mostrando la línea pretratamiento una mayor metilación que las células resistentes. Además, tanto las CpG hipermetiladas como las hipometiladas de CTC-MCC-41 mostraron una distribución similar según la región génica y el contexto de CpG, y solo un pequeño porcentaje de estas CpGs estaban en TFBS y regiones *enhancer*. Asimismo, el estudio del enriquecimiento de rutas KEGG y el análisis de vías Panther para los genes diferencialmente metilados entre CTC-MCC-41 y las líneas de CTCs resistentes, reveló un predominio de vías relevantes relacionadas con la señalización de PI3K/AKT y MAPK, el metabolismo, la señalización de Wnt, la señalización de cadherinas e integrinas, la inflamación y la apoptosis, entre otras.

Las diferencias en el perfil de metilación entre las células pretratamiento CTC-MCC-41 y el resto de células resistentes (CTC-MCC-41.4 y CTC-MCC-41.5 [A-G]), concuerdan con los datos de expresión previamente publicados para estas líneas en un trabajo de Cayrefourcq *et al.*,

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

en los que ambos grupos muestran un programa de expresión diferente. En base a esto, y con la idea de identificar biomarcadores asociados con la resistencia a la terapia, evaluamos la relación entre la metilación de las islas y regiones *shore* de los promotores entre estos dos grupos de CTCs (pretratamiento y resistentes). De este modo, identificamos seis genes candidatos que presentaban una mayor metilación y menor expresión en CTC-MCC-41 que en las líneas resistentes a tratamiento: *AP2M1, FKBP11, GALNT6, BDNF, COL9A3* y *NR4A1*. A continuación, el estado de metilación de los seis genes seleccionados se validó mediante pirosecuenciación en las nueve líneas celulares, de modo que en todos los genes analizados pudimos confirmar la presencia esperada de hipermetilación en CTC-MCC-41 con respecto a las líneas celulares de CTCs resistentes. De entre los seis genes seleccionados, *A2PM1* fue el gen que mostró mayores diferencias de metilación entre la línea celular pretratamiento y las líneas de CTCs resistentes, por lo que exploramos más en detalle la asociación entre su metilación y sus niveles transcripcionales, que fueron analizados mediante RT-qPCR en todas las líneas celulares. Este análisis reveló una asociación negativa entre los niveles de metilación del promotor *A2PM1* y sus niveles transcripcionales. Se pudo confirmar, por tanto, que la línea pretratamiento CTC-MCC-41 presenta un perfil de metilación global diferente a las líneas de CTCs resistentes y que en *AP2M1* existe una correlación entre la metilación y expresión en todas las líneas de CTCs analizadas. Tanto *AP2M1* como el resto de candidatos (*FKBP11, GALNT6, BDNF, COL9A3* y *NR4A1*) representan potenciales biomarcadores de la resistencia al tratamiento de las CTCs y dianas terapéuticas para el CCRm.

Finalmente, en el tercer capítulo, aislamos CTCs de pacientes con CCRm con el fin de estudiar su metiloma en relación con la respuesta a la terapia basada en FOLFOX en monoterapia o en combinación con el anti-VEGF bevacizumab. Para ello, en primer lugar, comparamos dos métodos de

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

enriquecimiento de CTCs mediante el análisis de las muestras de sangre periférica de una parte de nuestra cohorte de pacientes con CCRm. Por un lado, empleamos un sistema basado en propiedades biológicas, el sistema CellSearch, que aísla las CTCs basándose en expresión del marcador epitelial EpCAM. Por otro lado, empleamos un sistema basado en propiedades físicas, el sistema Parsortix, que es un sistema de microfluídica que aísla las CTCs basándose en su tamaño y deformabilidad. A pesar de la alta correlación entre los dos métodos, el sistema CellSearch fue el que detectó un mayor número de CTCs, por lo que resultó ser el sistema empleado en el total de los pacientes de nuestra cohorte. En línea con esto, nuestros resultados indicaron que la presencia de ≥3 CTCs/7,5 mL en sangre periférica antes del tratamiento se asocia con un mayor número de metástasis, la presencia de metástasis en los ganglios linfáticos y con una peor respuesta al tratamiento. Además, se observó un aumento significativo en los recuentos de CTCs con el aumento de los niveles en sangre de CEA y cfDNA. El análisis univariante y multivariante mostró además que la presencia de ≥3 CTCs/7,5 mL de sangre periférica es un factor pronóstico independiente de la supervivencia libre de progresión y de la supervivencia global. En línea con esto, las curvas de supervivencia de Kaplan-Meier mostraron que los pacientes con CCR avanzado con un número ≥ 3 CTCs/7,5 mL de sangre periférica tenían una supervivencia libre de progresión y supervivencia global significativamente menores que los pacientes que tenían <3 CTCs/7,5 mL de sangre periférica.

Las muestras de CTCs enriquecidas empleando el sistema CellSearch presentan, además de las células de interés, una gran cantidad de linfocitos. Es importante para análisis posteriores efectivos y más precisos, conseguir aislar estas CTCs libres de otros tipos celulares. Por este motivo, en colaboración con el grupo de la Dra. *Alix- Panabières,* el *Laboratoire Cellules Circulantes Rares Humaines* (LCCRH), las muestras enriquecidas

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

mediane el Sistema CellSearch se sometieron a un segundo proceso de aislamiento con el sistema DEPArray. Esta tecnología, basándose en las propiedades dielectroforéticas de las células, nos permitió recuperar las CTCs de manera pura para, posteriormente, ser tratadas para aislar el ADN y modificarlo con bisulfito de sodio. Posteriormente, este ADN bisulfitado se empleó para la preparación de librerías para el análisis de metiloma completo mediante secuenciación masiva. Con los resultados obtenidos tras la secuenciación, se realizaron tres comparativas para caracterizar el metiloma de estas CTCs en relación con la respuesta y resistencia al tratamiento. Por un lado, se comparó el perfil de metilación de las CTCs antes del inicio de la terapia de un paciente respondedor con el de un no respondedor al tratamiento con FOLFOX en monoterapia. Por otro lado, se comparó el perfil de metilación de las CTCs antes del inicio de la terapia de un paciente respondedor con el de un no respondedor al tratamiento con FOLFOX y bevacizumab. Y, finalmente, en un mismo paciente se evaluó el metiloma de las CTCs antes del inicio de la terapia combinada con FOLFOX y bevacizumab con respecto al metiloma de las CTCs obtenidas después del inicio de la terapia y en el momento de la progresión. Estos resultados revelaron que el perfil de metilación del DNA de las CTCs a basal podría ser útil para la selección de pacientes que se podrían beneficiar al tratamiento con FOLFOX en monoterapia y en combinación con bevazizumab. La presión selectiva del tratamiento con FOLFOX y bevacizumab indujo cambios en el perfil de metilación de las CTCs asociado con el desarrollo de resistencia a terapia y además, se detectaron también cambios relevantes en la metilación en el DNA mitocondrial asociados con respuesta y resistencia a tratamiento.

Los resultados obtenidos en esta tesis doctoral ayudan, por lo tanto, a revelar nueva información de la biología de las CTCs y del desarrollo de las metástasis, así como a revelar genes aberrantemente metilados relacionados con la formación de metástasis, e identificar nuevas dianas terapéuticas,

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

mecanismos de resistencia y potenciales biomarcadores que podrían beneficiar a los pacientes con CCRm en la práctica clínica.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

SUMMARY *IN EXTENSO*

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

SUMMARY *IN EXTENSO*

Colorectal cancer (CRC) is the third most frequently diagnosed cancer and the second leading cause of cancer-related deaths in the world. In addition, it is estimated that both incidence and mortality rates will continue to increase in the next years. Most CRCs generally develop from polyps, which are protrusions in the lumen of the colon due to localized growths or aggregations of abnormal cells within the colonic mucosa. However, only a small percentage of polyps acquire malignant characteristics, and the progression to cancer results from different genetic and epigenetic changes that accumulate over time and can take up to a decade to complete. CRC can be classified according to very diverse criteria, such as the originating factors, the location of the tumour, histology, molecular pathways, mutational status, cell origin, or gene expression. However, CRC is a very heterogeneous disease, which makes it difficult to classify it in way that is relevant for the clinic.

Systemic therapy for advanced and metastatic CRC is based on the use of chemotherapy, targeted therapies and immunotherapy, including various drugs either in combination or as single agents. For most patients with metastatic CRC (mCRC), treatment consists of initial administration of a fluoropyrimidine (*e.g.* 5-Fluorouracil) usually in combination with oxaliplatin (*e.g.* FOLFOX) or irinotecan (*e.g.* FOLFIRI) to improve response rates, these can be considered standard options for first-line treatment. It is quite common to add anti-EGFR (*e.g.* cetuximab) or anti-VEGF (*e.g.* bevacizumab) targeted therapies to these regimens, which have been shown to improve results in some cases. However, despite therapeutic advances in the treatment of CRC, drug resistance, whether its primary or acquired, is a major obstacle in improving survival rates for patients with advanced disease.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

In this sense, the metastatic disease remains a challenge and is responsible for more than 90% of cancer deaths. Metastasis, which is the last stage of cancer progression, is a succession of biological processes by which cells from the primary tumour spread and colonize a distant organ. Once the tumour cells are in the lumen of the blood or lymphatic vessels, they can spread through the circulation known as circulating tumour cells (CTCs). These CTCs, derived from a primary or metastatic tumour, are responsible for the development of metastases. Their study was what initially coined the term liquid biopsy, a term that today is also refered to the analysis of other tumour-derived components such as free circulating nucleic acids, tumour-educated platelets or extracellular vesicles such as exosomes, isolated mainly from blood but also from different biological fluids.

Liquid biopsy is a tool with great potential for improving the clinical management of patients, since it overcomes the limitations of tissue biopsy, which is the standard for tumour characterization. Liquid biopsy is a minimally invasive technique that allows to question the dynamic evolution of tumours repeatedly and in real time, reflecting the molecular heterogeneity of the different coexisting tumour lesions in a patient. Thus, it could help with early diagnosis, prognosis prediction, evaluation and monitoring of response to therapy, and patient stratification based on detection of therapeutic targets or resistance mechanisms. However, despite its numerous advantages, the detection of CTCs in the peripheral blood of cancer patients represents a technical challenge, since they are extremely rare cells in blood and very heterogeneous. Technologies for the isolation of CTCs must have high analytical sensitivity and specificity and these are generally based on characteristics that differentiate CTCs from blood cells. They can be based on biological characteristics, such as viability or protein expression on the cell surface, or on physical characteristics such as size, deformability, and dielectrophoretical properties, or a combination of both.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The study of these CTCs could help to better understand the biology of the metastatic process and, both the enumeration and the molecular characterization, could be implemented in the clinic as a valuable tool with prognostic significance, for real-time monitoring of treatment or identification of biomarkers, targets and resistance mechanisms. Currently, the only clinically validated use of CTCs is based on their enumeration in the context of metastatic disease. In this sense, CellSearch is the only platform approved by the United States Food and Drug Administration (FDA) clinically validated for the detection of CTCs in different types of cancer, including CRC. However, beyond their enumeration, the molecular characterization of CTCs has also provided valuable information not only for understanding the metastatic process but also to transfer to the clinical practice. However, the study of epigenetics on CTCs in CRC patients remains very limited.

Epigenetics is defined as the study of changes in gene function that are mitotically and/or meiotically heritable and do not involve changes in the DNA sequence. Among its characteristics, it is of great relevance that epigenetic marks are reversible and can be modulated by environmental factors. The most widely studied epigenetic mark in the human genome is DNA methylation. This mark consists of the covalent modification of DNA by incorporating a methyl group (CH_3) to the 5′ carbon of cytosines in cytosine-phosphate-guanine dinucleotides (CpG) to produce 5-methylcytosine (5mC). It is essential for the regulation of gene expression in normal cell development and differentiation. Cytosine methylation occurs almost primarily in the context of CpG dinucleotides, which tend to cluster in regions called CpG islands (CpGI). Generally, CpGI methylation in the promoter of a gene is inversely associated with its transcription, with hypermethylation associated with gene silencing and hypomethylation with gene expression.

In addition to CpGIs, DNA methylation has also been described in other regions that, according to their distance from CpGIs, are called CpGI

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

shores, CpGI shelves, and open sea. Alterations in DNA methylation patterns have been linked to a wide number of human diseases, including cancer. In general, two types of changes occur in tumour cells with respect to normal cells, which are generalized hypomethylation of the genome and hypermethylation of specific sites in CpGIs of gene promoters. DNA methylation is a promising biomarker due to its early and frequent occurrence in cancer, its relatively easy detection by well-established techniques, its stability, its presence in a wide range of body fluids, and its tissue specificity.

In light of all the above, the objective of this thesis was the epigenomic characterization of the DNA methylation profile of CTCs in mCRC, in order to provide new information on the biology of these cells and the development of metastases; as well as reveal new therapeutic targets, resistance mechanisms and potential biomarkers that could benefit patients with mCRC in clinical practice.

To this end, in the first chapter we analyzed the DNA methylome of the CTC-MCC-41 CTC cell line in comparison with cell lines of primary (HT29, Caco2, HCT116, RKO) and metastatic (SW620 and COLO205) colorectal cancer. Cancer cell lines are widely used to study the molecular mechanisms of tumour cell biology and to identify new biomarkers, therapeutic targets, and resistance mechanisms. Specifically, this CTC-MCC-41 line was established by Cayrefourcq L. *et al.* from a patient with mCRC, being the first cell line of colorectal cancer CTCs. This cell line of CTCs, as well as the rest of the primary and metastatic cell lines, were placed in culture and collected for DNA extraction, which was later treated with sodium bisulfite, a widely used technique for the study of DNA methylation that converts unmethylated cytosines to uracil. Bisulfited DNA from the lines was hybridized on Infinium MethylationEPIC BeadChips, an array that allows interrogation of the methylation status of more than 850,000 CpG sites throughout the human genome with a resolution of a single nucleotide. Results obtained in the array

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

revealed a different methylation profile between CTC-MCC-41 and HT29. Both hypomethylated and hypermethylated CpGs of CTC-MCC-41 were similarly distributed according to CpG context and gene region and, in a very small percentage, were also distributed at transcription factor binding sites (TFBS) and enhancer regions. When focusing the analysis on CpGI and shore of the gene promoters, which are regions in which epigenetic deregulation has been described as a highly relevant characteristic in cancer, a methylation profile capable of clearly differentiating CTC-MCC-41 from HT29 primary tumour cells was also identified. In addition, the analysis of Panther pathways carried out with the differentially methylated genes of these CpGs in the promoters, revealed an enrichment of various biological processes in cancer development and metastasis, such as regulation of transcription, cell adhesion and differentiation, apoptosis, metabolic processes and DNA repair, among others; and different pathways such as: Wnt signaling, the cadherin pathway, inflammation, angiogenesis, and integrins signaling.

The methylation profile of the CTC-MCC-41 cell line was also evaluated with respect to the metastatic colorectal cancer cell line COLO205. In this case, a methylation profile that clearly differentiates both cell lines was also identified. Most of the differentially methylated CpGs found in CTC-MCC-41 with respect to COLO205 were hypomethylated, and also showed a similar distribution according to CpG context and gene region, and a small percentage of these were located in TFBS and enhancer regions. Similarly, when the analysis was focused on the CpGI and shore of gene promoters, CTC-MCC-41 also showed a different methylation profile compared to the COLO205 tumour line. Gene ontology analysis revealed enrichment of key biological processes, such as regulation of transcription, cell adhesion and differentiation, apoptosis, cell cycle and proliferation, metabolic processes, and DNA repair. In addition, the Panther pathway analysis revealed an enrichment of important pathways related to cancer development and

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

metastasis, such as Wnt signaling, the cadherin pathway, inflammation, angiogenesis, and integrin signaling, among others. In line with this, to confirm the existence of a CTC-specific methylation signature in CRC, other CRC cell lines representing primary (Caco2, HCT116, RKO) and distant metastatic tumours (SW620) were included in the analysis. Thus, we obtained a specific methylation profile that allowed CRC cells to be classified according to their cell type: CTC, primary or metastatic tumour cell line.

The expression data of the CTC cell line CTC-MCC-41 in comparison with the primary tumour cell line HT29, had been previously published in a work by Cayrefourq *et al.* These data were used to determine if DNA methylation changes in CpGI and shore regions of gene promoters could regulate gene expression in this CTC cell line. In this sense, we focused our analysis on those genes whose expression had already been validated (downregulated expression in CTC-MCC-41: *TGFB2*, *DKK1*, *GJB6*, *PTGS2*, *PDGFC*, *SMARCA1* and *GATA2*; upregulated in CTC-MCC-41: *BMP7*, *BCL11A*, *SEMA6A*, *FN1*, *ABCB1*, *CCND2* and *GAL*). The methylation values obtained in the array for these genes were validated by pyrosequencing or quantitative methylation-specific PCR (qMSP), obtaining the expected methylation values in all the genes, with the exception of *GATA2*. Furthermore, to determine the correlation between expression and methylation, the CTC-MCC-41 cell line was treated with 5-Aza-2′-deoxycytidine, an epidrug that inhibits methylation. After treatment, cells were harvested for RNA extraction, and after retrotranscription, the expression of the genes of interest was analyzed by quantitative reverse transcription PCR (RT-qPCR). These results showed that the expression of the epigenetically silenced genes *TGFB2*, *DKK1*, *GJB6* and *PTGS2* was completely restored, evidencing that the epigenetic program of the hypermethylated genes in colorectal CTCs was reversed, favoring their expression. These results, therefore, made it possible to highlight that the CTC cell line studied has a

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

methylation profile both globally and in the CpG promoters that is different from that of the primary and metastatic tumour cell lines. In addition, the aberrantly methylated genes that have been identified are potential biomarkers and therapeutic targets for mCRC.

Following the analysis with cell lines derived from CTCs from CRC patients, in the second chapter we question whether the DNA methylation profile of CTCs could undergo changes in relation to resistance to therapy and disease progression. To this end, we studied the methylome of a unique collection of nine human CTC cell lines, established by Cayrefourcq *et al.* and Soler *et al*., from a patient with mCRC at different times during treatment and disease progression: The CTC-MCC-41 line, obtained before start of treatment; the CTC-MCC-41.4 cell line, obtained after first-line treatment with FOLFIRI combined with bevacizumab and after second-line treatment with FOLFOX and bevacizumab; and the cell lines CTC-MCC-5A, CTC-MCC-5B, CTC-MCC-5C, CTC-MCC-5D, CTC-MCC-5E, CTC-MCC-5F e CTC- MCC-5G (CTC-MCC-5 [A-G]), which were obtained one week before the patient's death. All these CTC cell lines were placed in culture and DNA extraction was performed which, after being treated with sodium bisulfite, was hybridized in Infinium Methylation EPIC BeadChips. The array results revealed a different methylation profile of the pre-treatment cell line CTC-MCC-41 with respect to all cell lines of CTCs resistant to treatment. The differentially methylated CpGs between this pretreatment line and the resistant cell lines showed a broad distribution along all the chromosomes of the genome, with the pretreatment line showing higher methylation than the resistant cells. Furthermore, both hypermethylated and hypomethylated CpGs from CTC-MCC-41 showed a similar distribution according to gene region and CpG context, and only a small percentage of these CpGs were in TFBS and enhancer regions. Likewise, the study of the enrichment of KEGG and Panther pathways of the differentially methylated genes between CTC-MCC-41 and resistant CTC cell

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

lines, revealed a predominance of relevant pathways related to PI3K/AKT and MAPK signaling, metabolism, Wnt signaling, cadherins and integrins signaling, inflammation and apoptosis, among others.

The differences in the methylation profile observed between the CTC-MCC-41 pre-treatment cell line and the resistant cell lines (CTC-MCC-41.4 y CTC-MCC-41.5 [A-G]), are in line with the expression data previously published for these lines in a work of Cayrefourcq *et al.*, in which both groups presented a different expression program. Based on this, and with the idea of identifying biomarkers associated with resistance to therapy, we evaluated the relationship between the methylation of the CpGI and shore regions of gene promoters between these two groups of CTCs (pretreatment and resistant). In this way, we identified six candidate genes more hypermethylated and downregulated in CTC-MCC-41 than in cell lines resistant to treatment: *AP2M1*, *FKBP11*, *GALNT6*, *BDNF*, *COL9A3* and *NR4A1*. Next, the methylation status of the six selected genes was validated by pyrosequencing in the nine cell lines, and we were able to confirm the expected presence of hypermethylation in CTC-MCC-41 with respect to the cell lines of resistant CTCs. Among the six selected genes, *A2PM1* was the gene that showed the greatest methylation differences between the pretreatment cell line and the resistant CTCs lines, so we explored in more detail the association between its methylation and its transcriptional levels, which were analyzed using RT-qPCR in all cell lines. This analysis revealed a negative association between the methylation levels of the *A2PM1* promoter and its transcriptional levels. It was therefore possible to confirm that the pre-treatment cell line CTC-MCC-41 presents a global methylation profile different from the resistant CTC cell lines and that in *AP2M1,* there is a correlation between methylation and expression in all the CTC cell lines analyzed. Both *AP2M1* and the rest of the candidates (*FKBP11*, *GALNT6*, *BDNF*, *COL9A3* and *NR4A1*) represent

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

potential biomarkers of resistance to treatment of CTCs and therapeutic targets for mCRC.

Finally, in the third chapter, we isolated CTCs from patients with mCRC in order to study their methylome in relation to the response to FOLFOX-based therapy either in monotherapy or in combination with anti-VEGF treatment (bevacizumab). For this end, we first compared two methods of CTC enrichment by analyzing peripheral blood samples from a part of our cohort of patients with mCRC. On one hand, we used a system based on biological properties, the CellSearch system, which isolates CTCs based on the expression of the epithelial marker EpCAM. On the other hand, we used a system based on physical properties, the Parsortix system, which is a microfluidic system that isolates CTCs based on their size and deformability. Despite the high correlation between the two methods, the CellSearch system was the one that detected the highest number of CTCs, so it was the system selected to be used in all the patients of our cohort. In line with this, our results indicated that the presence of ≥3 CTCs/7.5 mL in peripheral blood before treatment is associated with a higher number of metastases, the presence of lymph node metastases, and with a worse response to treatment. Furthermore, a significant increase in CTC counts was observed with increasing blood levels of CEA and cfDNA. The univariate and multivariate analysis also showed that the presence of ≥3 CTCs/7.5 mL of peripheral blood is an independent prognostic factor for progression-free survival and overall survival. In line with this, the Kaplan-Meier survival curves showed that patients with advanced CRC with a number ≥3 CTCs/7.5 mL of peripheral blood had significantly lower progression-free survival and overall survival than patients who had <3 CTCs/7.5 mL of peripheral blood.

The CTC samples enriched using the CellSearch system present, in addition to the cells of interest, a large number of lymphocytes. It is important for effective and more precise downstream analysis to be able to isolate these

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

CTCs from other cell types. For this reason, in collaboration with the group led by Dr. Alix-Panabières, the *Laboratoire Cellules Circulantes Rares Humaines* (LCCRH), the samples enriched using the CellSearch System were subjected to a second isolation process with the DEPArray system. This technology, based on the dielectrophoretic properties of the cells, allowed us to recover pure CTCs to later be treated for DNA isolation and to modify it with sodium bisulfite. Subsequently, this bisulfited DNA was used to prepare libraries for whole methylome analysis by massive sequencing. With the results obtained after sequencing, three comparisons were made to characterize the methylome of these CTCs in relation to the response and resistance to treatment. On one hand, the methylation profile of the CTCs was compared in a responding and a non-responding patient to treatment before therapy initiation with FOLFOX in monotherapy. On the other hand, the methylation profile of the CTCs of a responding and non-responding patient was compared before the start of therapy with FOLFOX and bevacizumab. And, finally, in the same patient, the methylome of the CTCs was evaluated before the start of combined therapy with FOLFOX and bevacizumab with respect to the methylome of the CTCs obtained after treatment initiation and at the time of progression. These results revealed that selective pressure of treatment with FOLFOX and bevacizumab was able to induce changes in the genomic and mitochondrial DNA methylation profile of CTCs associated with response and resistance to therapy. Results also showed that the DNA methylation profile of the CTCs at baseline may be useful for the selection of patients who could benefit from treatment with FOLFOX in monotherapy and in combination with bevacizumab.

The results obtained in this doctoral thesis help, therefore, to reveal new information on the biology of CTCs and the development of metastases, as well as to reveal aberrantly methylated genes related to metastasis formation, and to identify new therapeutic targets, resistance mechanisms

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

and potential biomarkers that could benefit mCRC patients in the clinical practice.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

INTRODUCTION

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

INTRODUCTION

1. Colorectal cancer

1.1 Epidemiology

Colorectal cancer (CRC) is the third most common cancer type worldwide, and with almost two million cases per year is responsible for a ~10% of all cancer diagnosis. Moreover, it is the second type of cancer when referring to mortality rates, being responsible for a ~9% of all cancer related deaths[1]. Although incidence and mortality rates are similar for both sexes, males show slightly higher rates than females with ~1 million and 800,000 of new cases per year; and ~500,000 and 400,000 deaths respectively[2].

There is a wide geographic variation in the distribution of incidence and mortality of CRC worldwide[3]. Developing countries are showing a rapid tendency increase of these rates, however, the rates are the highest in highly developed countries, in which the trend seems to have stabilized or even decreased[4] mainly due to treatment improvements and screening programs[5]. Despite this, it is estimated for these rates to increase worldwide a 60% and 70% for incidence and mortality respectively by 2040[2].

There are multiple risk factors for CRC development, that can be grouped in modifiable and non-modifiable factors[6]. Some of the modifiable risk factors include alcohol intake, dietary habits like high consumption of red meat, smoking, obesity and sedentarism. Non-modifiable risk factors include age, inherited genetic risk, personal history of Inflammatory bowel disease, family history of CRC and personal or family history of adenomatous polyps[6,7].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1.2 Development

Most colorectal cancers develop through a multistep process involving a series of morphological, histological and genetic changes that get accumulated over time[8]. Generally, colorectal cancer develops from polyps, which are protrusions into the colon lumen due to localized growths or aggregations of abnormal cells within the colon mucosa, although only a small percentage of polyps acquire malignant features[9,10]. Progression from polyp to cancer results from different genetic and epigenetic changes, and it can take even a decade to complete[8].

There are two main types of polyps with malignant potential, adenomas and sessile serrated polyps, which include sessile serrated adenomas, traditional serrated adenomas, and mixed polyps[11]. Adenomas are characterized by dysplasia, which is an abnormal intracellular and intercellular organization and more than 90% of CRCs are adenocarcinomas, following the classic adenoma-carcinoma sequence model of colorectal tumourigenesis[8]. On the other hand, approximately a 10% of all CRCs follow an alternative pathway in which serrated polyps replace the traditional adenoma as the precursor lesion, which is the serrated polyp-carcinoma sequence[11].

Molecularly, CRC can develop though a combination of three different pathways, which are often overlapped in CRC[12]: the chromosomal instability (CIN), the microsatellite instability (MSI) and the CpG island methylator phenotype (CIMP) pathways[12]. The classical adenoma-carcinoma sequence pathway is mainly controlled by the CIN pathway and *KRAS* mutations, whilst in the serrated pathway it is the CIMP pathway and *BRAF* mutations. The MSI pathway can be detected in both the adenoma-carcinoma sequence and the serrated pathway[11].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The CIN pathway accounts for approximately 65-70% of sporadic CRCs and it is characterized by imbalances in chromosome number leading to loss of heterozygosity and aneuploidic tumours[12–14]. CIN pathway cancers accumulate mutations in oncogenes and tumour suppressor genes like *APC, TP53, KRAS, PIK3CA* and *BRAF*, but mutations in *ARID1A, SOX9,* and *AMER1* are also frequent[15].

Typically, mutations in the *APC* gene occur first, affecting chromosome segregation during cell division. This is followed by oncogenic mutations, generally in the *KRAS* oncogene, affecting cell growth, motility, differentiation and survival. This can eventually cause a loss of function of the p53 gene, regulator of transcription and apoptosis, resulting in carcinogenesis[9,15].

The MSI pathway accounts for 15% of CRCs and represents a specific type of genomic instability, caused due to loss of DNA repair mechanisms and characterized by frequent microsatellite mutations[14,15]. Genes frequently mutated in this type of tumours are *PMS1, PMS2, MLH1, MSH2* and *MSH6*[13] or silencing of the *MLH1* promoter by hypermethylation[12]. These cancers tend to arise in proximal colon, tend to exhibit poor differentiation, mucinous cell type, and prominent lymphocytic infiltration[15].

Finally, the CIMP pathway accounts for 17% of CRCs and present a particular type of epigenetic instability[13,15]. It is characterized by hypermethylation of numerous promoter CpG island loci and inactivation of tumour suppressor genes. Mutations in the *BRAF* gene appear to be an early event in this type of tumours[12]. CIMP tumours tend to appear at older ages, in the proximal colon and more commonly in females[12].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1.3 Classification

CRC is a very heterogeneous disease product of the interaction of both environmental and genetic factors, and it is developed due to the accumulation of genetic and epigenetic changes[16]. Classification of CRC can be done according to very diverse criteria such as the originating factors, the tumour location, histology, molecular pathways, mutational status, cell origin or gene expression[14,16,17].

According to the originating factors, most colorectal carcinomas are sporadic (70-80%) but there are other CRCs that arise through inheritance (<1%-5%) and also other CRCs with an associated hereditary component, that are known as familial CRC [16]. When classification is made based on location of the tumour, CRC can be located in four major anatomic colon divisions: the right or ascending colon, the middle or transverse colon, the left or descending colon and the sigmoid colon, although these areas can also be subdivided[18]. The location of the primary tumour has already been described as a prognostic factor in metastatic colorectal cancer (mCRC), being the location of the primary tumour in the right colon associated with a worse prognosis[19,20], although it is the less common location[21]. In relation to histology, the World Health Organization (WHO) has reported a histological classification of colorectal tumours in which several types are considered depending on morphology[2], however, more than the 90% of CRCs are adenocarcinomas, originating from epithelial cells[14], other rare types of colorectal carcinomas are mentioned like the squamous cell or the small cell carcinoma and also several histological variants of adenocarcinoma are listed, such as mucinous adenocarcinoma, signet-ring cell, adenosquamous or medullary carcinomas[2]. However, these histopathological criteria have very modest prognostic value and little clinical significance[14,18].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

When classifying CRC according to molecular pathways, as it was previously mentioned, three groups can be defined: the CIN, the MSI and the CIMP[13,14,22]. According to mutational status, CRC could be also classified by the presence of several mutations in oncogenes or tumour suppressor genes, including *APC, KRAS, BRAF, NRAS, TP53, PI3KCA* and *CTNNB1*[14]. In relation to cell origin, five subtypes of CRC can be described whether if it is originated in stem-like cells, inflammatory cells, transit-amplifying cells, goblet-like cells or enterocytes[23] and in relation to gene expression, CRC can be grouped in Colon Cancer Subtype 1, 2 and 3 (CCS1, CCS2 and CCS3) that correlate with MSI, CIN and partially with CIMP respectively[24].

Along time, several other attempts to classify CRC have been made[25] with the expectation to be consistent in the clinic. However, as it was previously mentioned, CRC is a very heterogeneous disease, which makes it difficult to classify it in a way that is relevant for the clinic[26]. Recently, an international consortium, the CRC Subtyping Consortium (CRCSC), was formed to classify CRC with a translational view to the clinic and elaborate a robust classification system for CRC, describing four consensus molecular subtypes (CMS): CMS1 or MSI Immune subtype (14%), characterized for microsatellite instability, strong immune activation and hypermutated features; CMS2 or Canonical subtype (37%) characterized for epithelial features with marked Wnt and Myc signaling activation, chromosomal instability; CMS3 or Metabolic subtype (13%) characterized for epithelial and metabolic dysregulation and CMS4 or Mesenchymal subtype (23%) characterized for epithelial-mesenchymal transition (EMT) characteristics, stromal invasion, angiogenesis and marked transforming growth factor β-activation[27,28]. Some samples with mixture features (13%) are believed to have an intermediate phenotype or intra-tumoural heterogeneity[27].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1.4 Staging

Although there are several systems used for CRC staging, like the Duke´s, the Jass´s or the Astle-Coller systems[29,30], the most widely used for cancer staging worldwide and to determine the extent of the disease is the Tumour, Node and Metastasis (TNM) staging system [31]. This system was first developed between 1940-1950s by Pierre Denoix and then developed by the American Joint Committee on Cancer (AJCC) together with the Union for International Cancer Control (UICC)[32], and it is continuously being revised to incorporate new evidences for cancer management.

The TNM staging system classifies the disease according to: The size of the primary tumour and invasion of adjacent tissues (T) with values that range from Tx when the main tumour cannot be assessed, Tis when a carcinoma in situ is identified or T0 to T4, referring to progressive size or extent of the tumour; The affection of regional lymph nodes (N) with values that range from Nx when cancer in lymph nodes cannot be assessed or N0 to N3 indicating the progressive degree of spread in lymph nodes; and the presence of distant metastases (M) with values that range from Mx when metastasis cannot be assessed, M0 when there are no distant metastases or M1 when the tumour has spread. Other parameters like the grade of differentiation of cancer cells can be used to assess the cancer stage[33].

The information determining the extent of the disease by the TNM system, can be used for cancer stage grouping between Stage 0, indicating carcinoma in situ, and Stages I to IV associated with progressive disease severity and decreased survival rates[33].

1.5 Treatment

In spite of the therapeutic advances in CRC treatment, metastasis continues to be a challenge and prognosis of mCRC patients is still poor[34].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The best treatment strategy for mCRC is selected for each patient according to both tumour and disease characteristics like clinical presentation, tumour biology, metastasis location, dynamics of progression, symptoms and prognostic markers but also according to patient related factors like comorbidity and patients´ expectations[35].

Systemic therapy for advanced and metastatic CRC, is based in the use of chemotherapy, targeted therapies and immunotherapy, including diverse active drugs either in combination or as single agents[35,36]. These agents can be classified as cytotoxic chemotherapy: Fluorouracil (5-FU), capecitabine, oxaliplatin, irinotecan and trifluridine-tipiracil; targeted therapy directed against the epidermal growth factor receptor (anti-EGFR) (cetuximab, panitumumab), or directed against the vascular endothelial growth factor/vascular endothelial growth factor receptor (anti-VEGF/VEGFR) (bevacizumab, zivaflibercept, ramucirumab, regorafenib) and immunotherapy based against programmed cell death-1 receptor/programmed cell death ligand 1 (anti-PD1/anti-PD-L1), such as pembrolizumab or nivolumab[36].

In relation with chemotherapy, fluoropyrimidines 5-FU and capecitabine are the standard of most treatments for mCRC[36,37]. Antitumour activity of these fluoropyrimidines is determined by inhibiting the activity of the enzyme thymidylate synthase, necessary for pyrimidine nucleotide synthesis, inhibiting the incorporation of nucleotides into both RNA and DNA[38,39]. Irinotecan is converted to an active metabolite, which binds to the enzyme topoisomerase I, necessary for DNA replication and transcription, causing DNA fragmentation and tumour cell death[36]. Oxaliplatin is a platinum analog that acts as an alkylating agent forming cross-links between DNA strands, causing DNA damage and tumour cell death and in CRC it is usually used with a fluoropyrimidine[36]. As it was previously mentioned, these agents can be given in combination, and a current standard is oxaliplatin given

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

together with 5-FU (FOLFOX)[40] and 5-FU in combination with irinotecan (FOLFIRI), which has shown improved response rates[34].

In relation with targeted therapy, molecular targets include EGFR, overexpressed in majority of CRC, and VEGF/VEGFR[36]. The epidermal growth factor (EGF) ligand activates the EGFR, activating downstream cell signaling pathways, like the RAS/RAF/MAPK pathway[34] involved in cell proliferation, migration, resistance to apoptosis and angiogenesis[36]. Cetuximab an panitumumab are monoclonal antibodies (mAbs) that target EGFR, blocking the binding of EGF inhibiting downstream pathways[36]. The VEGF is an angiogenic factor, produced by many tumour and stromal cells, that promotes endothelial cell proliferation, and migration, and increases vessel permeability[41]. VEGF binds to VEGFRs on the vascular endothelial cells and some cancer cell, activating intracellular signaling pathways that promote angiogenesis[36]. The agents bevacizumab, which is a humanized mAb, and zivaflibercept act preventing the stimulation of angiogenesis by its binding to VEGF. Also, ramucirumab, is a mAb that prevents angiogenesis by binding to VEGFR[36].

In the case of immunotherapy in mCRC, benefits are still very restricted[40] although agents like pembrolizumab and nivolumab, both PD-1 blocking antibodies, have being approved by the U.S Food and Drug Administration (FDA) to use in patients with certain types of MSI tumours[36]. Based on the consensus molecular subtypes of CRC, these are expected to have different responses to immunotherapy. In the case of CMS1 and CMS4, which have high levels of CD8+ and CD68+ macrophage infiltration these could be targeted by immunological checkpoints like PD-1[42].

1.5.1 Treatment response and resistance

Drug resistance is still a major obstacle in the improvement of survival rates of patients with advanced CRC. Some patients fail to respond

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

to anticancer drugs as a result of either pre-existing resistance in the patient (intrinsic or primary resistance), or of resistance induced by treatment pressure (acquired resistance)[43,44]. Mechanisms of resistance can exist either for traditional cytotoxic therapy or for targeted therapy and understanding them is key to develop novel and more effective treatments to help guide future therapies[45,46].

In the case of cytotoxic therapy, resistance is generally accomplished by decreased drug delivery to the cancer cells, either by decreased or increased efflux by ATP-dependent transporters or by changes in enzymes involved in metabolism, however, resistance can also appear due to genetic or epigenetic modifications that alter drug sensitivity[46,47].

Resistance to 5-FU is closely related to the expression of thymidylate synthase, since the primary active metabolite of these drug inhibits this enzyme, key in synthesis of thymidine[46]. Since thymidylate synthase is primary target of 5-FU, there is an inverse association between low tumour thymidylate synthase expression and improved response rates to this drug[48,49]. Moreover, 5-FU is converted to other active metabolites through different enzymatic pathways involving the action of different enzymes like the thymidine phosphorylase, uridine phosphorylase, orotate phosphoribosyl transferase and dihydropyrimidine dehydrogenase, all involved in the metabolism and degradation of 5-FU. Several studies have described variable number of tandem repeats and single nucleotide polymorphisms of the genes encoding these enzymes contributing to resistance to this drug[46,48]. In the case of the enzyme dihydropyrimidine dehydrogenase, it contributes to the degradation of 5-FU and its overexpression has been associated with 5-FU resistance [45]. Inhibition of the activity of these enzymes could allow enhanced sensitivity to 5-FU[45,46,50,51].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

In the case of capecitabine, many of the mechanisms of resistance are shared with 5-FU, since capecitabine is an oral drug that through different enzymatic reactions is converted into 5-FU. In particular, thymidine phosphorylase, a critical enzyme for the conversion of capecitabine to 5-FU, plays a key role in its resistance and patients with higher expression levels of thymidine phosphorylase, have better responses to capecitabine and loss of function confers the resistance[45].

Irinotecan is converted intracellularly to its active metabolite, SN-38 (7-ethyl-10-hydroxycamptothecin), to exert its anticancer effect[46,52]. This active metabolite forms a topoisomerase-inhibitor-DNA complex affecting the DNA function. Therefore, the higher the concentration of topoisomerase I, the more sensitive cells become to the drug[45,53]. Also determining drug resistance to irinotecan, we can find uridine diphosphate glucuronosyltransferase, carboxylesterases, hepatic cytochrome P-450 enzymes, ATP-binding cassette transporter protein and β-glucuronidase, all involved in the uptake and metabolism of irinotecan[45,54]. In relation with epigenetic changes, changes in histone acetylation, such as H4K16 acetylation, is associated with resistance to irinotecan[45].

In the case of oxaliplatinum, resistance is related to the nucleotide excision repair (NER) pathway. Gene expression levels of *ERCC1, XRCC1* and *XDP* are correlated with resistance to oxaliplatin, and can be used together as a drug sensitivity predictor index[45,55]. Besides NER, the WBSCR22 protein is quite a novel oxaliplatin resistance biomarker[56]. Transforming growth factor-β1 (TGF-β1), secreted by a wide range of cells in the tumour microenvironment, is thought to help induction of resistance to oxaliplatin via EMT[57]. Thus, interfering with TGF-β1 could potentially sensitize tumour cells to this drug[45,57].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Besides traditional chemotherapy, targeted agents are also subject of resistances, and these occur by different mechanisms including alterations of the target itself, bypass mechanisms, upregulation and activation of downstream effectors, or crosstalk between related pathways which activate cell survival and growth pathways[46,58,59].

Benefit of anti-VEGF therapy in mCRC patients is limited due to acquired resistance. During treatment with the anti-angiogenic drug bevacizumab, there is a decrease in VEGF-A and an increase in VEGFR1 contributing to drug resistance[60]. Moreover, during acquired resistance, hepatocyte growth factor (HGF) levels decrease, suggesting the potential of HGF-ligand inhibition as alternative strategy[60]. It has also been proposed the transcription factor HOXB9 as one of the key mechanisms of anti-VEGF resistance, however the major mechanism of resistance to anti-VEGF therapy is not totally elucidated[45,61].

The anti-EGFR antibodies cetuximab and panitumumab, inhibit downstream signaling pathways, inhibiting proliferation and induction of apoptosis. *KRAS* or *NRAS* mutations and also the *BRAF* V600E mutation are well-known predictors of resistance to anti-EGFR therapy since they lead to constitutive activation of RAS/RAF/MAPK pathway[36,62]. Patients *KRAS* mutated do not benefit from combined therapy, however, these agents do improve response in patients with wild-type *KRAS*[63,64]. In particular, in the case of patients with a G13D *KRAS* mutation in exon 2 codon 13, they do benefit from the use of cetuximab[65]. Also, several studies address that *PTEN* and *PIK3CA* have predictive value for anti-EGFR treatment efficacy[66,67]. It is therefore supported *KRAS* and *BRAF* mutation assessment for mCRC patients before treatment initiation with anti-EGFR therapy[45].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

2. The metastatic cascade

Metastasis is a succession of biological processes by which cells in the primary tumour site disseminate and colonize at a distant organ[68]. Metastasis is the last stage of cancer progression and the cause, including in CRC patients, of more than 90% of cancer deaths[69].

The metastatic cascade (**Figures 1-3**) can be mainly divided in three main events: Tumour cells have to (1) invade the local tissue surrounding the primary tumour site, (2) intravasate to the lymph or the bloodstream, disseminate and arrest in the circulation, extravasating into a tissue and (3) colonize a new site to form a metastatic tumour site[70,71].

2.1 Local invasion

Local invasion of the surrounding tumour-associated stroma by the primary tumour cells, is the step that initiates the metastatic process (**Figure 1**)[70,72]. Tumour cells can invade either as individual cells with amoeboid or mesenchymal type movement or as clusters, sheets or strands using collective migration[73]. During invasion, tumour cells undergo EMT, the dominant program in colon cancer, in which epithelial carcinoma cells undergo changes to highly motile and invasive mesenquimal cells[71]. During EMT, cell junctions to other cells and also to the extracellular matrix (ECM) are destructed, promoting cell migration[74]. There is also active proteolysis, effected principally by matrix metalloproteinases (MMPs) which degrade the ECM and cleave E-cadherin[72]. Moreover, actin cytoskeleton remodeling occurs, forming cellular protrusions that will play a role in cell invasion and ECM degradation[74].

Diverse transcription factors such as SNAIL, TWIST and ZEB, among others, are upregulated in colon cancer cells undergoing

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

metastasis[75] and will act suppressing the expression of epithelial markers such as E-cadherins, cytokeratins and claudins and inducing the expression of mesenquimal markers such as N-cadherin, fibronectin and vimentin[72,74]. These transcription factors are under DNA methylation, histone modifications, and RNA-mediated epigenetic regulation[74].

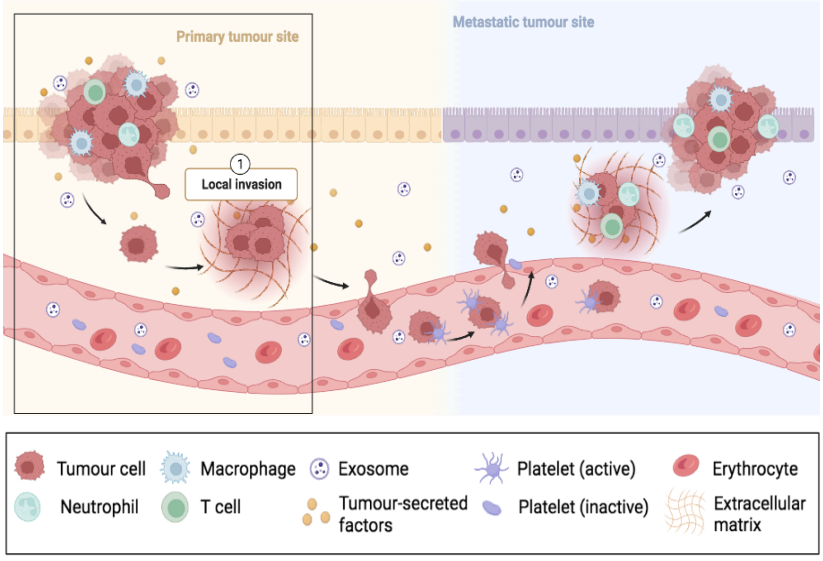


Figure 1. Overview of the initial step of the metastatic cascade. Primary tumour cells, undergoing EMT, locally invade the surrounding tumour-associated stroma. Created with Biorender.com

2.2 Intravasation, dissemination and extravasation

Once tumour cells have dissociated from the primary tumour, they have to intravasate and migrate into the lumina of lymphatic or blood vessels[72]. In this process (**Figure 2**), EMT is also required, since it mobilizes processes like angiogenesis and endothelial integrity disruption[74]. Cancer

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

cells stimulate the formation of new blood vessels, this abnormal neovasculature with high density and hyperpermeability leads to efficient access to blood and lymph[72,74]. Particularly in CRC, Notch signaling plays an essential role in enhancing intravasation and therefore promoting hematogenous metastasis of colon cancer cells[76]. Moreover, Interleukin 6 (IL-6), highly expressed in CRC, promotes the migration and invasion of CRC cells through the IL-6/JAK/STAT signaling[77].

Once tumour cells are in the lumina of blood or lymph vessels, they can disseminate widely through circulation[72]. They are known now as circulating tumour cells (CTCs), cells en route between primary tumour sites and sites of dissemination, acting as metastatic intermediates[70]. To reach new organ sites, these CTCs have to survive in the bloodstream and overcome diverse barriers like the lack of a substratum, velocity-induced forces, and the immune system[70]. When epithelial cells lose cell adhesions, they undergo a special cell death induction called anoikis[78], however, tumour cells are able to evade it due to several factors like the EMT transition, growth factors, oxidative stress, stemness, autophagy, epigenetics and signaling pathways[79]. Moreover, these CTCs have also to overcome mechanical damage caused by shear forces of bloodstream and the immune system, especially natural killer cells. This is done by forming interactions with blood platelets, that will act as a shield [80] and make this platelet coated CTCs more likely to persist in circulation and arrest at distant sites[72]. Arrest of tumour cells can be produced either due to physical constraints or adhere to the layers of a vessel[81]. Once the CTCs reach the secondary site, these exit the vessel and entry the organ parenchyma in a process called extravasation. Extravasation is not always required for metastasis, sometimes CTCs initiate intravascular growth and form a microcolony that disrupts the walls of surrounding vessels, placing tumour cells in contact with tissue parenchyma[72]. For extravasation, cells undergo the opposite

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

phenomenon to intravasation, and cells undergo the mesenchymal-epithelial transition to then proliferate and form a secondary tumour[72].

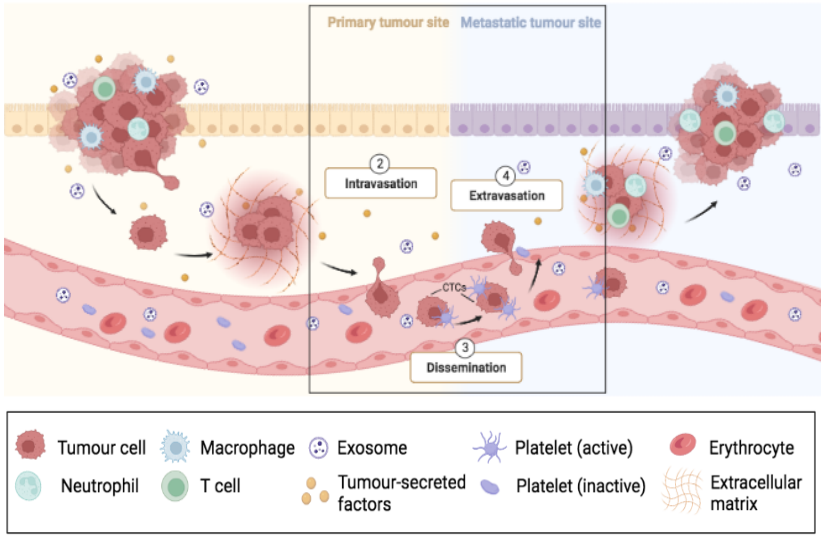


Figure 2. Intermediate steps of the metastatic cascade. Dissociated tumour cells intravasate into the lumina of blood vessels and migrate, as CTCs, to extravasate the vessel and enter another organ. Created with Biorender.com

2.3 Colonization

Survival in the new microenvironment of distant tissues is a key step for extravasated tumour cells to develop metastases[72] (**Figure 3**). In the case of CRC, the liver is the most colonized organ mainly due to its positioning and the blood supply through the major vascularity and microvascularity[82]. Tumour cells arriving to a secondary site need to adapt to an environment that tends to differ from the primary site in its ECM composition, the stroma cells, soluble factors, vascularity or metabolic and nutrient components[72,83]. However, it is believed that environment at target

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

sites is adapted before the arrival of tumour cells to make distant sites more suitable for them, forming the so-called, pre-metastatic niche[72,83]. There are three crucial elements in the establishment of this niche, which are the primary tumour, that will secrete diverse factors to mobilize bone marrow derived cells to the target site, these bone marrow derived cells and the local stroma of the metastatic site[84]. These three elements would work together to make the target site more hospitable through molecular and cellular components like VEGF, TGF-β or cancer-associated fibroblasts[84]. In CRC, colon cancer cells secrete VEGF-A, stimulating CXCL1 expression and recruiting CXCR2 positive Myeloid derived suppressor cells to the premetastatic liver to promote survival and metastasis[82]. Besides this pre metastatic niche formation, tumour cells also need cell-autonomous programs to adapt to the new environment[72]. Sometimes these cells can persist in dormancy as micrometastases, which is a non-proliferative state, paired with the activation of cellular stress programs, allowing them to accumulate alterations that will help with colonization[82].

Survival of tumour cells in the new environment does not guarantee its proliferation to form macroscopic metastases[72], disseminated cells must have high self-renewal capacity. In this sense, it has been proposed that only the tumour initiating cells (TIC) have it, which is limited to a subpopulation present in the tumour[72]. Molecules that promote entrance into the TIC state are EMT-promoting transcription factors like TWIST, SNAIL and ZEB1 or microRNAs involved in regulating EMT[72]. Also, tumour growth must be sustained by the development of new blood vessels to form a macroscopic tumour [85].

The adaptative programs governing metastatic colonization are determined by the tissue of origin of the disseminated cells but also by the identity of the organ site at which metastatic colonization occurs[72]. There are also specific organ environments intrinsically more or less hospitable for

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

survival and proliferation of certain types of disseminated tumour cells[72]. The formation of solid growing macroscopic metastases represents the endpoint of the metastasic cascade [72].

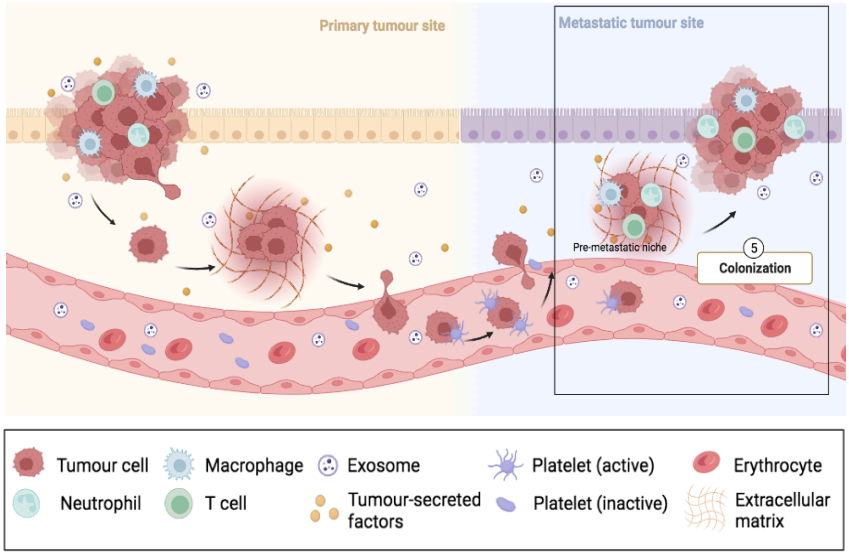


Figure 3. Overview of the final step of the metastatic cascade. Extravasated cells adapt and colonize the new tissue to form macroscopic metastasis. Created with Biorender.com

3. Liquid biopsy

3.1 Concept and characteristics

In oncology, the term liquid biopsy is refered to the analysis of tumour-derived material isolated from biological fluids, mainly blood, but including others such as urine, saliva, ascites or cerebrospinal fluid[86]. Although initially this term was mostly referred to the analysis of CTCs, it is also extended to the analysis of other material like cell-free circulating nucleic acids (DNA, mRNA, non-coding RNA such as micro-RNA or long non-coding RNA),

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

tumour educated platelets or extracellular vesicles like exosomes[87] (**Figure 4**).

The current gold standard for tumour profiling typically involves tissue biopsy, however, it has some limitations since it is not always possible to perform due to inaccessibility to the tumour; it is an invasive technique, which bears the risk of complications; and it doesn't capture tumour heterogeneity and its evolution during treatment[88]. In this sense, liquid biopsy is presented with great potential to overcome these limitations since it is minimally invasive and allows repeated and real time questioning of the dynamic evolution of tumours[89], reflecting also the intratumoural and intermetastatic molecular heterogeneity of the tumours coexisting in a patient.

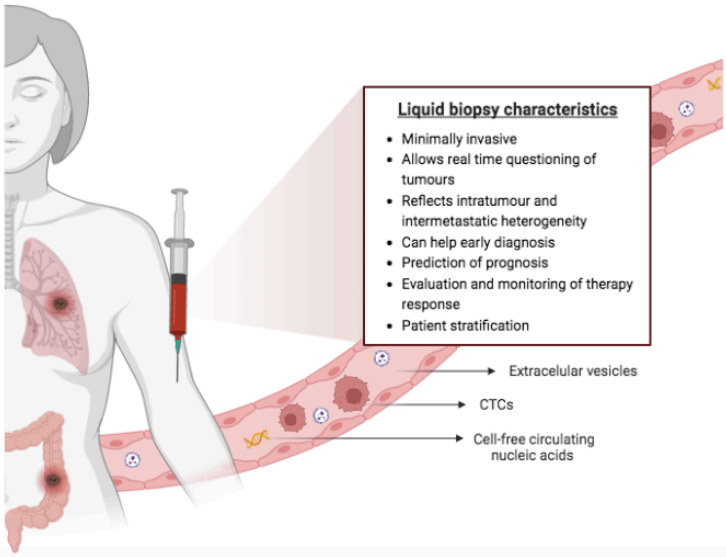


Figure 4. Characteristics and clinical applications of liquid biopsy. Liquid biopsy includes the analysis of tumour-derived material, such as CTCs, extracellular vesicles or cell-free circulating nucleic acids isolated from biological fluids, mainly blood. Created with Biorender.com

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Liquid biopsy has potentially a very relevant role in oncology, since it could help with early diagnosis, prediction of prognosis, evaluating and monitoring systemic therapy response. Moreover, it could also assist with stratification of patients based on the detection of therapeutic targets or resistance mechanisms[90], helping to achieve a precision medicine.

3.2 Circulating tumour cells

The presence of circulating tumour cells in peripheral blood was first described in the mid of the 19th century in patients with cancer after their death[91]. Since that, the field of CTCs has evolved and major advances have been made, mainly due to the development, especially in the last 15 years, of specialized technologies for their capture and enumeration and also for their analysis at a molecular level[92].

CTCs are cells shed from a tumour, primary or metastatic, that migrate through the circulatory system and are responsible for the development of metastasis at distant sites[93] or for returning to the tumour of origin in what is called "tumour self-seeding"[94]. These CTCs are extremely rare and are present at a frequency of 1 CTC per 10^6 –10^7 leukocytes or less when not in an advanced stage[93]. These tumour cells, are characterized as non-leukocytic, nucleated cells, typically epithelial in origin and with larger diameters and membrane surfaces than normal blood cells, however, morphological features may differ depending on the tumour type, stage or timing (pre- or post-treatment) and these can also even lose their epithelial markers in favor of EMT markers[95].

Phenotypically, CTCs can be present in blood as single CTCs or forming clusters. CTC clusters, which are multicellular aggregates, can be heterotypic, when formed with other blood cells like platelets or neutrophils

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

or homotypic, when formed with other CTCs[96]. These CTC clusters appear to be critical mediators in cancer metastasis[97].

The analysis of CTCs is especially interesting since it offers, unlike other liquid biopsy elements, an entire range of analyses at DNA, RNA and protein level and also functional studies, *in vivo* or *in vitro*, on viable cells[98].

3.2.1 CTCs isolation methods

Detection of CTCs in peripheral blood of cancer patients is technically challenging since they are extremely rare in blood. Technologies for CTC isolation need to have high analytical sensitivity and specificity[99] and they are generally based on a combination of positive or negative enrichment strategies to increase the concentration of CTCs, followed by different strategies for their detection, using immunologic, molecular or functional assays[100]. There are several technologies for CTC enrichment and detection, some already validated and others still in development[101]. These technologies rely on characteristics that differentiate CTCs from blood cells, according to biological characteristics, like viability or cell surface protein expression; physical characteristics such as density, size, deformability and electric charges or a combination of both [97,99].

3.2.1.1 Methods based on biological characteristics

Concerning technologies based on biological characteristics, like cell surface protein expression, these rely mostly on immunoaffinity-based separation using either magnetic or microfluidic approaches or a mix of both[102]. These technologies can use positive selection, when using antibodies targeting tumour-associated antigens or negative selection when targeting antigens typically expressed by leukocytes, like CD45[99]. However, the lack of reliable target antigens is a limitation of these technologies, which

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

mostly use the epithelial marker EpCAM, putting aside a subset of EpCAM low/negative CTCs that already underwent EMT and express mostly mesenquimal markers[103].

In the case of immunomagnetic approaches, these methods target a specific CTC antigen using an antibody coupled with a magnetic bead and generally after union, CTCs can be isolated using a magnetic field[99]. A technology that relies on this basis is the CellSearch System (Menarini-Silicon Biosystems, Italy), the only FDA approved platform for CTC detection and the most widely used[101]. In this system, CTC enrichment is performed using immunomagnetic antibodies against EpCAM and identification is performed, after immunofluorescent labeling with antibodies against the cytokeratins (CK) CK8, CK18 and CK19, and against CD45 and DAPI. Results are expressed in number of CTCs per 7.5 ml of whole blood and CTCs are defined as positive for CKs and negative for CD45[101]. Other immunomagnetic assays using anti-EpCAM antibodies are AdnaTest (Quiagen, Germany), MACS Technology (Magnetic cell sorting) (Miltenyi Biotec, Germany) or the CellColector (Gilupi, Germany), that consists of a funcionalized device that is inserted into the venous system capturing CTCs *in vivo* [102].

In the case of immuno-microfluidic approaches, these are based on devices with microestructures coated with antibodies, that generate microvortices to enhance the number of interactions between CTCs and the antibodies of the surface[101]. Technologies using this strategy are the CTC-Chip, with thousands of microposts coated with anti-EpCAM antibodies [104] and OncoCEE chip (Biocept Inc., USA), which uses a cocktail of antibodies that includes anti-EpCAM, tumour associated and mesenquimal markers[102].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.2.1.2 Methods based on physical characteristics

Among the technologies based on physical characteristics, we can find the Parsortix Cell Separation System (Angle, USA), which is a semi-automated microfluidic system capable of capturing and harvesting viable CTCs based on size and deformability. The blood sample runs through a separation cassette, under constant and controlled pressure conditions retaining the CTCs, that can either be stained inside the cassette or harvested for subsequent analysis[105]. Other system based on physical properties is the Vortex Chip, a device with a narrow channel and several reservoirs in which CTCs will remain trapped due to the generation of inertial forces[106]. The main disadvantages that these technologies present are the possibility of losing small CTCs and the retention of leukocytes.

Other methods based on differential cell size include the ISET (Isolation by size of epithelial tumour cells) (Rarecells, France) in which blood is passed through a pored filter via vacuum filtration[102]. Other methods using pored filters are ScreenCell, CellSieve, CellOptics[103] or the FaCTChecker device which, in line with most novel microfiltration approaches, uses 3-dimensional microfilters to improve sample purity and enrichment[102]. These filter-based methods have the loss of small CTCs and membrane clogging as their major disadvantages[103].

Another physical property used by some of these technologies is cell density, used in techniques like Ficoll Paque and OncoQuick. In these techniques, blood is layered on top of a density gradient medium and is centrifuged to separate different cell populations. Erythrocytes and granulocytes pellet at the bottom and mononuclear cells and CTCs, due to their lower density, remain on top on an interphase layer. In the case of OncoQuick, this uses also a porous barrier to avoid layer mixing[102]. Although

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

these methods are easy and inexpensive, they have low specificity and may have inter-layer cross contamination[103].

CTCs can also be isolated based on electrophoretic properties, which depend on differential electrical features of CTCs compared to leukocytes[103]. This is the case of the DEPArray System (Menarini-Silicon Biosystems, Italy), which is an automated instrument that identifies CTCs and recovers them at single-cell level, based on a cartridge with a microelectronic chip, microfluidic chambers and valves. Inside the Main Chamber of the cartridge, in response to the voltage applied, dielectrophoretic cages will be formed, which will be able to trap and move these CTCs[107]. Other device based on electrophoretic properties is the ApoStream platform (ApoCell, USA), which attracts CTCs to a charged electrode on the floor of the device´s chamber, discarding leukocytes and showing high recovery efficiency and recovering CTCs with high viability[102,103].

3.2.1.3 Methods based on biological and physical characteristics

There are different technologies based on a combination of both physical and biological properties. This is the case of the RosetteSep CTC Enrichment Cocktail technology (StemCell Technologies, Canada), an immunodensity methodology that uses antibodies targeting hematopoietic cells, establishing links with leukocites that will form complexes with high density that will be depleted after gradient density centrifugation[102].

3.3 Relevance of the study of CTCs

The study of CTCs could not only help to better understand the biology of the metastatic process but also, both enumeration and molecular characterization, could have translation into the clinic as a valuable tool for

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

prognostic significance, real-time monitoring of therapies or identification of biomarkers and therapeutic targets, including resistance mechanisms[108].

Currently, the only clinically validated use of CTCs is based on their enumeration in the context of the metastatic disease[102]. In this sense, the CellSearch is the only FDA cleared platform with standardization and reproducibility with clinical validation for CTC detection in colorectal, prostate and breast cancers[109]. It has been demonstrated that counts ≥5CTCs/7.5 mL of blood in breast[110] and prostate cancer[111] and counts ≥3CTCs/7.5 mL of blood in CRC[112] are related with poor progression-free (PFS) and overall survival (OS) rates. Several studies have also analyzed the prognostic value of CTC enumeration in other types of tumours like non-small cell lung cancer[113], small-cell lung cancer[114] and pancreatic adenocarcinoma[115] among others.

Although CTC enumeration with prognostic value is only clinically validated for metastatic disease, CTCs are detected at early stages of tumour development[116] being additional prognostic information for patient stratification into low-, intermediate-, or high-risk groups. This could result in better treatment decision and therefore in better patient outcome[102]. It is also important to mention that presence of CTC-clusters have also been related to poor prognosis[117–119] and failure to clear CTCs or increasing counts during treatment is also prognostic factor of worse survival[120]. Since CTCs can be detected at early stages of tumour development, there are several studies using CTCs for early cancer diagnosis[120–122], however, clinical utility of CTCs in relation to cancer detection or early diagnosis is still under debate.

According to monitoring of therapeutic response, there are several studies in breast cancer patients that associate decreased CTC counts after one chemotherapy cycle to have better prognosis than woman with unchanged elevated CTC counts[110]. This was also described for metastatic

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

colon cancer patients, showing improved PFS and OS in those patients with lower CTC counts during chemotherapy[112]. Also, baseline CTC number and changes of CTC numbers after chemotherapy has shown to be of prognostic factor in small cell lung cancer patients[123], in castrate resistant prostate cancer patients treated with docetaxel[124] and in epithelial ovarian cancer patients treated with standard taxol/carboplatin therapy[125].

Besides CTC enumeration, molecular profiling of CTCs has also proven utility for prognosis prediction, patient stratification and disease monitoring[92]. The presence of CTCs HER2 positive in patients with breast cancer was associated with poor PFS[126], and also in breast cancer, CTCs expressing CD47 and PD-L1 showed prognostic value[127]. Moreover, CTCs expressing mesenchymal or stemness-related markers are also associated with poorer survival[128]. Also, in the case of CRC, detection of CTCs with CK20 is an independent negative prognostic marker[129].

In relation with CTCs and patient stratification and therapy response, in metastatic melanoma patients, CTC analysis of several markers by RT-qPCR has been related with differences in treatment response[130]. Also related to prediction of response to chemotherapy is the copy number aberrations profiling of CTCs in small cell lung cancer[131] or the detection of androgen receptor splice variant 7 in CTCs of castration-resistant prostate cancer patients[132]. In CRC, expression of CD44v6 in CTCs seems to be a useful tool for evaluating tumour response to different first-line treatments[133]. In this sense, there are many trials assessing the clinical utility of CTCs, such as the phase III trial VISNU-1, which explores the role of CTC counts as a biomarker to select CRC patients for therapy with FOLFOXIRI-bevacizumab[134].

The analysis of CTCs may provide information about the molecular characteristics of occult minimal residual disease in patient without

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

detectable metastases[135]. Thus, CTCs hold great amount of information, been very promising as surrogate biomarker in clinical practice for patient stratification, prognosis prediction, cancer screening or treatment monitoring.

3.4 CTC cell lines as models for the study of CTCs

Cancer cell lines are widely used to study molecular mechanisms of tumour cell biology[136] and to identify new biomarkers, therapeutic targets and mechanisms of resistance[137]. The development of CTC cell lines shows great promise not only for the development of personalized cancer medicine, but also for performing studies to help understand the mechanisms of the development of cancer metastasis. This will allow to identify proteins and pathways involved in cancer cell stemness and dissemination, and therapeutic resistance[138,139].

The establishment of CTC lines from patients is a very challenging task, since there are very low numbers of these cells in the bloodstream. Their culture relies on a large number of CTCs, which is mainly reduced to patients with advanced disease. It is also a challenge to distinguish viable cells suitable for culture and also knowing the favorable environment for their growth[138].

3.4.1 Colorectal CTC cell lines

In 2015 Cayrefourcq L. *et al* [140] established the first CTC CRC cell line, that was designated as CTC-MCC-41. The CTC-MCC-41 cell line was established from a CRC patient with metastatic disease with numerous abdominal and mediastinal metastatic lymph nodes as well as metastatic lesions in the liver. The primary tumour was *KRAS* wild-type (codons 12 and 13) and *BRAF* mutated (V600E). The patient did not respond to any of the treatments: first under the combination of FOLFIRI and bevacizumab and

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

then combining FOLFOX and bevacizumab. CTC-MCC-41 was characterized at genomic, transcriptomic, proteomic and secretomic level, showing a wide variety of chromosomal aberrations across the genome typical of CRC. It also presented an intermediate epithelial-mesenchymal phenotype, epithelial properties with stem cell-like characteristics and an osteomimetic signature characteristic from a bone marrow origin. CTC-MCC-41 also showed the capacity of inducing angiogenesis *in vitro* and *in vivo* tumour formation after being xenografted in immunodeficient mice.

Besides the CTC-MCC-41 cell line, in a posterior work, Soler *et al.*[139] established other eight CTC cell lines from the same patient with samples collected at different time points during the clinical follow-up (**Figure 5**). At the time of the second relapse the cell line CTC-MCC-41.4 was established, and at the time of the third relapse seven CTC lines, that started to proliferate at different time points, were established: CTC-MCC-41.5A, CTC-MCC-41.5B, CTC-MCC-41.5C, CTC-MCC-41.5D, CTC-MCC-41.5E, CTC-MCC-41.5F, and CTC-MCC-41.5G. All these cell lines were also characterized at genomic, transcriptomic and proteomic level showing all cell lines strong expression of epithelial markers and a wide variety of chromosomal aberrations across the genome typical of CRC. Functional assays were also performed showing all cell lines capacity of inducing *in vitro* angiogenesis.

This compendium of cell lines represents the unique model of CTC cell lines of CRC that have been established until now.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

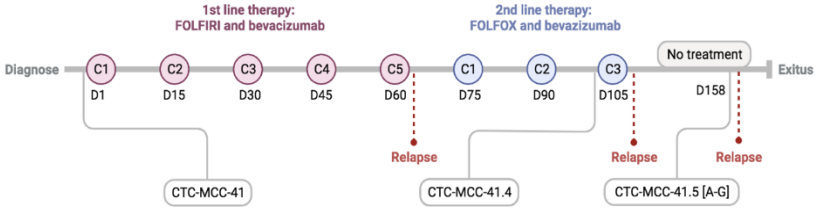


Figure 5. Timeline of the establishment of the nine CTC cell lines from the peripheral blood of a mCRC patient. The cell line CTC-MCC-41 was established before treatment initiation, CTC-MCC-41.4 at the time of the second relapse and at the end of second-line therapy, and CTC-MCC-41.5 [A-G] at the time of the third relapse before the patient´s death. D, day; C, treatment cycle (Adapted from Soler *et al.*[139]). Created with Biorender.com

4. Epigenetics

The term epigenetics was first introduced by Waddington in the 1940s, and since then, the meaning has gradually evolved. Nowadays, we can define epigenetics as the study of changes in gene function that are mitotically and/or meiotically heritable and do not involve changes in the DNA sequence[141].

Epigenetic modifications are responsible for regulating tissue and cell-type specific gene expression patterns[142]. These epigenetic gene patterns have an important role in different biological processes like genetic imprinting, embryonic development and inactivation of the X chromosome[143].

Epigenetic marks are reversible and they can be modulated by environmental factors like diet, behavioral factors like stress or the exposure to diverse toxins or pharmacological compounds[144]. These epigenetic marks can provide the link between environmental factors and phenotypical changes during the whole lifetime of an organism[142].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.1 Epigenetic mechanisms

The main epigenetic processes responsible for epigenetic regulation include DNA methylation, post-translational modifications of histones and nucleosome positioning and non-coding RNAs (ncRNAs)[143,145,146].

4.1.1 DNA Methylation

Covalent modifications of the DNA bases were described long ago, but they were not linked with gene expression until the end of 1960s[147]. DNA methylation, which is the most widely studied epigenetic mark in the human genome, is a covalent modification of DNA that consists in the incorporation of a methyl group (CH_3) to the 5′ carbon of cytosines in cytosine-phosphate-guanine (CpG) dinucleotides to produce 5-methylcytosine (5mC)[145,146], being essential for gene expression regulation in normal cell development and differentiation[148].

DNA methylation results from the role of the enzymes DNA methyltransferases (DNMT), which transfer a methyl group from S-adenosyl methionine to DNA. In the case of mammals, this family is composed by DNMT1, DNMT2, DNMT3a, DNMT3b and DNMT3L[147]. Out of these, the only ones with methyltransferase activity are DNMT1, responsible for methylation maintenance and DNMT3a and DNMT3b, responsible for the *de novo* methylation of unmethylated sites[147].

Cytosine methylation occurs almost mainly in the context of CpG dinucleotides, which tend to cluster in regions called CpG islands (CpGIs)[145]. These CpGIs regions are characterized for having more than 500 bp and with at least a 55% of CpG content[149]. They are associated with ~60% of the human gene promoters and they are usually unmethylated in normal cells even though some become methylated in a tissue specific way during early development or in differentiated tissues[145]. Generally, methylation of CpGIs

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

in the promoter of a gene is inversely associated with its transcription[142], associating methylation with gene silencing and unmethylation with gene expression (**Figure 6**). DNA methylation can inhibit gene expression by recruitment of methyl-CpG-binding domain proteins or by preventing the recruitment of DNA binding proteins from their target sites, while unmethylation generates a chromatin structure favorable for gene expression[142,145].

Methylation of DNA has also been described in regions out of the CpGIs context[150]. According to their distance from CpGIs, these are called CpGI shores, which are regions in close proximity, up to 2 kb upstream or downstream from CpGIs, presenting lower CpG density; CpGI shelves, between 2 and 4 kb from CpGIs and the open sea regions with a distance over 4 kb from CpGIs[151,152]. In CpGI shores, methylation is generally associated with transcriptional inactivation, however, in gene bodies methylation may be positively correlated with gene expression[145], which makes the context for DNA methylation distribution vital in gene regulation[152].

DNA methylation can affect either intragenic or intergenic regions of the genome[153]. In this sense, there is a significant amount of methylated CpGs in repetitive elements, including satellite DNA and parasitic elements such as long interspread transposable elements (LINE) and short interspread transposable elements (SINE)[154]. Methylation would maintain chromosomal integrity by preventing their activation, that would lead to chromosomal instability, translocations and gene disruption[145].

DNA methylation can be reversed by ten-eleven translocation (TET) enzymes, which induce the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) in a process defined as DNA hydroxymethylation[155]. This

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

demethylation mechanism can modulate gene expression by adjusting the methylation levels[156].

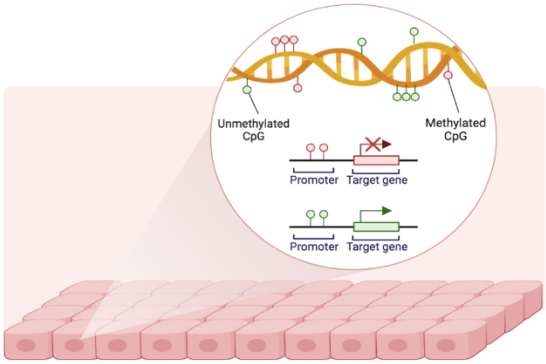


Figure 6. Role of DNA methylation in regulating gene expression. In promoter CpGIs of genes, hypermethylation is usually inversely associated with silencing of its transcription, and hypomethylation with activation of gene expression. Created with Biorender.com

4.1.2 Other epigenetic mechanisms

4.1.2.1 Post-translational modification of histones

In the mid 1960s it was already proposed that histone modifications might have an influence on the regulation of transcription[157]. Nowadays it is greatly known that histones and their post-translational modifications play a key role in epigenetics[152]. Histones consist mostly of a globular domain and a N-terminal tail, which protrudes from the nucleosome making it accessible for binding to the DNA, for protein interactions and covalent attachment of molecules[158].

A wide range of covalent post-translational modifications have been described[159]. These histone modifications alter nucleosome structure and DNA accessibility[160] and are involved in transcriptional regulation, DNA

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

replication, DNA repair, alternative splicing and chromosome condensation[145] (**Figure 7**).

All histones are subject to post-translational modification[147] and although these modifications take place mostly on the tail, they can be modified at different sites simultaneously. Histones forming the nucleosomes can each have several modifications and the combination of all marks in a nucleosome or region specifies outcome and not a single histone mark by itself[145]. There are many enzymes catalyzing and removing covalent post-translational modifications such as histone methyltransferases, demethylases, acetyltransferases and deacetylases [147].

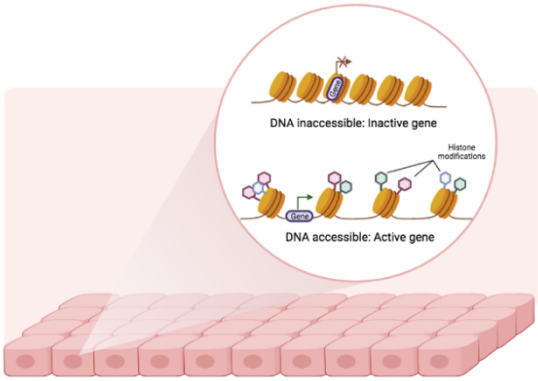


Figure 7. Role of histones in gene expression regulation. Histone modifications regulate nucleosome structure and DNA accessibility and are involved in biological processes such as transcriptional regulation. Created with Biorender.com

4.1.2.2 Nucleosome positioning

Nucleosomes are the smallest packaging unit of the chromatin fiber, consists of 147 bp of DNA wrapped around a histone octamer[142]. Positioning of these nucleosomes across the genome is not random and plays essential

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

roles in gene regulation[161]. Nucleosomes are a barrier to transcription, blocking the access of activators and transcription factors to DNA, and they also inhibit the elongation of the transcripts by engaged polymerases[145].

It is in particular the exact position of the nucleosomes around the transcription start sites (TSSs) what has an important influence on the initiation of transcription[145]. The loss or the addition of a nucleosome directly upstream of the TSS is correlated with gene activation and repression respectively[145] (**Figure 8**). Nucleosome positioning has also an important role in shaping the methylation pattern and at the same time, histone variants regulate nucleosome positioning and gene expression[145].

Different structurally related ATP-dependent chromatin remodeling proteins and complexes known to move, destabilize, eliminate or remodel nucleosomes have been identified[145,162].

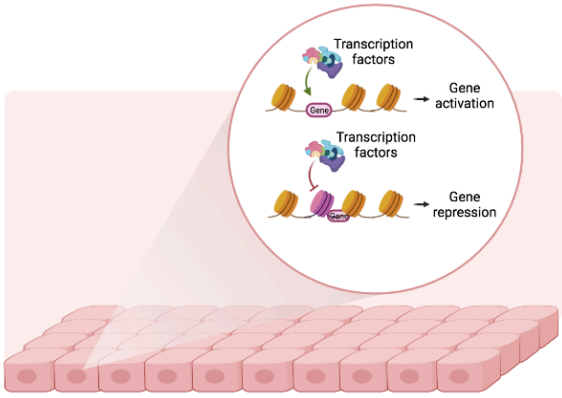


Figure 8. Role of nucleosomes in gene expression regulation. Positioning of nucleosomes regulates the access of activators and transcription factors to the DNA. Created with Biorender.com

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.1.2.3 Non-coding RNAs

It is well established that although practically the entire genome is transcribed to RNA only a 2% approximately is subsequently translated to proteins[152]. Out of these RNA transcripts that do not encode for proteins, the ncRNAs, there is a subset involved in gene regulation through different ways, such as epigenetic silencing or post-transcriptional regulation of mRNA stability[163] (**Figure 9**). These ncRNAs are generally classified according to their size range, distinguishing between small non-coding RNAs (small ncRNAs), when their size is under 200 nt, and long non-coding RNAs (lncRNAs), when their size is over 200 nt[164].

In the case of the small ncRNAs, members of these group can act as key regulators of gene expression in many different cellular pathways and systems[165]. These tend to show high degree of sequence conservation across species and are mainly involved in transcriptional and post-transcriptional gene silencing through specific base pairing with their targets[152]. LncRNAs show poor cross-species sequence conservation[152], and they are a more diverse group of regulatory RNAs in relation with characteristics, localization and modes of action[165].

Besides, it is also important to mention that despite the recent advances in the field of ncRNAs, there are still a vast majority of ncRNA completely uncharacterized[165].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

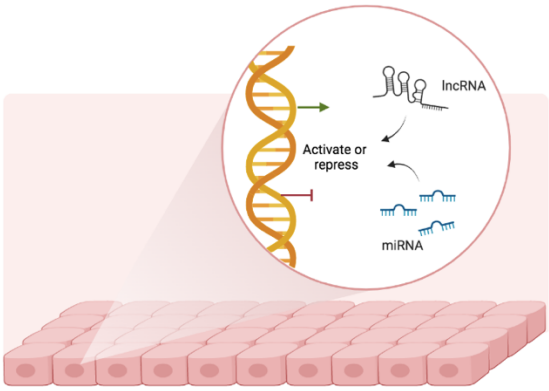


Figure 9. Role of ncRNAs in gene expression regulation. Regulatory ncRNAs are involved in gene regulation in different ways such as epigenetic silencing or post-transcriptional regulation of mRNA stability. Created with Biorender.com

4.2 Interplay between epigenetic mechanisms

Epigenetic modifications are tightly interconnected and form a network of regulatory pathways and feedback loops[166]. DNA organization and DNA packaging into chromatin is mediated by interplay between all epigenetic mechanisms, and it is this interplay amongst DNA methylation, histone modifications and ncRNAs what determines the specific structure of chromatin and the establishment of the diverse patterns of gene expression[166,167]. Examples of this interplay amongst epigenetic machinery are the regulation of DNA methylation by the interaction between DNMTs and other epigenetic factors, such as histone modifications, that direct machinery to specific parts in the genome[145,158,168]; the association of ncRNAs both in normal and pathological conditions, with regulation of deposition, alterations and removal of histone marks[167] and the regulation of the expression of different types of ncRNAs by histone modification

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

patterns[169,170]. Moreover, the disruption of ncRNA expression can also alter histone mark and DNA methylation patterns[171].

4.3 DNA Methylation and cancer

Cancer initiation and progression are accompanied by changes in DNA methylation[172]. Two types of general changes occur in tumour cells in comparison with normal cells and it is genome-wide hypomethylation and site-specific CpGI promoter hypermethylation[173]. This is also seen in CRC, where malignant cells have 10–40% lower levels of absolute methylation compared with normal colonic tissue, however, some CRCs have extensive and co-ordinate patterns of promoter hypermethylation[174].

Genome-wide hypomethylation is refered to a global genomic decrease in the methylcytosine content[175]. This loss of DNA methylation at CpG dinucleotides was the first epigenetic change to be identified in cancer cells, firstly in colon carcinoma[176,177] and latter in a wide range of cancers in comparison to matched normal tissues[175,178]. Global DNA hypomethylation has a significant role in tumourigenesis and occurs at different genomic sequences that are normally methylated in normal cells, like retrotransposons, repetitive elements, introns and CpG-poor promoters[172].

In normal cells, repetitive sequences are silenced, ensuring genomic integrity and stability. However, hypomethylation of these repetitive elements, seen in a wide range of tumours, leads to increased genomic instability by promoting chromosomal rearrangements[179,180]. Thus, transposable elements are reactivated and can integrate at other genomic regions leading to mutagenesis and genomic instability[172,180]. DNA hypomethylation of individual genes in cancer cells is quite uncommon[180] however loss of methylation in CpGI promoter regions is associated with the expression of certain cancer-associated genes[181]. Moreover, hypomethylation can also lead to loss of imprinting in tumours, leading to the

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

activation of genes involved in growth promotion[182]. Thus, DNA hypomethylation leads to aberrant activation of genes and non-coding regions through different mechanisms contributing to cancer development and progression[172,180].

On the other hand, also site-specific hypermethylation contributes to tumourigenesis by silencing tumour suppressor genes through methylation of CpGIs in regulatory regions, which are unmethylated in normal cells. These genes play a role in cellular processes that are integral to cancer development and progression, like DNA repair, cell cycle related, apoptosis, cell adhesion and signal transduction[172]. Moreover, DNA hypermethylation can also silence different genes in an indirect way by silencing transcription factors or genes involved in different receptor-mediated signaling pathways[172,180].

It is also important to mention that although there are certain genes that are hypermethylated in diverse tumour types, promoter methylation is cancer-cell-specific, and each cancer type can be defined by a specific methylome. Moreover, differences in DNA methylation patterns may reflect the stage of tumourigenesis or even difference between the particular histological features[180,183].

4.3.1 Relevance of the study of DNA methylation in cancer

DNA methylation is a promising biomarker due to its early and frequent occurrence in cancer, its relatively easy detection by well-established techniques, its stability, its presence in a wide range of body fluids and its cell-type specificity[184–186]. These characteristics make DNA methylation a great tool for assessing cancer risk, early detection, prognosis, and predicting therapeutic response[184–186].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Changes in the DNA methylation pattern usually occur early in cancer development, therefore, the abnormal methylation of cancer genes could be served as biomarkers for early cancer detection[186]. In this sense, in CRC, methylation of the *VIM* promoter in feces[187] and methylation of the *SEPT9* promoter in plasma[188] have great specificity and sensitivity for detection of CRC, leading both to the development of commercial tests: ColoSure and mSEPT9 respectively[185]. In line with this, the Cologuard test includes DNA assays for aberrant *NDRG4* and *BMP3* methylation with high sensitivity in people at average risk for CRC[186,189].

DNA methylation can also be used as a molecular prognosis biomarker. In CRC, although none have reached clinical validation, hypomethylation of LINE-1 and *IGF2* is associated with adverse prognosis; and a gene methylation panel comprising *APC*, *p14* and *RASSF1A* identifies a subset of CRC patients with poor prognosis[190]. In addition, DNA methylation biomarkers can also predict responses to treatment, like hypermethylation of *MGMT*, which has shown correlation with 5-FU response[191].

Besides for the development of new biomarkers, the study of DNA methylation has been crucial for the identification of therapeutic targets and the development of efficient anti-tumour drugs and treatments. The epigenome, including methylation, is potentially reversible making it possible to be reprogramed, representing a promising therapeutic strategy by changing the expression of certain genes[192–194]. Although epigenetic mechanisms and their role in gene expression are complex there are available drugs that target epigenetic mechanisms[194], the so-called epidrugs, which are chemical compounds that alter DNA and chromatin structure and are able to reverse the epigenetic state including the DNA methylation pattern[192]. In line with this, there are currently two DNMT inhibitors, which are the nucleoside analogues 5-azacitydine (azacitidine)

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

and 5-aza-2′-deoxycytidine (decitabine), approved by the FDA and the European Medicines Agency (EMA) against different hematological malignancies (myelodysplastic syndromes, acute myeloid leukemia, and chronic myelomonocytic leukemia).

Next-generation epigenetic drugs keep being developed and different clinical trials are testing the efficacy of combining different epidrugs among themselves and also with traditional therapies[195]. Moreover, besides the approved epidrugs, it has been shown that 1% of all drugs that have been approved by the FDA show significant epigenetic activity and silence promoters in colon cancer cells, evidenciating the importance of the study of the epigenome and of investing on epigenetic therapies[193,196].

5. CTCs, DNA methylation and CRC

Alterations in DNA methylation have been detected in a wide range of liquid biopsy components and also in a wide range of biological fluids, being blood the most explored[197,198].

Analysis of cell-free DNA (cfDNA) has started to be applied to oncologic therapy selection and a plethora of epigenetic biomarkers based on altered methylation have been evaluated in cfDNA for the management of CRC. However, CTCs have shown until now limited translation into routine clinical practice. In spite of this, epigenetic markers of CTCs, such as DNA methylation of tumour suppressor and metastasis suppressor genes, have been explored to better understand cancer progression[199], although the number of studies focused on this circulating tumour population is still very low[200].

The first study to determine the DNA methylation profile of cancer-related genes in CTCs of CRC patients revealed that these CTCs are characterized by hypermethylation of the *SFRP2* promoter and exon 1

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

of *VIM*[198], which are genes related to EMT and CRC metastasis[201,202]. Comparison of methylation levels of *VIM* and *SFRP2* in CTCs of CRC patients with those of tumour tissues, showed a strong correlation between the methylation status of the *SFRP2* promoter in CTCs and that of the corresponding tissue, but a weaker correlation in the case of *VIM*. Therefore, although the CTC population can show a different methylation pattern than tumour cells located in primary tissue and metastasis due to the gain of specific characteristics in the bloodstream, these data demonstrated that methylation profiling of CTCs in CRC patients represents a promising non-invasive approach for tumour detection[198].

The analysis of single CTCs with genome-wide methylation[119] or other multiplex PCR approaches[203] has provided relevant information on cancer dissemination, however, the study of epigenetic alterations in CTCs from CRC patients is still limited. Thus, despite the difficulty of translating CTCs into the clinical context, their molecular characterization has provided valuable information for understanding how colorectal tumour cells are able to disseminate, implant at distant locations, and generate metastasis[204,205].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

HYPOTHESIS

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

HYPOTHESIS

Advanced CRC continues to be a major health burden despite advances in detection and therapy. Great number of patients have metastases at the time of diagnosis. In CRC, one of the great challenges continues to be patient selection to optimize treatment response and also the development of treatment resistances. All these evidenciates the need of new therapeutic targets and biomarkers to benefit patients in the clinical practice. In line with this, an increasing number of studies have shown that DNA methylation is involved in the regulation of cancer-related genes in colorectal malignant cells, and it is associated with metastasis development and also with treatment efficacy. This, together with the fact that it is a reversible mark, makes it of great interest as target for drug development and a promising biomarker to assess therapeutic response. Alterations of DNA methylation have been detected, not only in tumour tissues, but also in CTCs, which are cells that spread from primary tumours or metastatic sites to colonize distant tissues. The study of these CTCs has provided in recent years valuable information on the development of cancer metastasis and its clinical applications. However, the disruption of DNA methylation in colorectal CTCs is still poorly explored and has been only studied in a few particular genes. Therefore, in this context, we hypothesize that the analysis of the DNA methylome of CTCs could provide new information about the biology of CTCs and the development of metastasis, allowing the identification of new potential therapeutic targets and epigenetic biomarkers that help guide therapeutic decisions in patients with advanced disease. Thus, the epigenomic insights of CTCs provided in this work could bring great benefits to the clinical practice, helping to achieve a more personalized medicine in CRC.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

OBJECTIVES

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

OBJECTIVES

The main objective of this doctoral thesis was the epigenomic characterization of the DNA methylation profile of CTCs in mCRC, with the aim of revealing new information about the biology of these cells and the development of metastasis, as well as new therapeutic targets, resistance mechanisms and potential biomarkers useful to guide therapeutic decisions. To this end, we followed these specific objectives:

O1. To evaluate the DNA methylome of a metastatic-competent CTC cell line derived from a mCRC patient in comparison with colorectal primary and metastatic tumour cells. This evaluation was carried out to identify aberrant DNA methylated genes related to the formation of metastasis and to identify new potential biomarkers and therapeutic targets of CTCs.

O2. To study the DNA methylome of a unique collection of nine CTC cell lines obtained from a mCRC patient before initiation of a 5-fluorouracile-based therapy and at different time points during the treatment of the disease. Using this *in vitro* model, this study aimed to reveal new mechanisms and biomarkers linked to tumour progression and therapy resistance.

O3. To isolate CTCs from peripheral blood of mCRC patients and study their DNA methylome in relation to therapy response and the progression of the metastatic disease. This study aimed to identify aberrantly methylated genes and new potential biomarkers in CTCs mainly associated with the response and resistance to therapy based on FOLFOX in monotherapy or in combination with anti-VEGF treatment.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER I

Epigenomic characterization of the metastasis-competent cell line CTC-MCC-41 respect to primary and metastatic CRC cell lines

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Epigenomic characterization of the metastasis-competent cell line CTC-MCC-41 respect to primary and metastatic CRC cell lines

1. Abstract

Metastasis formation is the main cause of cancer-related deaths in CRC patients, and both CTCs and alterations in epigenetic mechanisms are involved in its development. In this chapter, we characterize the DNA methylation profile of metastasis-competent CTCs in CRC, using Infinium DNA MethylationEPIC arrays to compare the human CRC CTC-derived cell line CTC-MCC-41 with primary and metastatic CRC cell lines. Relevant differentially methylated CpGs obtained from the EPIC array analysis were validated with pyrosequencing and quantitative methylation-specific PCR (qMSP), and association between DNA methylation and the transcriptional profile of CTC-MCC-41 was also evaluated. To confirm the epigenetic regulation of hypermethylated genes, CTC-MCC-41 cells were treated with the demethylating agent 5-aza-2′-deoxycytidine (AZA) and gene expression was analyzed with quantitative reverse transcription PCR (RT-qPCR). The CTC-MCC-41 cell line showed a global DNA methylation profile different to primary and metastatic CRC cells. This specific DNA methylation pattern was associated with the regulation of its transcriptional profile and AZA was able to restore the expression of the epigenetically silenced genes in CTC-MCC-41. This study provides relevant new insights into the epigenomic landscape of metastasis-competent cancer CTCs in CRC, revealing important biological information for metastasis development, as well as new potential biomarkers and therapeutic targets in CRC.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

2. Introduction

Metastasis is the last stage of cancer progression and the cause of more than 90% of cancer deaths[69]. In the development of metastasis, CTCs have a fundamental causal role[206,207], since these cells spread from primary or metastatic tumour sites and colonize distant tissues to form metastasis[128,206]. Although in recent years relevant advances on the molecular properties of CTCs have been made, the molecular characterization of CTCs continues to be very challenging due to the low presence of these cells in peripheral blood of cancer patients[206]. Therefore, and since cancer cell lines are widely used to study molecular mechanisms of tumour cell biology[136], it is important to take advantage of CTC-derived cell line models to increase our knowledge about the metastatic process[208].

In this sense, Cayrefourcq *et al.* established the first human CTC-derived cell line of CRC, designated CTC-MCC-41. This cell line was isolated from the peripheral blood of a nonresectable metastatic colorectal patient before starting first-line chemotherapy. CTC-MCC-41 cells are metastasis-competent cells with a stable intermediate epithelial-mesenchymal phenotype and stem cell-like properties[140]. This cell line has been characterized at transcriptional level, revealing the differential expression of genes involved in relevant cancer-related pathways, such as energy metabolism and DNA repair[209].

Epigenetic mechanisms, such as DNA methylation, contribute to cancer progression and there is great evidence of aberrant DNA methylation patterns in tumour cells, which may be different between primary and metastatic tumours[210,211]. In the case of CRC, it has recently been described that colorectal CTCs may undergo aberrant methylation in some particular genes[198]. However, the DNA metylome of colorectal CTCs with the ability to initiate and support metastasis formation is still unexplored, since molecular

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

analysis of CTCs have rather been focused on transcriptional, genomic or even proteomic level[212]. Therefore, it is important to characterize the DNA methylation pattern of this subset of CTCs to better understand the metastatic process and obtain new clinical applications for cancer patients.

Thus, in this chapter, we analyze the DNA methylome of the metastasis-competent CTC-MCC-41 cell line in comparison with primary (HT29, Caco2, HCT116, RKO) and metastatic (SW620 and COLO205) CRC cell lines. The aim was to identify aberrant DNA methylated genes related to the formation of metastasis and also to identify new potential biomarkers and therapeutic targets of CTCs.

3. Materials and methods

3.1 Cell culture

3.1.1 Primary and metastatic tumour cell lines

A total of four cell lines of primary tumour origin and two cell lines of metastatic tumour origin were included in the study (**Table 1**). All of them were purchased from the American Type Culture Collection (ATCC) (Rockville, MD). These cell lines were cultured according to ATCC recommendations with the following media: RKO and Caco2 with EMEM, HT29 and HCT116 with McCoy's 5A, COLO205 with RPMI-1640, and SW620 with Leibovitz's L-15. All medias were supplemented with 1% penicillin/streptomycin solution (Gibco, Thermo Fisher Scientific, USA) and 10% Fetal bovine serum (FBS) (Eurobio, France), except for the culture of Caco2 were 20% FBS was used. All cells were grown at 37 °C, 5% CO_2 and culture medias were changed once or twice a week.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

For cell collection, once cells reached confluency, media was discarded for adherent cells and collected for those cells that grow also in suspension. After that, cells were washed with phosphate buffered saline (PBS) (Merck, Germany) before adding 1 mL of trypsin-EDTA solution (Merck, Germany). Three passages of each cell line were collected. Once cells were recovered, they were centrifuged 5 min at 600 x g, and a dry pellet was kept at -80 ºC until DNA extraction.

3.1.2 CTC-MCC-41 cell line

The CRC CTC cell line CTC-MCC-41 (**Table 1**) grows in suspension and was cultured using Corning Ultra-Low attachment cell culture flasks (Merck, Germany), and RPMI-1640 medium (ATCC, USA) supplemented with 10% FBS (Eurobio, France), 1% penicillin/streptomycin solution (Gibco, Thermo Fisher Scientific, USA), 1% L-glutamine (Merck, Germany), 1% of insulin-transferrin-selenium (ITS-G) (Gibco, Thermo Fisher Scientific, USA), basic human fibroblast growth factor (bFGF) (Miltenyi Biotec, Germany) at a final concentration of 10 ng/mL and human epidermal growth factor (hEGF) (Miltenyi Biotec, Germany) at a final concentration of 20 ng/mL. Cells were grown at 37 °C, 5% CO_2 and culture medias were changed one or twice a week. Three passages of this cell line were collected. For this, once cells reached confluency, they were centrifuged 8 min at 600 x g, keeping the dry pellet for at -80 ºC until DNA extraction.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Table 1. General characteristics of the human CRC cell lines used in the study for DNA methylation analysis.

Cell line	Disease	Tissue	Age (yrs)	Gender
HT29	Colorectal adenocarcinoma	Primary tumour (colon)	44	Female
Caco2	Colorectal adenocarcinoma	Primary tumour (colon)	72	Male
HCT116	Colorectal carcinoma	Primary tumour (colon)	48	Male
RKO	Colorectal carcinoma	Primary tumour (colon)	-	-
COLO205	Colorectal adenocarcinoma (Dukes´ type D)	Metastasis (ascitic fluid)	70	Male
SW620	Colorectal adenocarcinoma (Dukes' type C)	Metastasis (lymph node)	51	Male
CTC-MCC-41	Colorectal adenocarcinoma	CTC (peripheral blood)	57	Male

Yrs, years.

3.2 Cell line treatment with AZA

With the purpose of evaluating the epigenetic regulation of gene expression by DNA methylation, the CTC-MCC-41 cell line was treated with the demethylating agent AZA (Sigma-Aldrich, USA). For these experiments, $2.5 \cdot 10^6$ cells were plated in Corning Ultra-Low attachment cell culture flasks (Merck, Germany) for both controls and treated cells. For control cells, DMSO at 1% (Merck, Germany) was added to a final concentration of 5 μM, and for treated cells, AZA was added to a final concentration of 5 μM. Cells were

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

transferred to the incubator at 37 °C, 5% CO_2 and incubated for 72 h. After 36 h of treatment initiation, AZA was renewed, and at the end of the experiment cells were collected and resuspended in 1mL of TRIzol Reagent (Invitrogen, USA) and kept at -80 ºC until RNA extraction.

3.3 Nucleic acids extraction and quantification

3.3.1 DNA extraction

For DNA extraction, cell pellets were resuspended in 450 µL of lysis buffer (100 mmol/L Tris pH 8.0, 50 mmol/L EDTA, 0.1 mol/L NaCl), 50 µL of SDS 10% and 2.5 µL of proteinase K (Qiagen, Germany), followed by an overnight incubation at 37 ºC. After cell lysis, 0.2 µL of RNAse A (Qiagen, Germany) was added and incubated 1 h at 37 ºC and also 169 µL of NaCl were added before centrifugation at 19,000 x g at 4 ºC. The supernatant was then transferred to a new tube, and 540 µL of isopropanol was added for DNA precipitation. After centrifugation at 19,000 x g at 4 ºC, the supernatant was removed, the DNA was washed with 70% ethanol and finally resuspended in nuclease free water.

3.3.2 DNA quantification

All DNA samples were quantified by the fluorometric method Qubit 4 (Invitrogen, USA), using the Qubit dsDNA BR (Broad Range) Assay Kit (Invitrogen, USA). For each assay, Qubit working solution was previously prepared by diluting the Qubit dsDNA BR Reagent 1:200 in Qubit dsDNA BR Buffer. For standards, 190 µL of working solution and 10 µL of each kit standard was mixed, and for samples, 199 µL of working solution and 1 µL of DNA was mixed. Both kit standards and DNA samples were measured after a two-minute incubation.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.3.3 RNA extraction

RNA extraction from cells was carried out using the TRIzol Reagent (Invitrogen, USA) following manufacturer´s instructions. Briefly, after unfreezing the samples, they were incubated at room temperature for 5 min to allow the dissociation of the nucleoproteins complex and 200 µL of chloroform was added before centrifugation to separate the RNA. The aqueous phase, containing the RNA, was transferred to a new tube and 500 µL of isopropanol was added for RNA precipitation. After centrifugation, supernatant was removed, and the pellet was washed with 75% ethanol and air dried to be finally resuspended in RNase-free water.

3.3.4 RNA quantification

Concentration and purity of all RNA samples was determined using the spectrophotometer NanoDrop One (Thermo Fisher Scientific, USA). Absorbance was measured at a waveleght of 260 nm using 1 µL of sample and using nuclease free water as blank. Purity was assessed according to the ratios A260/A280 and A260/A230, in which a value close to 2 was considered optimal for RNA.

3.4 DNA methylation analysis

3.4.1 Bisulfite treatment

Bisulfite treatment of methylated DNA is the most used technique to identify methylation. Incubation with bisulfite converts unmethylated cytosines into uracil, keeping methylated cytosines unchanged during the treatment[213]. For each sample, 500 ng of genomic DNA was treated with sodium bisulfite using the kit EZ DNA Methylation kit (Zymo Research, USA) following manufacturer's instructions for the Illumina Infinium Methylation Assay (Illumina, USA).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Samples were modified using a ProFlex PCR System (Thermo Fisher Scientific, USA) under the incubation conditions of 95 °C for 30 sec and 50 °C for 60 min, for a total of 16 cycles, and eluted in a volume of 12 µL.

3.4.2 Infinium MethylationEPIC array

After the bisulfite treatment, DNA samples were hybridized to the Infinium MethylationEPIC BeadChips (Illumina, USA), following the Infinium HD methylation assay and according to the manufacturer´s instructions. This microarray assay allows to interrogate the methylation status of over 850,000 CpG sites all across the human genome at a single-nucleotide resolution[214]. Briefly, DNA samples were amplified before enzymatic fragmentation and DNA precipitation, once this fragmented DNA was resuspended it was hybridized by sequence complementarity to the probes on the BeadChips. After hybridization, BeadChips were washed and stained and single base extension of labeled nucleotides occurred.

Once the BeadChips were hybridized and stained, they were scanned with a HiScan System (Illumina, USA). Image intensities were extracted using GenomeStudio (v2011.1) Methylation module (1.9.0) software from Illumina. Data quality control was assessed with GenomeStudio and BeadArray Controls Reporter (Ilumina), based on the internal control probes present on the array. The methylation score of each CpG from samples that passed this quality control was represented as β-value and previously normalized for color bias and background level adjustment, and quantile normalization across arrays. β -values were obtained as the ratio of fluorescent signal between methylated (M) probe relative to the sum of the M and unmethylated (U) probes (β = M/(M+U)). The β-values range from 0 (no methylation) to 1 (completely methylated). Probes and sample filtering involved a two-step process for removing SNPs and unreliable β-values with a high detection P-value >0.01. After this filtering step, the remaining CpGs were considered

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

valid for the study. Unsupervised hierarchical clustering heatmaps with the ComplexHeatmap package, scatter plots and principal component analysis (PCA) of β-values were carried out using R environment. The gene ontology (GO) enrichment analysis of biological and Panther pathways for the methylation profiles were evaluated using GENECODIS[215].

3.4.3 Validation of the Infinium MethylationEPIC array data

The methylation data obtained with the methylation EPIC array was validated using pyrosequencing. This methodology was used for all the selected genes except for one, that was not possible to analyse by pyrosequencing and quantitative methylation-specific PCR (qMSP) was used.

3.4.3.1 Pyrosequencing

Pyrosequencing is a technology based on the principle of sequence by synthesis, that uses fluorescently labeled nucleotides to give quantitative information of the methylation pattern at single or multiple CpGs in the genes of interest[216].

The design of the primer sequences (**Table 2**) was performed with PyroMark Assay Design 2.0 (Qiagen, Germany). Bisulfite-converted DNA of the samples was previously amplified performing standard PCR for all candidate genes. PCR was conducted in a final volume reaction of 29 µL which contained 10 ng of bisulfte-converted DNA, 3 µl of Immobuffer 10x, 0.9 µl of $MgCl_2$ 50 mM, 3 µl of dNTPs 2 mM, 0.12 µl of Immolase Hot start Polymerase, 0.3 µl of both forward and reverse primers at 10 µM and nuclease-free water. The PCR reaction was performed using a ProFlex PCR System (Thermo Fisher Scientific, USA) under the following conditions: 1 cycle at 94 °C for 10 min, 50 cycles at 94 °C for 30 s, 52-64 °C for 30 s, and 72 °C for 30 s, and a final extension at 72 °C for 8 min.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Following PCR, sequencing reactions were performed using a PyroMark Q24 System (Qiagen, Germany) with the PyroMark Gold SQA™ Reagent Kit (Qiagen, Germany) and the sequencing primer at 10 μM, following manufacturer's instructions. Methylation values of the CpG sites of interest were obtained using PyroMark Q24 Software 2.0 (Qiagen, Germany). Methylation assays were conducted in triplicate and 100% Human Methylated & Non methylated DNA (Zymo Research, USA) were used as positive and negative controls respectively and water as no template control.

Table 2. Primer sequences for DNA methylation analysis by pyrosequencing.

Gene	Primer	Sequence (5´- 3´)
ABCB1	Forward	[Btn] GGTTGATTGGTTGGGTAGGAATA
	Reverse	ACACTTAAAAACTATCCCATAATAACTC
	Sequencing	ATCCCATAATAACTCCC
BCL11A	Forward	[Btn] TGGATGTTAAAAGGTATTGATGAAGATA
	Reverse	TCTCCCCCTCRCTTTTACTTCTAATCCT
	Sequencing	ACTTAAATACTAAAATTTACCTTA
BMP7	Forward	TGGTYGATTTTAGTTTGGATAA
	Reverse	[Btn] ATAAATACCRAATTATACTTACC
	Sequencing	ATTTTTTTTATTTTGGGTTTGTT
CCND2	Forward	GGGAGGAAGGAGGTGAAGAAA
	Reverse	[Btn] CACCACCCCTTCCTTTCC
	Sequencing	TTTGTAAAGATAGTTTTGATTTAAG
DKK1	Forward	TGGGGATTTAGGTTTGTAAAGTGA
	Reverse	[Btn] AATTCAAAATAACRCTCACTCCCAACAA
	Sequencing	TTTGAGTTTTTTTGAGATGAT
FN1	Forward	GGAAGAGGGGATTTGTAGTTATAATT
	Reverse	[Btn] ACAAATTACCACCAAATTTACTTCC
	Sequencing	AGGTATTAGAAGGGATTTT

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Table 2. Continuation

Gene	Primer	Sequence (5´- 3´)
GATA2	Forward	[Btn] GGGTAGGGGATTTAGTTAGG
	Reverse	AACAAATAATTCYCCTTTTTCTCCCAC
	Sequencing	ATTACCCTCCCCCTC
GJB6	Forward	TTAAGGGTYGGAAAGAGG
	Reverse	[Btn] AACCYAAACTATCCCTCC
	Sequencing	TGTATTTTTTAGGGGTTTTTTAT
PDGFC	Forward	GTTTTGGATGGGATTATGTGGA
	Reverse	[Btn] CCTAAATCCCCTATCTCTAACC
	Sequencing	ATGGGATTATGTGGAAA
PTGS2	Forward	AGGGAGAGAAATGTTTTAAGGT
	Reverse	[Btn] TCACCAATATAAAATTTCCTACCTTCTAA
	Sequencing	GGAGAGAAATGTTTTAAGGTATA
SEMA6A	Forward	AGATAAGTATATATGTAGGAATGATAAAGG
	Reverse	[Btn] CCCTCAAACTAAAATCCAACTACC
	Sequencing	ATAATGGATTATTAAATGGGATATA
SMARCA1	Forward	AGGGGAGTTTAGGTTTATTGTTGG
	Reverse	[Btn] TCCCCTTCCCCACCCAAAA
	Sequencing	GTGGGAAAGGAGATT
TGFB2	Forward	[Btn] ATTAGGGATTTGTTTTAGGAGAAG
	Reverse	ACATATCTATTTTATAAACCTCCT
	Sequencing	CATATCTATTTTATAAACCTCCTT

Btn, Biotin.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.4.3.2 Quantitative methylation-specific PCR (qMSP)

Quantitative methylation-specific PCR (qMSP) is a qPCR-based method that uses primers to detect both unmethylation and methylation of the region of interest. Primers (**Table 3**) were designed using Primer3 v.0.4.0[217] and all reactions were performed using 10 ng of bisulfited-converted DNA, in a final volume reaction of 20 µl, using 0.3 µL of the 10 µM forward and reverse primers and 10 µL of *Power* SYBR Green PCR Master Mix (Applied Biosystems, USA). All reactions were performed in triplicates both for methylation and unmethylation detection using a Step One Plus thermocycler system (Applied Biosystems, USA), using the following thermocycling conditions: 95ºC for 10 sec. and 50 cycles at 95ºC for 15 sec. and 60ºC for 60 sec. Amplification data was obtained with the software StepOne (Applied Biosystems, USA) and the analysis of the data was performed based on a formula described by Cottrell *et al.*[218]:

$$\text{Methylation (\%)} = 100/[1+2^{(CT_{CG} - CT_{TG})}]$$

In this formula, CT_{CG} represents the threshold cycle of the methylation status analyzed by the methylation primers and CT_{TG} represents the threshold cycle of the methylation status analyzed by the unmethylation primers.

Table 3. Primer sequences used for the qMSP analysis.

Gene	Status	Primer	Sequence (5´- 3´)
GAL	M	Forward	GTGAGCGTTTTAGGTCGTTA
		Reverse	ACTACGACTACGCCCGAAT
	U	Forward	GTGAGTGTTTTAGGTTGTTAGAGTT
		Reverse	CTTAAACTACAACTACACCCAAAT

M, methylation; U, unmethylation.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.5 Expression analysis

3.5.1 DNAse treatment and retrotranscription

Previous to the synthesis of cDNA, RNA was purified using the Turbo DNA-free Kit (Invitrogen, USA) following manufacturer´s instructions and using 10 µg of RNA as starting material. Briefly, TURBO DNase Buffer and Enzyme were added and incubated at 37 ºC. After this, DNase inactivation reagent was added and incubated to then centrifugate and recover the supernatant containing the purified RNA.

Retrotranscription of this purified RNA was performed using the SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen, USA) following manufacturer´s recommendations. 1-2 µg of RNA were used as starting material and then added 1 µL of 10 mM dNTP mix and 1 µL of 0.5 µg/µL oligo$(dT)_{12–18}$ and incubated at 65 ºC for 5 min. After this, 9 µL of the reaction mix was added to each sample, which contained 2 µL of 10X RT buffer, 4 µL of 25 mM $MgCl_2$, 2 µL 0.1 M DTT and 1 µL of RNaseOUT™ (40 U/µL) and incubated at 42 ºC for 2 min before adding 1 µL of SuperScript™ II RT. Samples were then incubated 50 min at 42 ºC and 15 min at 70 ºC. Finally, 1 µl of RNase H was finally added to the samples and after a 20 min incubation at 37 °C, cDNA samples were kept at -20 ºC.

3.5.2 Quantitative reverse transcription PCR (RT-qPCR)

The expression analysis of the selected genes was performed with RT-qPCR using the reagent *Power* SYBR Green PCR Master Mix (Applied Biosystems, USA) in a Step One Plus thermocycler system. The thermocycling conditions were the following: 95 ºC for 10 sec and 40 cycles at 95 ºC for 15 sec and 60 ºC for 60 sec. The PCR reaction was performed with 25-200 ng of cDNA, in a final volume reaction of 20 µl, using 0.3 µL of the 10 µM forward and reverse primers and 10 µL of SYBR Green reagent.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

All reactions were performed in triplicates for each gene, using the gene beta-2-microglobulin (*B2M*) as a housekeeping. Primer sequences were obtained from a previous work of Laure Cayrefourcq *et al.*[140] (**Table 4**). Data analysis was performed based on the relative quantification of the $2^{-\Delta\Delta CT}$ method[219].

Table 4. Primer sequences for the RT-qPCR analysis.

Gene	Primer	Sequence (5´-3´)
DKK1	Forward	GGTATTCCAGAAGAACCACCTTG
	Reverse	CTTGGCCAGAAGTGTCTAGCAC
GJB6	Forward	GAAACCACTCGCAAGTTCAGGC
	Reverse	AGGCTGCTTCAAAGATGATTCGG
PDGFC	Forward	TGAACCAGGGTTCTGCATCCAC
	Reverse	TAAGCAGGTCCAGTGGCAAAGC
PTGS2	Forward	CGGTGAAACTCTGGCTAGACAG
	Reverse	GCAAACCGTAGATGCTCAGGGA
SMARCA1	Forward	AACCTGGCAAGTGCTGATGTGG
	Reverse	CAGTGTTGTCAGTGATGAGACGG
TGFB2	Forward	AAGAAGCGTGCTTTGGATGCGG
	Reverse	ATGCTCCAGCACAGAAGTTGGC
B2M	Forward	GTCTTTCAGCAAGGACTGGTCT
	Reverse	TTACATGTCTCGATCCCACTTAAC

3.6 Statistical analysis

For genome-wide DNA methylation analysis, differentially methylated CpGs (DMCpGs) were determined using empirical Bayes methodology. P-values were corrected for multiple testing (False discovery rate, FDR) using Benjamini-Hochberg method, and a threshold of p-value <0.05 was selected

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

for significance. For the validation assay of DNA methylation levels and for the evaluation of expression differences after AZA treatment, the Kolmogorov–Smirnov test was first used to evaluate the normality of the distribution, and the Mann–Whitney U test or Student´s t -test was used accordingly. The GraphPad Prism 8.0 software (GraphPad Software, USA) and the R statistical environment version 4.2.0 were used for statistical analysis and graphical representation. All expressed p-values were calculated with two-tailed tests and were considered significant when the p-value <0.05.

4. Results

4.1 Evaluation of the DNA methylation profile of CTC-MCC-41 cells respect to primary tumour cell line HT29

The DNA methylation profile of CTC-MCC-41 was evaluated respect to the primary CRC cell line HT29 using the EPIC array technology. After performing normalization and the quality control of the data, we obtained 852,917 valid CpGs for the analysis (**Figure 10A**). Principal component analysis (PCA) of these CpGs classified both cell types into different groups, revealing a different methylation profile between CTC-MCC-41 and HT29 (**Figure 10B**), also reported in a scatter plot representing methylation of CTC-MCC-41 respect to HT29 (R^2 = 0.68) (**Figure 10C**). Moreover, we identified 188,185 significant differentially methylated CpGs (DMCpGs) (p-value <0.05; FDR adjusted p-value <0.05) with a difference of methylation (Δβ-value) greater than 0.20 (Δβ-value >|0.20|) between CTC-MCC-41 and HT29 tumour cells. In addition, a hierarchical clustering analysis revealed a DNA methylation profile able to clearly differentiate the CTC-MCC-41 from HT29 cancer cells (**Figure 10D**). **Supplementary Table 1** shows the top 20 DMCpGs found between CTC-MCC-41 and HT29.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

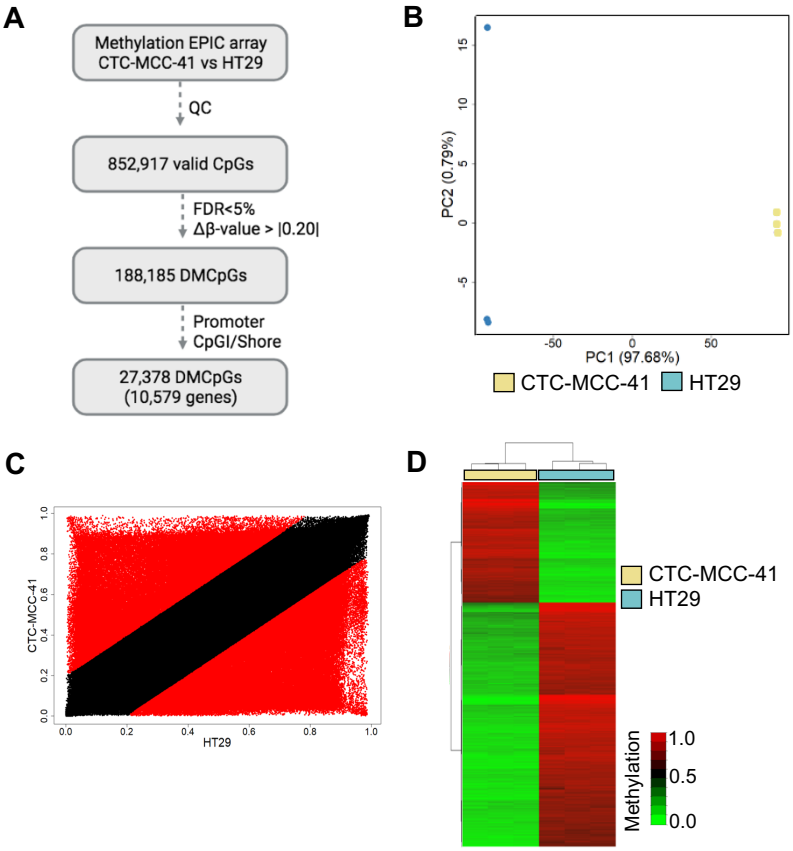
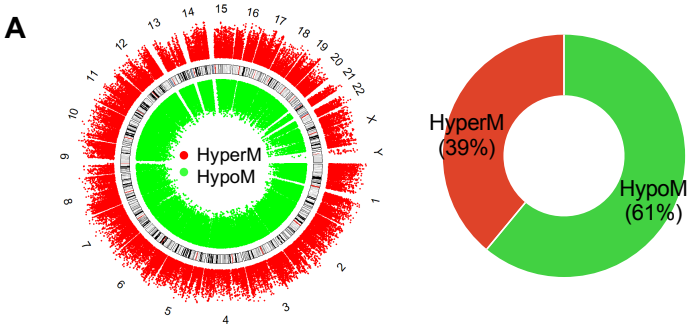


Figure 10. Genome-wide DNA methylation analysis of CTC-MCC-41 respect to the primary tumour cells HT29. (A) Schematic flowchart used to identify significant DMCpGs in CTC-MCC-41 compared to HT29. Three different passages of each cell line were analyzed **(B)** Principal component analysis (PCA) of the DNA methylation data obtained in CTC-MCC-41 and HT29. **(C)** Scatter plot representing mean normalized levels of DNA methylation (β-values) in CTC-MCC-41 and HT29. Dots in red show significant differentially methylated CpGs. **(D)** Hierarchical clustering heatmap of the 10,000 most DMCpGs (FDR adjusted p-value <0.05) between CTC-MCC-41 and HT29. QC, quality control; FDR, false discovery rate; CpGI, CpG island, HypoM, hypomethylated; HyperM, hypermethylated.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The 188,185 DMCpGs previously identified showed a wide distribution throughout all the chromosomes of the genome, showing a higher number of hypomethylated CpGs (114,621 CpGs; 61% of all DMCpGs) than hypermethylated CpGs (73,564 CpGs; 39% of all CpGs) in CTC-MCC-41 respect to HT29 tumour cells (**Figure 11A**). Both hypo- and hypermethylated CpGs of CTC-MCC-41 showed a similar distribution according to their CpG context and gene region location (**Figure 11B**). In particular, most of these DMCpGs (64% of hyper- and 66% of hypomethylated) were mainly located in Open sea regions, which are characterised for its low CpG density, whereas the remaining DMCpGs were distributed in regions with a higher enrichment of CpGs, such as CpGIs and shore regions. Regarding the gene region, both the hypo- and hypermethylated CpGs were fairly homogeneously distributed throughout promoters, gene bodies and intergenic regions. In addition, although most of the CpGs located at transcription factor-binding sites (TFBS) did not show methylation changes, some of these CpGs were differentially methylated. These methylation differences corresponded with a higher number of hypomethylated CpGs located at TFBS (19,651 CpGs; 17% of all the DMCpGs at TFBS) of CTC-MCC-41 cells than hypermethylated CpGs (11,762 CpGs; 16% of all the DMCpGs at TFBS). In addition, enhancer regions also showed DMCpGs, but these differences corresponded with a small proportion of the CpGs (<4%) located at these regulatory regions (**Figure 11B**).



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

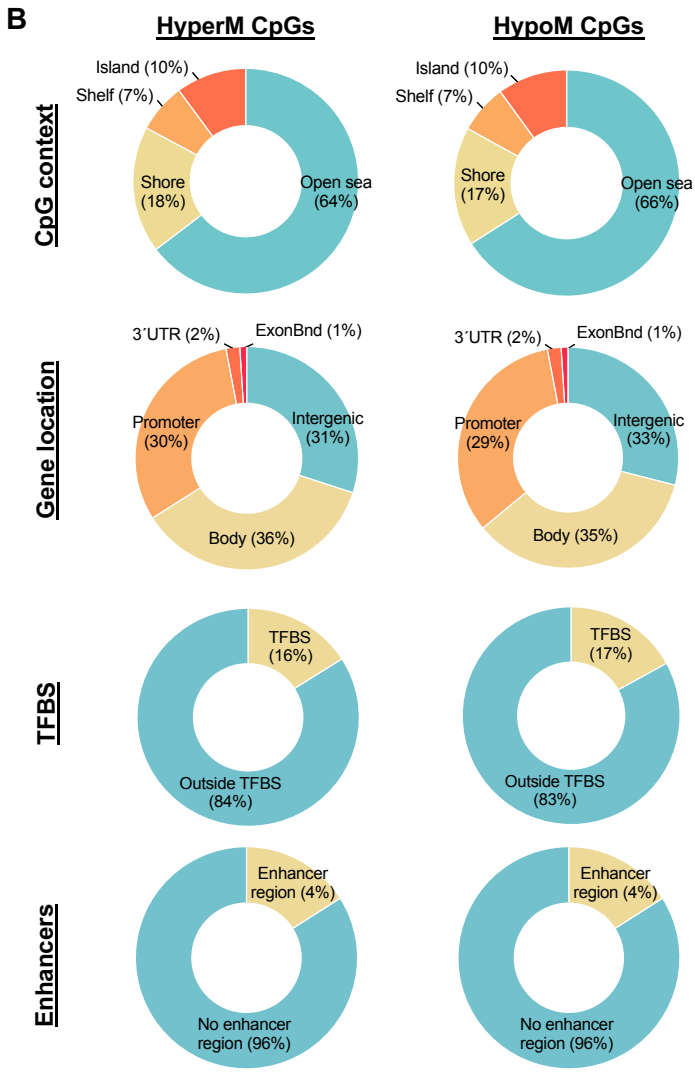


Figure 11. Description of the 188,185 DMCpGs found in CTC-MCC-41 respect to HT29. Distribution of the DMCpGs according to **(A)** chromosome location and methylation status, and **(B)** CpG context, gene location and TFBS. DMCpGs, differentially methylated CpGs; HypoM, hypomethylated; HyperM, hypermethylated; TFBS, transcription factor binding sites; ExonBnd, exon boundaries.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

On the other hand, we also focused our analysis in CpGs located at CpGIs and shores of gene promoters, regions in which epigenetic deregulation has been described as a very relevant feature in cancer[220,221]. In accordance with this, we identified 27,378 DMCpGs located in CpGIs or shore regions of gene promoters between CTC-MCC-41 and HT29, that correspond to 10,579 genes (**Figure 10A**). Moreover, using hierarchical clustering analysis we identified a methylation profile at these CpG-rich promoter regions able to clearly differentiate CTC-MCC-41 from HT29 tumour cells (**Figure 12A**). Top 20 DMCpGs from CpGIs or shore regions of gene promoters are represented in **Supplementary Table 2**. In addition, we carried out a GO and Panther pathway analysis with the differentially methylated genes from these CpG-rich promoters which revealed an enrichment of relevant biological processes such as regulation of transcription, cell adhesion and differentiation, apoptosis, cell cycle and proliferation, metabolic processes and DNA repair, among others; and pathways like Wnt signalling, cadherin pathway, inflammation, angiogenesis and integrin signalling, all related to cancer development and metastasis (**Figure 12B**).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

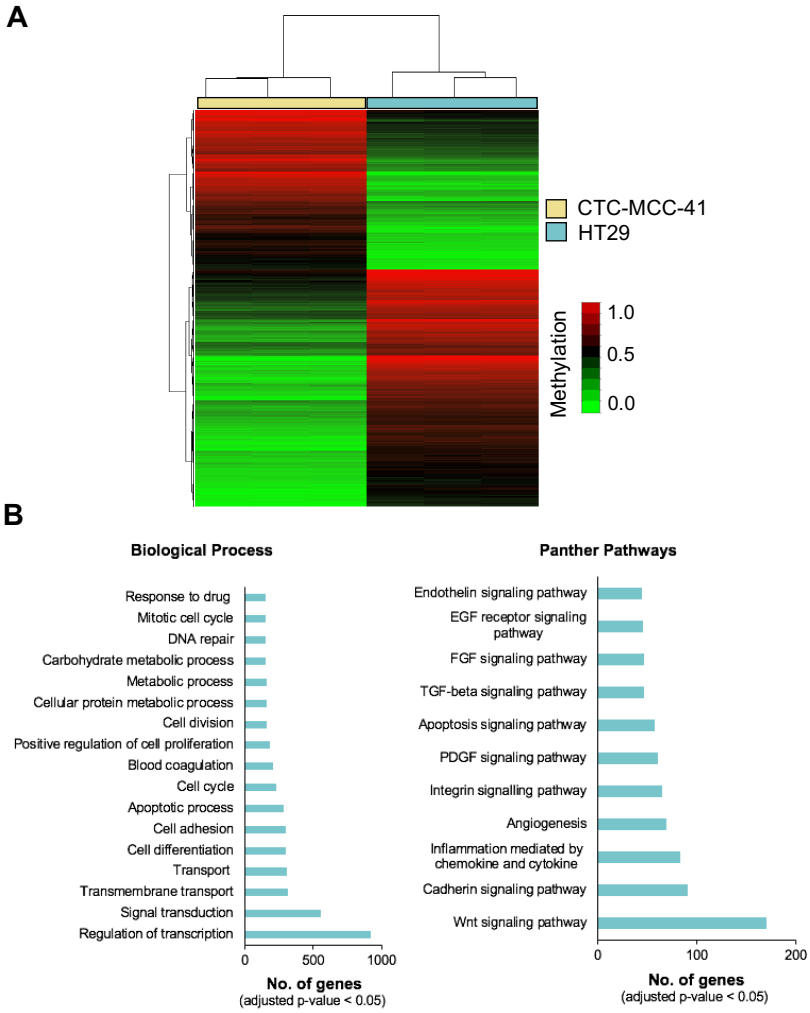
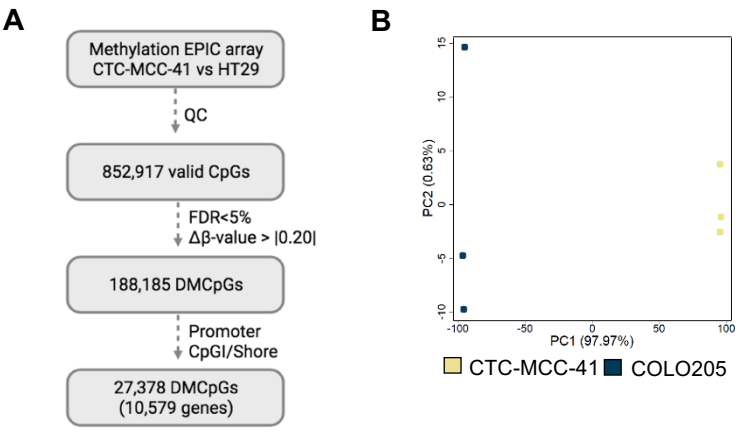


Figure 12. DNA methylation pattern of CpGIs and shore regions of gene promoters in CTC-MCC-41 compared to primary tumour cells HT29. (A) Hierarchical clustering heatmap of the 10,000 most DMCpGs (FDR adjusted p-value <0.05) in CTC-MCC-41 respect to HT29 and located at CpGIs and shore regions of gene promoters. **(B)** GO analysis representing some of the most cancer relevant biological processes and Panther pathways based on the 10,000 most DMCpGs of CTC-MCC-41 compared to HT29 and located at CpGIs and shore regions of gene promoters.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.2 Study of the DNA methylation profile of CTC-MCC-41 cells with respect to metastatic tumour cell line COLO205

In line with the previous analysis, the DNA methylation profile of CTC-MCC-41 was also evaluated respect to the metastatic CRC cell line COLO205 using the EPIC array technology, allowing the analysis of 852,917 valid CpGs (**Figure 13A**). The evaluation of the methylation levels of these CpGs using PCA and scatter plot (R^2 = 0.66) revealed that CTC-MCC-41 and COLO205 have different global methylation profiles (**Figures 13B-C**). Among the CpGs analyzed, we identified 196,748 significant DMCpGs (p-value <0.05; FDR adjusted p-value <0.05; Δβ-value >|0.20|) between CTC-MCC-41 and COLO205, and a hierarchical clustering analysis of these DMCpGs identified a methylation profile that clearly differentiates both cancer cell lines (**Figure 13D**). The top 20 DMCpGs in CTC-MCC-41 respect to COLO205 are represented in **Supplementary Table 3**.



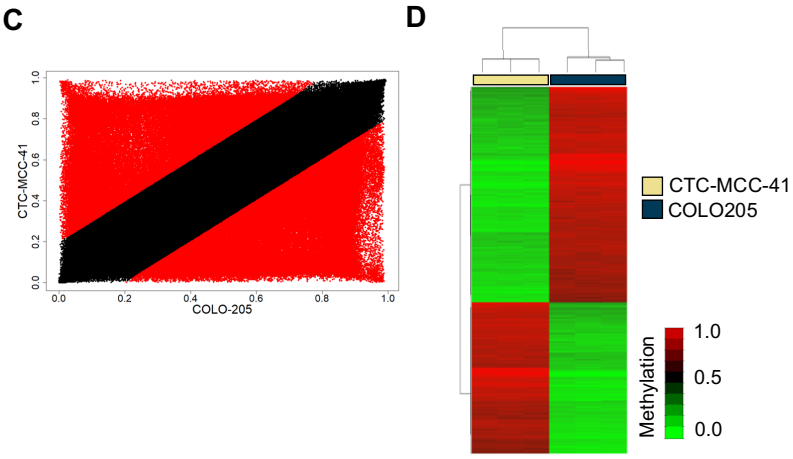
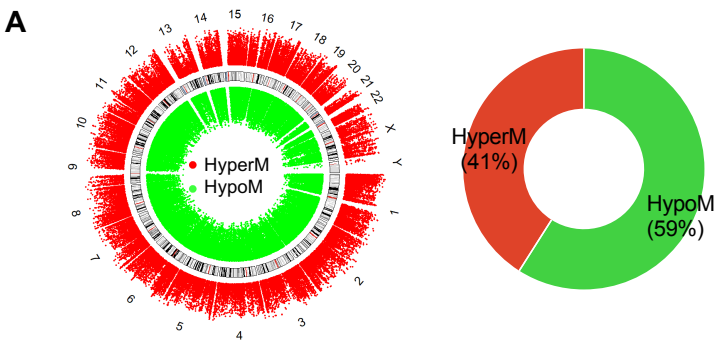


Figure 13. Genome-wide DNA methylation analysis of CTC-MCC-41 respect to the metastatic tumour cells COLO205. (A) Schematic flowchart used to identify significant differentially methylated CpGs in CTC-MCC-41 compared to COLO205. Three different passages of each cell line were analyzed. **(B)** Principal component analysis (PCA) of DNA methylation data obtained in CTC-MCC-41 and COLO205. **(C)** Scatter plot representing mean normalized levels of DNA methylation (β-values) in CTC-MCC-41 and COLO205. Dots in red show significant differentially methylated CpGs. **(D)** Hierarchical clustering heatmap of the 10,000 most differentialy methylated CpGs (FDR adjusted p-value <0.05) between CTC-MCC-41 and COLO205. QC, quality control; FDR, false discovery rate; CpGI, CpG island, HypoM, hypomethylated; HyperM, hypermethylated.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

The analysis of the 196,748 DMCpGs previously obtained, showed that methylation differences are distributed throughout all the genome. In line with our previous comparison between CTC-MCC-41 and HT29, most of the DMCpGs found in CTC-MCC-41 respect to COLO205 were hypomethylated (116,431 CpGs; 59% of all DMCpGs), while the remaining (80,317 CpGs; 41% of all DMCpGs) were hypermethylated (**Figure 14A**). These DMCpGs showed a similar distribution according to their CpG context and gene region location (**Figure 14B**), with most of the CpGs located in open sea, regions with low CpG density. Moreover, the distribution of these CpGs was fairly similar between promoters, gene bodies and intergenic regions (**Figure 14B**). In relation with TFBS, only a small percentage of DMCpGs were located at these sites, that corresponded with a higher number of hypomethylated CpGs (19,068 CpGs; 16% of the DMCpGs at TFBS) of CTC-MCC-41 cells than hypermethylated CpGs (13,468 CpGs; 17% of the DMCpGs at TFBS). In addition, a small percentage of the DMCpGs (<5%) between CTC-MCC-41 and COLO205 cells were located at enhancer regions (**Figure 14B**).



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

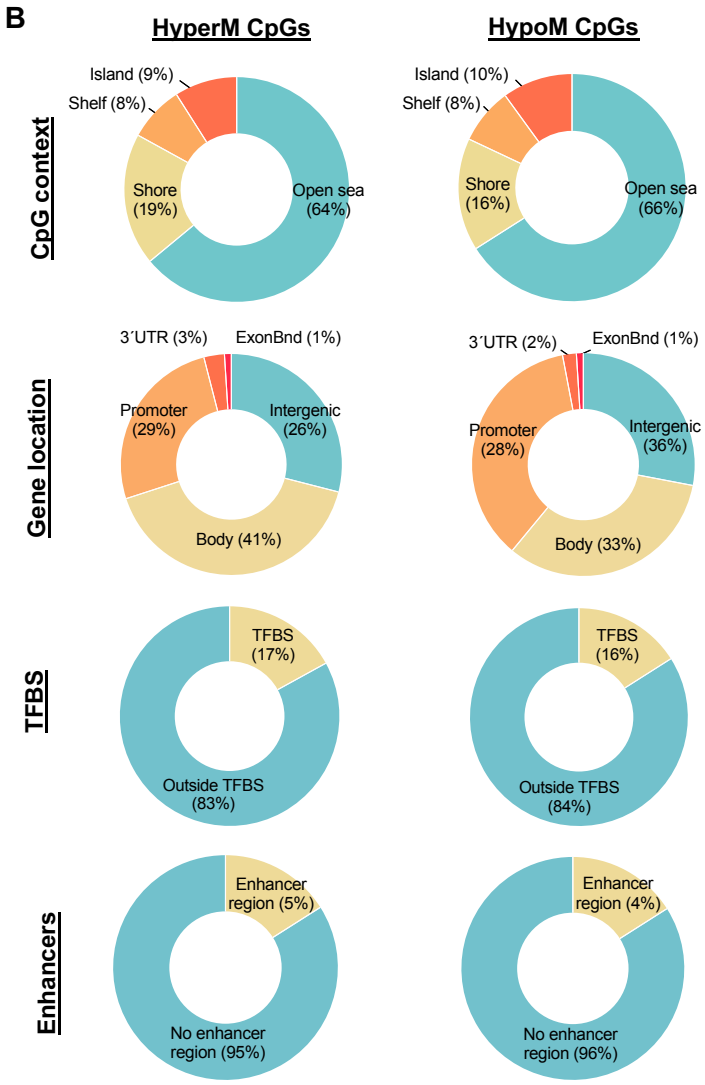
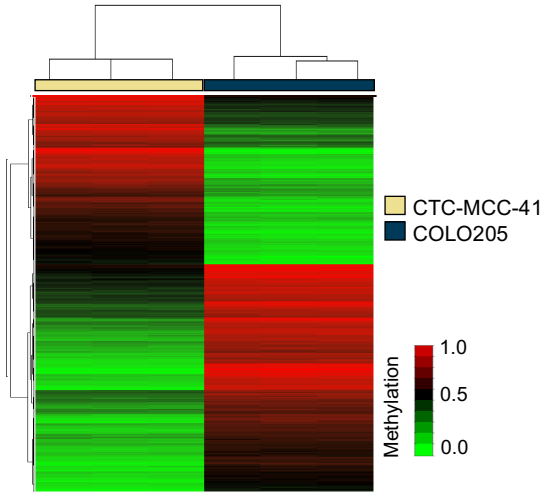


Figure 14. Description of the 196,748 DMCpGs found in CTC-MCC-41 respect to COLO205. Distribution of the DMCpGs according to **(A)** chromosome location and methylation status, and **(B)** CpG context, gene location and transcription factor binding sites (TFBS). FDR, false discovery rate; CpGI, CpG island, HypoM, hypomethylated; HyperM, hypermethylated; TFBS, transcription factor binding sites; ExonBnd, exon boundaries.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

When the analysis was focused at CpGIs and shores of gene promoters of CTC-MCC-41 and COLO205, we were able to identify 26,828 DMCpGs (**Figure 13A**), which corresponded with 10,569 genes. Using hierarchical clustering analysis, we observed that the CpGIs and shores of gene promoters in CTC-MCC-41 have a distinctive methylation profile to the COLO205 metastatic tumour cell line (**Figure 15A**). **Supplementary Table 4** shows the 20 most DMCpGs located at these regulatory regions. On the other hand, the GO analysis of the 10,569 differentially methylated genes revealed the enrichment of key biological process, including regulation of transcription, cell adhesion and differentiation, apoptosis, cell cycle and proliferation, metabolic processes and DNA repair. In addition, the Panther pathway analysis revealed an enrichment of important pathways related to cancer development and metastasis, such as, Wnt signaling, cadherin pathway, inflammation, angiogenesis and integrin signalling, among others (**Figure 15B**).

A



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

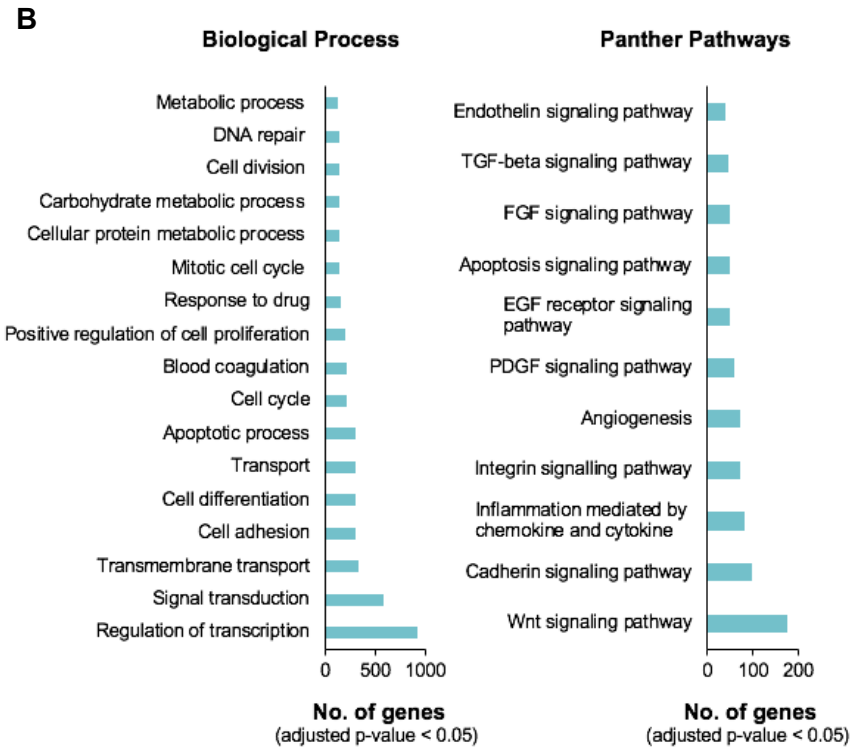


Figure 15. DNA methylation pattern of CpGIs and shore regions of gene promoters in CTC-MCC-41 compared to metastatic tumour cells COLO205. (A) Hierarchical clustering heatmap of the 10,000 most DMCpGs (FDR adjusted p-value <0.05) in CTC-MCC-41 respect to COLO205 and located at CpGIs and shore regions of gene promoters. **(B)** GO analysis representing some of the most cancer relevant biological processes and Panther pathways based on the 10,000 most DMCpGs of CTC-MCC-41 compared to COLO205 and located at CpGIs and shore regions of gene promoters.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.3 Analysis of the DNA methylation profile of CTC-MCC-41 cells compared to both primary and metastatic cancer cell lines

Once the DNA methylation profile of CTC-MCC-41 was evaluated in comparison with HT29 first and with COLO205 after, independently, we wondered if CTC-MCC-41 could have a specific methylation profile different from both primary and metastatic tumour CRC cell lines. Therefore, considering the most DMCpGs previously identified in CTC-MCC-41 respect to HT29 and COLO205, a hierarchical clustering analysis revealed an epigenetic signature of 17,827 CpGs that enable the distinction of CTC-MCC-41 from both HT29 and COLO205 cancer cells (**Figure 16A**). It is important to mention that the tree-based dendrogram obtained yielded two arms, one with the CTC-MCC-41 cell line and the other one with the primary and metastatic tumour cell lines.

Furthermore, to confirm the existence of a specific-CTC methylation signature in CRC, we included in our analysis other CRC cell lines representing primary (Caco2, HCT116, RKO) and metastatic distant (SW620) tumours. First, we analyzed the methylation levels of these cancer cell lines using the EPIC array technology, and we obtained the DMCpGs (Δβ-value >|0.20|) respect to CTC-MCC-41. Then, considering the DMCpGs obtained in CTC-MCC-41 respect to all primary (HT29, Caco2, HCT116, RKO) and metastatic (COLO205 and SW620) cancer cell lines analyzed, we obtained a specific methylation profile of 9,949 CpGs that enabled the classification of CRC cells according to their cell type: CTC, primary or metastatic tumour cell line (**Figure 16B**). It is important to note that the tree-based dendrogram provided two clear arms, CTC-MCC-41 was segregated in one branch and the other tumour cell lines in other branch with two well differentiated groups

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

containing all the primary cancer cell lines (HT29, Caco2, HCT116, RKO) and the metastatic ones (COLO205 and SW620).

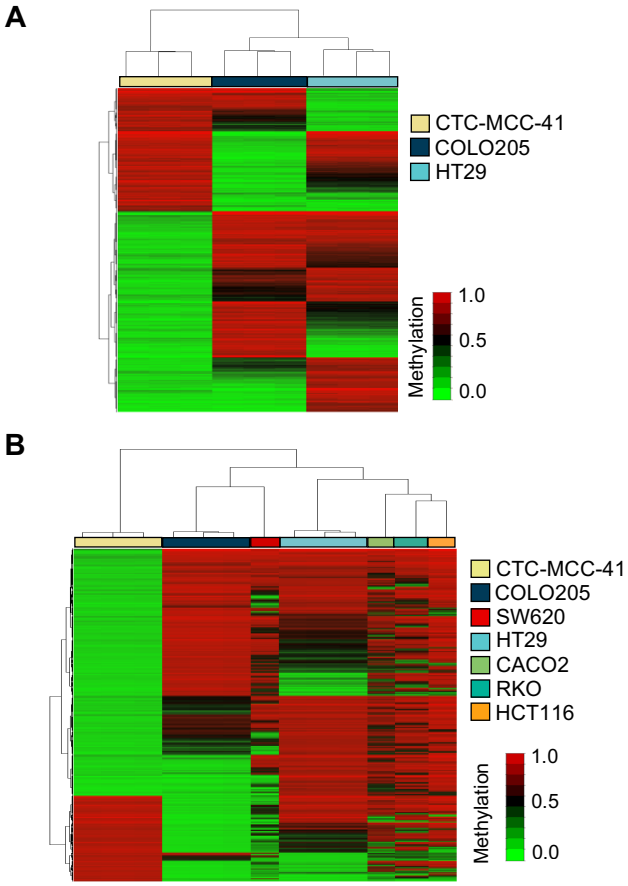


Figure 16. DNA methylation signature of CTC-MCC-41 respect to colorectal primary and metastatic tumour cells. (A) Hierarchical clustering heatmap with the 17,827 DMCpGs (FDR adjusted p-value <0.05) in CTC-MCC-41 compared to HT29 and COLO205, primary and metastatic tumour cells, respectively. **(B)** Hierarchical clustering heatmap of the 9,949 DMCpGs (FDR adjusted p-value <0.05) in CTC-MCC-41 compared to several colorectal

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

primary (HT29, Caco2, HCT116, RKO) and metastatic tumour cells (COLO205 and SW620).

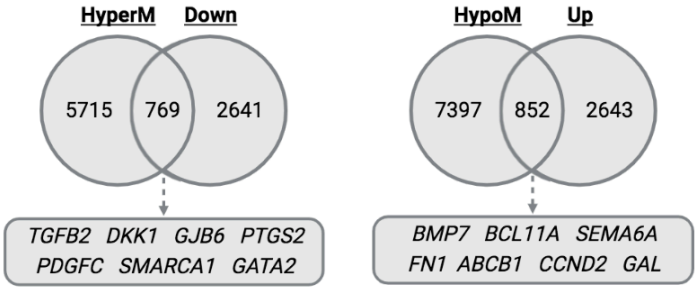
4.4 Evaluation of the association between the DNA methylation profile of CTC-MCC-41 cells and its transcriptional program

Data from a previous work published by Alix-Panabières *et al.*[209], where the transcriptional profile of CTC-MCC-41 in comparison to HT29 cancer cells was analyzed, revealed the differentially expressed genes (up- and down regulated) between both cancer cell lines. These data were used to test whether DNA methylation changes at CpGIs and shore regions of gene promoters were able to regulate the gene expression of metastasis-competent colorectal CTCs. In this sense, we evaluated the association between these differentially expressed genes and the differentially methylated genes in CpGIs and shore regions of gene promoters found in our analysis when we compared CTC-MCC-41 and HT29 cancer cell lines (**Figure 17A**). Venn diagram analysis enabled the identification of 769 genes whose hypermethylation was associated with transcriptional silencing, and 852 hypomethylated genes that were associated with transcriptional activation. Since expression of some of these genes had already been validated in the work of Alix-Panabières *et al*. [209], (downregulated in CTC-MCC-41: *TGFB2*, *DKK1*, *GJB6*, *PTGS2*, *PDGFC*, *SMARCA1*, and *GATA2*; upregulated in CTC-MCC-41: *BMP7*, *BCL11A*, *SEMA6A*, *FN1*, *ABCB1*, *CCND2*, and *GAL*), we focused our analysis on these to evaluate an association between methylation and expression. **Supplementary Tables 5-6** show the methylation levels of the hypermethylated and hypomethylated genes that were obtained in our genome-wide analysis by EPIC array and selected for validation. Methylation status of all the selected genes was validated using bisulfite pyrosequencing except for one, *GAL*, for which we used qMSP technology. In all the genes analyzed we were able to confirm the expected presence of hypermethylation or hypomethylation in CTC-MCC-41 respect to the HT29 cancer cell line. Only

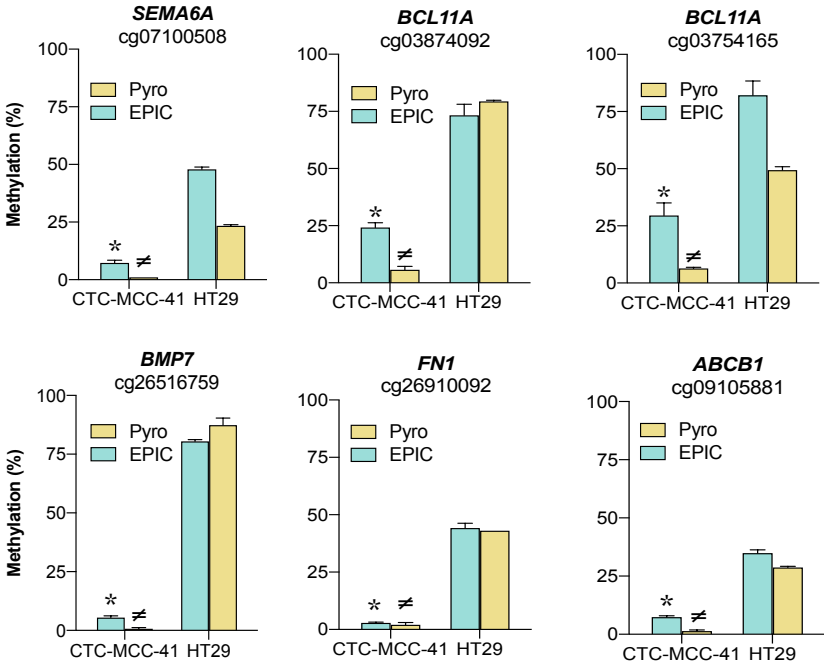
USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

in one gene, *GATA2*, the methylation level observed after bisulfite pyrosequencing in CTC-MCC-41 was not as high as the obtained by the EPIC array. **Figures 17B-C** show the validation of one representative CpG for each of the selected genes.

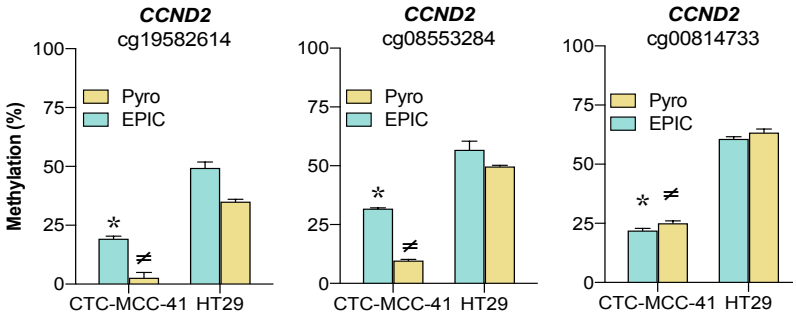
A



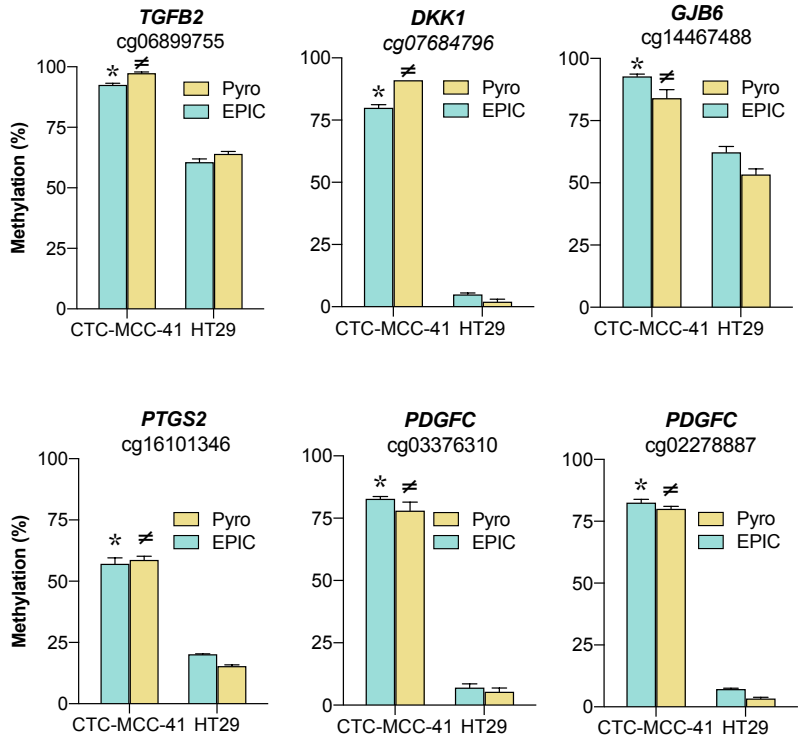
B



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



C



USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

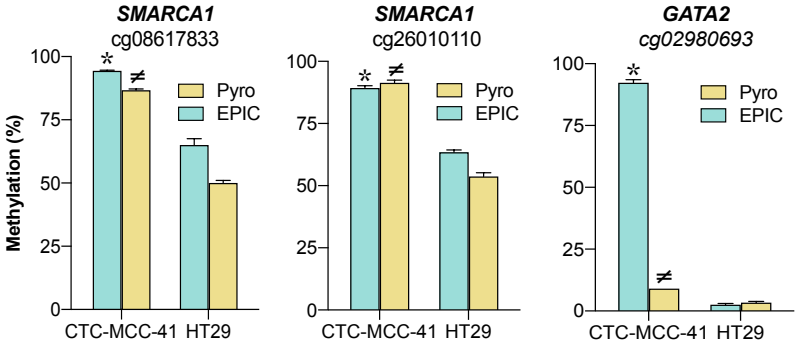


Figure 17. Impact of DNA methylation into the transcriptional program of CTC-MCC-41. (A) Association between the hypermethylated-downregulated genes and the hypomethylated-upregulated genes found in CTC-MCC-41 respect to HT29. HyperM, hypermethylated; HypoM, hypomethylated; Down, downregulated; Up, upregulated. **(B-C)** Validation of the methylation status of representative promoter CpGs from selected **(B)** hypomethylated-upregulated genes and **(C)** hypermethylated-downregulated genes. All the genes were analyzed by pyrosequencing except *GAL* that was analyzed by qMSP. EPIC array data are expressed as % of methylation. Methylation values are represented as mean ± SD. *, p-value <0.05 between CTC-MCC-41 and HT29 after analysis by EPIC array; ≠, p-value <0.05 between CTC-MCC-41 and HT29 after analysis by pyrosequencing or qMSP. Pyro, pyrosequencing.

Once the methylation levels of the selected genes in CTC-MCC-41 were confirmed, we wondered if the epigenetic program of the hypermethylated genes in colorectal CTCs could be reverted affecting its expression. For this purpose, we treated CTC-MCC-41 cells *in vitro* with the demethylating agent AZA, showing that the expression of the epigenetically silenced genes *TGFB2*, *DKK1*, *GJB6* and *PTGS2* was completely restored (**Figure 18**).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

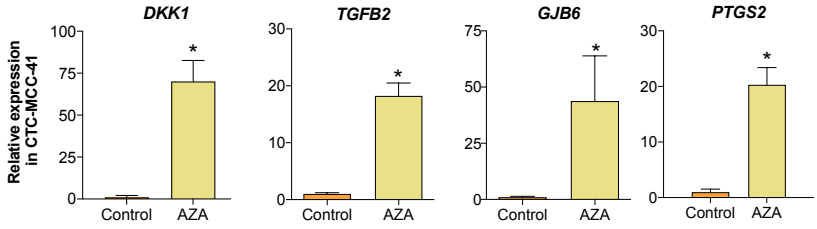


Figure 18. **Restored expression of hypermethylated genes (DKK1, TGFB2, *GJB6* and *PTGS2*) in CTC-MCC-41 after treatment with the demethylating agent 5-aza-2′-deoxycytidine (AZA).** Gene expression was analyzed by RT-qPCR. Expression values were determined from triplicates and are expressed as mean ± SD. *, p-value <0.05 between control and AZA treatment of CTC-MCC-41 cells.

In addition to these validated genes, it is noteworthy that among the 852 hypomethylated and downregulated genes identified in CTC-MCC-41, we also found genes involved in pathways relevant to cancer metastasis, such as Wnt signaling (e.g., *FZD5*, *LEF1*, *ACTA2*, *ARRB1*), the cadherin pathway (e.g., *CDH3*, *LYN*, *PCDHA1*, *PCDHA4*, *PCDHA5*, *PCDHA7*, *PCDHA12*), inflammation (e.g., *IKBKB*, *GNG4*, *CASK*) or angiogenesis (e.g., *DLL1*, *LPXN*).

5. Discussion

Epigenetic mechanisms regulate and contribute to every single hallmark in cancer, representing a major hallmark in tumourigenesis[222]. Disruption of epigenetic mechanisms, such as DNA methylation, can lead to altered gene function and neoplastic transformation[143], having relevant implications for CRC development and progression[137,223].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Also relevant in cancer progression are CTCs, since their dissemination from primary or metastatic tumours may colonize distant tissues and form metastasis[128,206], which are responsible for the majority of cancer-related deaths[224]. The analysis of the epigenome of these metastasis-competent CTCs may help not only to unravel important cancer biological processes, such as metastasis, but also to discover new clinical applications[69]. In this sense, quite recently, genome-wide sequencing studies have helped to study DNA methylation changes in metastasis on a broader scale[225]. Thus, in this study we compared the DNA methylome of the human colorectal metastasis-competent cell line CTC-MCC-41 with both primary (HT29, Caco2, HCT116, RKO) and metastatic (SW620 and COLO205) CRC cell lines, revealing that metastasis-competent CTCs display a completely different methylation program to that displayed by primary and metastatic tumour cells.

Human cancer-derived cell lines have demonstrated to be, since the first immortalized cell line (HeLa) was established[226], fundamental models used in laboratories to understand the biology of cancer and pursue clinical applications[137,227,228]. Recent findings, based on the characterization of hundreds of cell lines with omics technologies, reinforced the concept of cell line usefulness in medical research[229]. In this sense, epigenomic characterization of CTC lines, such as CTC-MCC-41, represents a great opportunity to get an insight into the biology of this subset of CTCs and also to provide clinical applications in cancer patients. Particularly, this CTC-MCC-41 cell line represents a unique model in colon cancer, as it is the only established CTC line obtained from a metastatic patient prior to antitumour therapy[140].

In the present study, the epigenomic analysis of CTC-MCC-41 revealed that metastasis-competent colorectal CTCs have a different global DNA methylation profile than both primary and metastatic cancer cells, and

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

are defined by a distinctive methylation signature respect to both tumour cell types. Although the global methylation differences observed between hyper- and hypomethylation were fairly similarly distributed, CTCs showed a slight predominance of hypomethylated CpGs. This hypomethylation was mainly distributed in CpG-poor regions, suggesting that the epigenetic regulation of this part of the genome may be important for metastasis-competent CTCs. In cancer, the hypomethylation of these regions has been associated with proto-oncogene expression, chromosomal instability, and malignant transformation of tumours[230,231]. In line with this, other studies have revealed that the cancer cell line CTC-MCC-41[140], and CTCs obtained from peripheral blood of patients from different cancer types[232,233], are characterized by the expression of proto-oncogenes and the presence of chromosomal aberrations. On the other hand, in relation with CpGs located at TFBS, we observed that the number of hypomethylated CpGs associated with TFBS in CTC-MCC-41 was higher than the hypermethylated ones. In this sense, a previous work identified that CTCs may undergo methylation changes in CpGs from TFBS[119]. Interestingly, although in a small proportion, part of the methylation differences were located at CpGs of enhancers, which are epigenetically regulated regions involved in cancer cell plasticity and tumour progression[234].

Aberrant methylation of promoter CpGIs and shore regions is frequently observed in cancer[200,235]. In this work, results evidence that promoter CpGIs and shore regions (CpG-rich promoters) of CTC-MCC-41 are characterized by a different methylation profile than that showed by primary and metastatic cancer cells. In this sense, aberrant promoter methylation of some particular genes, like *VIM* and *SFRP2*, has been recently described in CTCs from CRC patients[198]. Notably, the methylation differences that we found in CpGI and shore regions of promoters of CTCs were highly associated with transcriptional regulation, which is a relevant characteristic of epigenetic mechanisms[200]. Many of these methylation differences were also associated

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

with Wnt signaling pathway, suggesting that the alteration of this pathway is a relevant characteristic of metastasis-competent CTCs. In many malignancies, including CRC, the alteration of Wnt pathway is associated with the stemness of cancer cells[236], which is also a property previously described in CTC-MCC-41 cells[140]. Moreover, changes observed in CpGIs and shore regions of promoters of CTC-MCC-41 showed relation with other relevant biological processes and pathways, including apoptosis, metabolism and DNA repair, among others. In this sense, a previous work in CTC-MCC-41 showed that the transcriptional program of this cell line is associated with changes in the expression of genes involved in these pathways[209].

In cancer, aberrant DNA methylation of CpGI and shore promoters is usually linked to changes in gene expression, with a negative correlation between methylation and gene expression[137,237]. In line with this, the results obtained evidenced multiple genes of CTC-MCC-41 with a negative association between their DNA methylation and their transcriptional levels. Importantly, in several selected genes we confirmed their promoter hypermethylation (*TGFB2*, *DKK1*, *GJB6*, *PTGS2*, *PDGFC*, *SMARCA1*, *GATA2*) or hypomethylation (*BMP7*, *BCL11A*, *SEMA6A*, *FN1*, *ABCB1*, *CCND2*, *GAL*). One important characteristic of epigenetic mechanisms in cancer, including DNA methylation, is that they can be reversed to restore gene function by epidrugs such as the demethylating agent AZA. In this sense, AZA was able to reverse the epigenetic silencing of some of the hypermethylated genes in CTC-MCC-41, including *DKK1*, *TGFB2*, *GJB6* and *PTGS2*, confirming the epigenetic regulation of these genes by DNA methylation. Altogether these results indicate that DNA methylation is a regulator of the transcriptional program in metastasis-competent CTCs.

In addition, to provide relevant biological information, the aberrantly methylated genes identified in this work also represent potential biomarkers and therapeutic targets for metastasis-competent CTCs. For example, we

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

have identified the hypermethylation of *DKK1*, which is a tumour suppressor gene that works as a potent inhibitor of the canonical Wnt signaling pathway, avoiding epithelial-mesenchymal transition, cell proliferation and metastasis in cancer [238,239]. *DKK1* has been found hypermethylated in several types of tumours and its tumour suppressor activity can be reversed using demethylating agents[240,241], suggesting that the hypermethylation of this gene may be a new therapeutic target in metastatic-competent CTCs. In line with this, the use of DNA demethylating drugs, such as guadecitabine (SGI-110), is under study in metastatic CRC patients[242]. On the other hand, the hypermethylation of *DKK1* has been also identified as a prognostic factor in cancer patients [243,244], indicating that the hypermethylation of this gene in CTCs could have clinical implications as a prognostic biomarker.

Furthermore, we have identified other potential therapeutic targets in metastasis-competent CTCs that are involved in the activation of Wnt signaling pathway. This is the case of the genes *FZD5*, *LEF1*, *ACTA2*, and *ARRB1,* which showed transcriptional activation associated to hypomethylation in the CTC-MCC-41 cell line. Importantly, therapeutic targeting of the Wnt pathway is a promising strategy under study to impair cancer cells in several tumours including CRC[245,246]. This is the case for example of vantictumab (OMP-18R5), which is a fully human monoclonal antibody that inhibits the Wnt pathway through the binding of several FZD receptors, including FZD5[247], that has been evaluated in clinical trials with solid tumours (NCT01973309, NCT01957007, NCT02005315). In line with this, hypomethylation of *FZD5* could be a potential biomarker to select the patients who would benefit from anti-FZD5 therapies. Similarly, the inhibition of LEF1, which is a transcription factor with a central role in Wnt pathway, has shown antitumoural properties in several preclinical studies[248]. In line with this, we have identified in CTC-MCC-41 cells other genes with hypomethylated promoters involved in cadherin pathway (*CDH3*, *LYN*, *PCDHA1*, *PCDHA4*,

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

PCDHA5, *PCDHA7*, *PCDHA12*), inflammation (*IKBKB*, *GNG4*, *CASK*) or angiogenesis (*DLL1*, *LPXN*), suggesting that these genes could also be therapeutic targets for metastasis-competent CTCs in CRC patients. Importantly, the time this type of CTCs survive in the vascular system represents one of the most critical parameters that controls the formation of clinical metastases, indicating that the use of therapies, which are able to reduce the survival of this type of CTCs in the vascular system, could significantly reduce the risk of metastasis[207].

To summarize, in this study, the characterization of the epigenomic profile of CTC-MCC-41 revealed that metastasis-competent CRC CTCs display a globally different DNA methylation program than primary and metastatic tumour cells. These CTCs also have a specific DNA methylation profile of their CpG-rich promoters that has the ability to regulate the transcriptional program of these cells and it is associated with relevant cancer pathways. Besides, the aberrantly DNA methylated genes identified in this work represent potential biomarkers and therapeutic targets of metastasis-competent CTCs that may contribute to the development of new specific therapies directed against this type of CTCs in CRC.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

CHAPTER II

Epigenomic evaluation of colorectal CTC cell lines in association with therapy response

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Epigenomic evaluation of colorectal CTC cell lines in association with therapy response

1. Abstract

Resistance to therapy is a major limitation to the efficacy of CRC treatment, and epigenetic modifications, such as DNA methylation, are regulatory mechanisms driving drug resistance in cancer. Liquid biopsy approaches can capture modifications of the landscape of tumour cells under selective pressure of treatment. In line with this, molecular characterization of CTCs has potential to provide biomarkers to guide treatment decisions and mechanistic insight into drug resistance. In this chapter, we compared, using the EPIC array technology, the DNA methylation profile of a colorectal CTC cell line before fluoropyrimidine-based treatments initiation and different colorectal CTC lines after resistance development to this therapy. The association between the DNA methylation and expression pattern of these CTC lines was also analyzed. Relevant epigenetically regulated genes identified in resistant CTCs were validated using pyrosequencing. Expression analysis were carried out with RT-qPCR. Pre-treatment and resistant CTCs showed a different global DNA methylation profile, characterized by the presence in resistant CTCs of a more hypomethylated pattern. These resistant CTCs showed a distinctive DNA methylation signature related to relevant cancer pathways, such as PI3K/AKT, MAPK and metabolism. Several epigenetically deregulated genes, mainly *AP2M1*, associated with therapy resistance were identified in these CTC lines. Results obtained in this study delve into the epigenomic landscape of colorectal therapy resistance mediated by CTCs, providing novel mechanisms of resistance, potential biomarkers and therapeutic targets for advanced CRC.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

2. Introduction

Despite the advances on CRC treatment, the efficacy of therapy is still a major challenge in clinical oncology[249]. Some cancer patients fail to respond to anticancer drugs as a result of either intrinsic or acquired resistance[43,249]. In line with this, these mechanisms of resistance may be mediated by epigenetic changes, including DNA methylation alterations, which have been described in association with therapy resistance in CRC patients[250].

The longitudinal study of tumours after diagnosis and throughout the course of the disease is especially important to design effective therapeutic approaches that can improve cancer patient's survival[43]. Different resistance mechanisms can arise simultaneously within the same patient, and the analysis of a single-lesion biopsy, can fail to detect clinically relevant molecular alterations that can be responsible for treatment failure[251]. In this sense, CTCs and other liquid biopsy elements have the potential of detecting simultaneous resistance mechanisms coexisting in a single patient[251]. Identification of these mechanisms can give patients the opportunity to benefit from therapies designed to overcome resistance[43,251]. In this sense, molecular profiling of CTCs, including CTC lines, is a relevant approach to study the disease progression and to discover new biomarkers for monitoring treatment response[252,253].

Recently, a unique collection of nine human CTC-derived cell lines were established from the peripheral blood of a mCRC patient at different time-points during treatment and cancer progression: CTC-MCC-41, CTC-MCC-41.4, CTC-MCC-5A, CTC-MCC-5B, CTC-MCC-5C, CTC-MCC-5D, CTC-MCC-5E, CTC-MCC-5F and CTC-MCC-5G[139,140]. These cell lines were isolated from samples withdrawn before treatment initiation (CTC-MCC-41); after the first-line treatment with FOLFIRI and bevacizumab and second-line

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

treatment with FOLFOX and bevacizumab (CTC-MCC-41.4); and one week before patient´s death (CTC-MCC-41.5 [A-G]). Although these nine cell lines presented common traits, showing all epithelial-to-mesenchymal plasticity and stem cell-like characteristics, their molecular profiling at transcriptomic level, revealed expression changes linked to tumour progression and treatment pressure[253].

Thus, we wonder whether the DNA methylation portrait of these CTCs could be also associated with tumour progression and therapy response of CRC. For this, in this chapter we analyzed and compared the DNA methylation landscape of these nine CTC cell lines using a genome-wide methylation approach. Moreover, we also aimed to discover relevant biological information of CTCs related to mechanisms of resistance, as well as new epigenetic biomarkers linked to the progression of the disease after 5-fluorouracile-based therapies in CRC.

3. Materials and methods

3.1 Cell culture

For this study, besides the CTC cell line CTC-MCC-41, other eight CRC CTC cell lines were included: CTC-MCC-41.4 and CTC-MCC-5A, CTC-MCC-5B, CTC-MCC-5C, CTC-MCC-5D, CTC-MCC-5E, CTC-MCC-5F and CTC-MCC-5G. All this cell lines grow in suspension and were cultured using Corning Ultra-Low attachment cell culture flasks (Merck, Germany), and RPMI-1640 medium (Merck, Germany) supplemented with 10% Fetal bovine serum (FBS) (Merck, Germany), 1% penicillin/streptomycin solution (Gibco, Thermo Fisher Scientific, USA), 1% L-glutamine (Merck, Germany), 1% of Insulin-Transferrin-Selenium (ITS-G) (Gibco, Thermo Fisher Scientific, USA), basic human fibroblast growth factor (bFGF) (Miltenyi Biotec, Germany) at a final concentration of 10 ng/ml and human epidermal growth factor (hEGF)

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

(Miltenyi Biotec, Germany) at a final concentration of 20 ng/ml. Cells were grown at 37 °C, 5% CO_2 and culture medias were changed once or twice a week.

For cell recovery, cells were collected once reached confluency and they were centrifuged 8 min at 600 x g, keeping both a dry pellet for DNA extraction and a pellet resuspended in 1 mL of Trizol (Invitrogen, USA) for RNA extraction at -80 ºC.

3.2 Nucleic acids extraction and quantification

3.2.1 DNA extraction

DNA extraction was performed as established in the point *3.3.1 DNA extraction* of Chapter I. Briefly, cell pellets were resuspended in lysis buffer, SDS 10% and proteinase K and incubated overnight at 37 ºC. After cell lysis, RNAse A was added and incubated 1h at 37 ºC and also NaCl was added before centrifugation. The supernatant was transferred to a new tube, and isopropanol was added for DNA precipitation before centrifugation and washing the pellet with 70% ethanol and resuspend it in nuclease free water.

3.2.2 DNA quantification

DNA quantification was performed as established in the point *3.3.2 DNA quantification* of Chapter I. Briefly, samples were quantified by the fluorometric method Qubit 4 (Invitrogen, USA), using the Qubit dsDNA BR (Broad Range) Assay Kit (Invitrogen, USA), following manufacturer´s instructions.

3.2.3 RNA extraction

RNA extraction was performed as established in the point *3.3.3 RNA extraction* of Chapter I. Briefly, RNA extraction from cells was carried out using the TRIzol Reagent (Invitrogen, USA), unfreezed samples were incubated at

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

room temperature to allow dissociation of the nucleoproteins complex and chloroform was added before centrifugation to separate the RNA. The aqueous phase, containing the RNA, was transferred to a new tube and isopropanol was added for RNA precipitation. After centrifugation, the pellet was washed with 75% ethanol and resuspended in RNase-free water.

3.2.4 RNA quantification

RNA quantification was performed as established in the point *3.3.4 RNA quantification* of Chapter I. Concentration and purity of all RNA samples was determined using the spectrophotometer NanoDrop One (Thermo Fisher Scientific, USA).

3.3 DNA methylation analysis

3.3.1 Bisulfite treatment

Bisulfite treatment was performed as established in the point *3.4.1 Bisulfite treatment* of Chapter I. Briefly, 500 ng of each sample was treated with sodium bisulfite using the kit EZ DNA Methylation kit (Zymo Research, USA) following manufacturer's instructions for the Illumina Infinium Methylation Assay (Illumina, USA).

3.3.2 Infinium MethylationEPIC array

Infinium methylation EPIC Array was performed as established in the point *3.4.2 Infinium MethylationEPIC array* of Chapter I. Briefly, DNA methylated samples were hybridized to the Infinium MethylationEPIC BeadChips, following the Infinium HD methylation assay and according to the manufacturer´s instructions. Once the BeadChips were hybridized and stained, they were scanned with a HiScan System.

The methylation data were processed in the R statistical environment using RnBeads 2.0[254]. Raw intensity data files (IDATs) were imported into

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

RnBeads 2.0 for quality control and preprocessing. First, a greedycut algorithm was used to filter out low-quality probes. Probes overlapping with SNPs and probes whose sequences mapped to multiple genomic locations (cross-reactive) were removed. IDATs obtained in the array were normalized using the beta-mixture quantile (BMIQ) method. The DNA methylation level was represented as the average β-value, which was calculated as the ratio of the fluorescent signal intensity of the methylated probe to those of total (methylated and unmethylated) probes. Average β-values were used to calculate the mean methylation difference between groups. Unsupervised hierarchical clustering heatmaps of β-values were generated using the ComplexHeatmap package. The enrichment analysis of biological pathways was evaluated by Gene Ontology (GO) using Genecodis[255].

3.3.3 Validation of the Infinium MethylationEPIC array data

3.3.3.1 Pyrosequencing

Pyrosequencing was performed as established in the point *3.4.3.1 Pyrosequencing* of Chapter I and the design of the primer sequences (**Table 5**) was performed with PyroMark Assay Design 2.0 (Qiagen, Germany). Briefly, bisulfited-converted DNA of the samples was previously amplified performing standard PCR for all candidate genes using a ProFlex PCR System (Thermo Fisher Scientific, USA). Sequencing reactions were performed using a PyroMark Q24 System (Qiagen, Germany) with the PyroMark Gold SQA™ Reagent Kit (Qiagen, Germany) following manufacturer's instructions. Methylation values of the CpG sites of interest were obtained using PyroMark Q24 Software 2.0 (Qiagen, Germany). Methylation assays were conducted in triplicates and 100% Human Methylated & Non methylated DNA (Zymo Research, USA) were used as positive and negative controls respectively and water as no template control.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Table 5. Primer sequences for DNA methylation analysis by pyrosequencing.

Gene	Primer	Sequence (5´- 3´)
AP2M1	Forward	TGTTTGGGGGTTTTAGGTTTGAATTA
	Reverse	[Btn] CTCAACCAAACTCCAATTCCTATTTATTC
	Sequencing	AGGTAAGTTTTAGTGAGTTAT
BDNF	Forward	GGGGTTGGAAGTGAAAATATTTGTAAAAG
	Reverse	[Btn] ACCCCATCAACRAAAAACTCCATT
	Sequencing	GGTAGYGGAGGTATTATATGATAG
COL9A3	Forward	TTTTTGGGAAATGGATAT
	Reverse	[Btn] TCAAACCACRTCAATTACTA
	Sequencing	ATTTYGGGGAGGGGT
FKBP11	Forward	GGAGGAAGGGTTGTAGGA
	Reverse	[Btn] TCACCAAAATCTCCACTTAA
	Sequencing	GTTTAGTTTGGTTTGGG
GALNT6	Forward	AGGTATTGGGTGAGGTTT
	Reverse	[Btn] TTCCCTATCTCATCCTTCTT
	Sequencing	ATAGTTTTGGATATATTAAAGAGT
NR4A1	Forward	TAGGGAAGGTTTTAGGTG
	Reverse	[Btn] ATACCCCTCCCCCTAT
	Sequencing	GTTTGTGTTTAGTAGGGT

Btn, biotin.

3.4 Expression analysis

3.4.1 DNAse treatment and retrotranscription

DNAse treatment and retrotranscription were performed as established in the point *3.5.1 DNAse treatment and retrotranscription* of Chapter I. Briefly, RNA was purified using the Turbo DNA-free Kit and retrotranscription of this purified RNA was performed using the SuperScript

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

First-Strand Synthesis System for RT-PCR (Invitrogen, USA) both following manufacturer´s recommendations.

3.4.2 Quantitative reverse transcription PCR (RT-qPCR)

Expression analysis by RT-qPCR was performed as established in the point *3.5.2 Quantitative reverse transcription PCR (RT-qPCR)* of Chapter I. Briefly, RT-qPCR was performed with 25-200 ng of cDNA using the reagent *Power* SYBR Green PCR Master Mix (Applied Biosystems, USA) in a Step One Plus thermocycler system. All reactions were performed in triplicates for each target, using the gene beta-2-microglobulin (*B2M*) as a housekeeping. Primer sequences for target genes were designed with the Primer3 tool[217] except for *B2M* for which the same primer sequences as in Chapter I were used. Primer sequences are indicated in **Table 6**. Data analysis was performed based on the relative quantification of the $2^{-\Delta\Delta CT}$ method[219].

Table 6. Primer sequences for the RT-qPCR analysis.

Gene	Primer	Sequence (5´- 3´)
AP2M1	Forward	AGATGACATCGGGAGGAAC
	Reverse	CCATCACGTCACACATCTTA
BDNF	Forward	AGAAAAAAACACACACACACAC
	Reverse	AGCCAGTAAAGCAATGACAA
COL9A3	Forward	GGCATTGACGGAGAAGCTG
	Reverse	GTCCAGTCAGACCATCCAC
FKBP11	Forward	ATACACTACACGGGAAGCTTGGTA
	Reverse	GAAGACTCTGCTCCAGACCT
GALNT6	Forward	TGGCCCAAGAGAAATCCT
	Reverse	TCAGGTCAGGAACAAACATC
NR4A1	Forward	AAGGAAGTTGTCCGAACAGA
	Reverse	AGCTCCTGGAACTTGGAGTA

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.5 Statistical analysis

Hierarchical linear models performed with RnBeads 2.0 were used to obtain the methylation differences between groups. P-values were corrected for multiple testing (FDR) using Benjamini-Hochberg method and a threshold of p-value <0.05 was selected for significance. To compare methylation and expression values obtained with pyrosequencing and RT-qPCR, normality distribution of data was tested first using the Kolmogorov–Smirnov test following either the Mann–Whitney U test or Student´s t -test when corresponded. The GraphPad Prism 8.0 software (GraphPad Software, USA) and the R statistical environment (version 4.2.0) were used for statistical analysis and graphical representation. All expressed p-values were calculated with two-tailed tests and were considered significant when the p-value <0.05.

4. Results

4.1 Evaluation of the DNA methylation profile of the pre-treatment cell line CTC-MCC-41 respect to post-treatment CTC cell lines

The DNA methylation profile of the pre-treatment cell line CTC-MCC-41 was evaluated respect to the eight post-treatment CRC cell lines CTC-MCC-41.4 and CTC-MCC-41.5 [A-G] using the Infinium MethylationEPIC array technology. After normalization and performing the quality control of the data, we obtained 803,354 valid CpGs for the analysis (**Figure 19A**). The comparison between the pre- and post-treatment CTC cell lines with a scatter plot revealed a different methylation profile between these CTC cell lines (**Figure 19B**). In particular, we identified 34,085 significant DMCpGs (p-value <0.05; FDR adjusted p-value <0.05; Δβ-value >|0.20|) between CTC-MCC-41 and all post-treatment CTC cell lines, from which 4,018 DMCpGs were located in promoters of CpGIs and shore regions.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

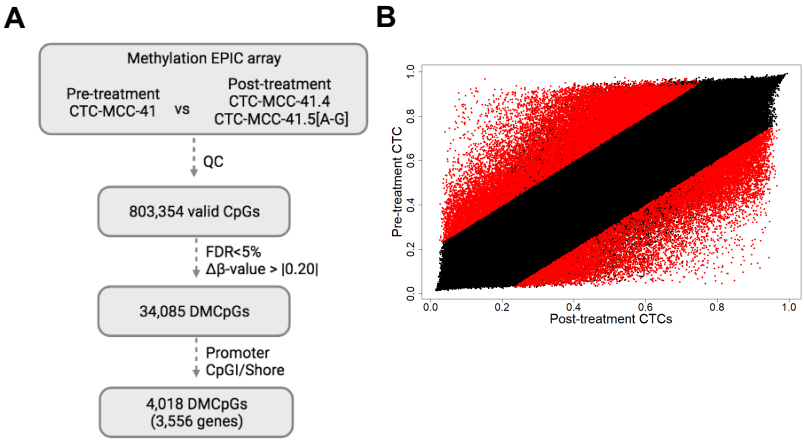
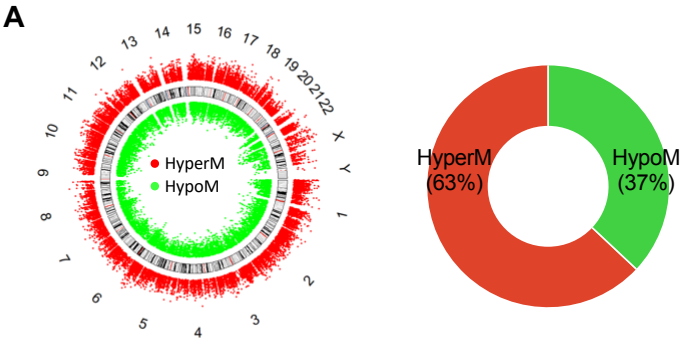


Figure 19. Genome-wide DNA methylation analysis of the pre-treatment cell line CTC-MCC-41 respect to the post-treatment cell lines CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]. (A) Schematic flowchart used to identify significant DMCpGs in CTC-MCC-41 compared to CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]. **(B)** Scatter plot representing mean normalized levels of DNA methylation (β-values) in CTC-MCC-41 and post-treatment cell lines CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]. Dots in red show significant DMCpGs. QC, quality control; FDR, false discovery rate; CpGI, CpG island.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

These 34,085 DMCpGs previously identified showed a wide distribution throughout all the chromosomes of the genome, showing a higher number of hypermethylated CpGs (21,409 CpGs; 63% of all DMCpGs) than hypomethylated CpGs (12,676 CpGs; 37% of all CpGs) in CTC-MCC-41 cells respect to all post-treatment CTC cell lines (**Figure 20A**). Both hyper- and hypomethylated CpGs in CTC-MCC-41 showed a similar distribution according to their gene region location and CpG context (**Figure 20B**). In particular, most of these DMCpGs (71% of hyper- and 65% of hypomethylated CpGs) were located in open sea regions, characterized for its low CpG density. Regarding gene region, both hyper- and hypomethylated CpGs were fairly homogeneously distributed throughout promoters, gene bodies and intergenic regions. In addition, we also observed, although in a small percentage of CpGs, methylation differences between CTC-MCC-41 and post-treatment cells in CpGs located at TFBS, corresponding with 2,369 hypomethylated CpGs (19% of all the DMCpGs at TFBS) and 3,591 hypermethylated CpGs (17% of all the DMCpGs at TFBS) in CTC-MCC-41. Moreover, enhancer regions also showed DMCpGs, corresponding these differences with a small proportion of the CpGs located at these regulatory regions being a 5% of the hypermethylated CpGs and a 3% of the hypomethylated CpGs of CTC-MCC-41 cells (**Figure 20B**).



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

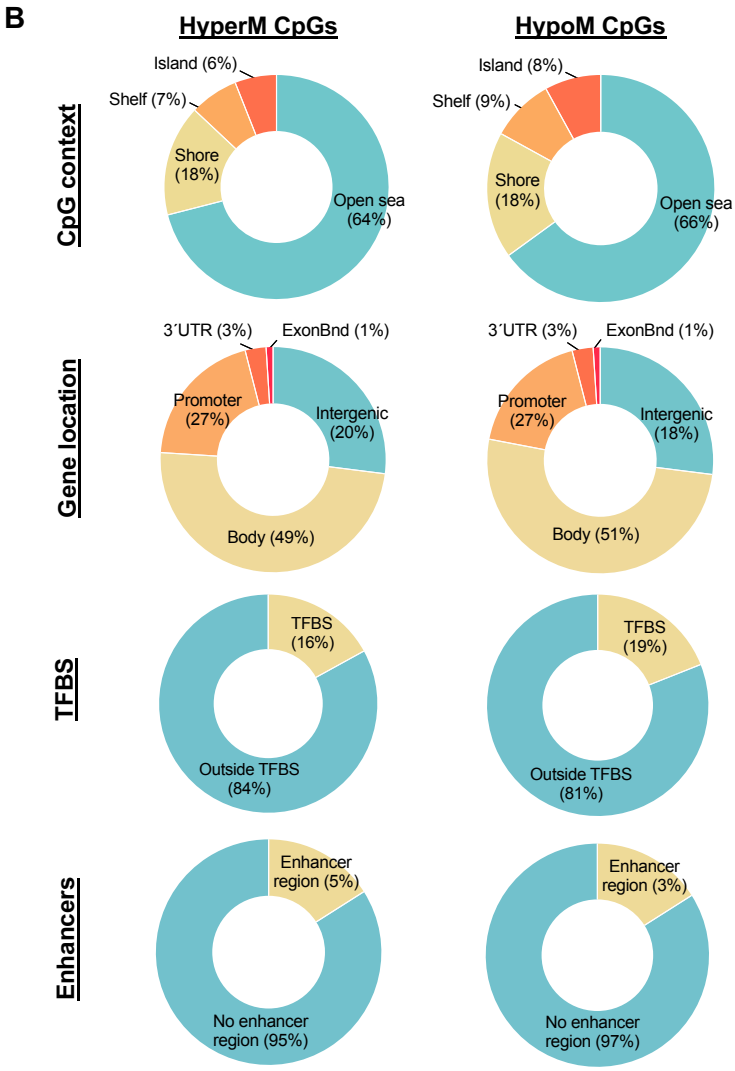
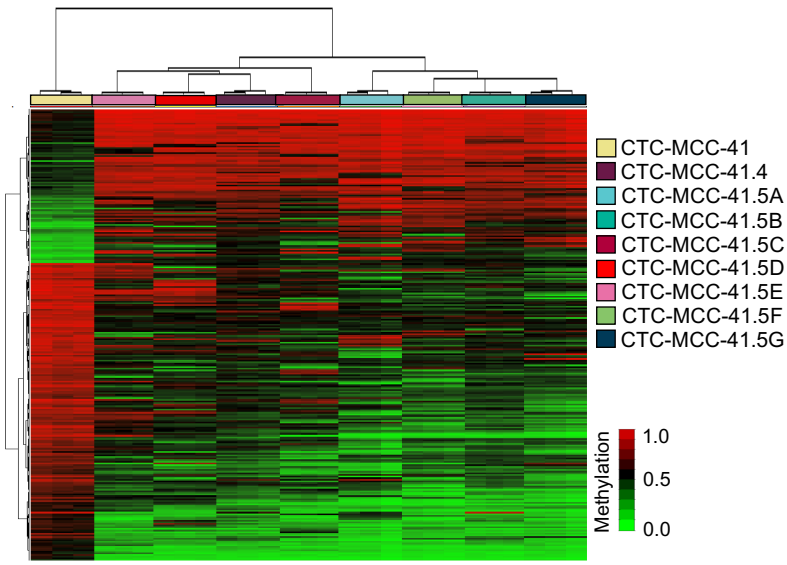


Figure 20. Description of the 34,085 DMCpGs found in the pre-treatment cell line CTC-MCC-41 respect to all post-treatment cell lines CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]. (A) chromosome location and methylation status, and **(B)** CpG context, gene location and transcription factor binding sites. FDR, false discovery rate; CpGI, CpG island, HypoM, hypomethylated; HyperM, hypermethylated; TFBS, transcription factor binding sites; ExonBnd, exon boundaries.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Noteworthy, a hierarchical clustering analysis of the 10,000 most DMCpGs previously found displayed a DNA methylation profile able to clearly differentiate the pre-treatment CTC-MCC-41 cells from all the post-treatment cancer cells (**Figure 21A**). Significantly, hierarchical clustering analysis was able to separate post-treatment cells in two branches: on one hand CTC-MCC-41.4 and CTC-MCC-41 5 [CDE] and on the other hand CTC-MCC-41 5 [ABFG]. **Supplementary Table 7** shows the 20 most DMCpGs found between CTC-MCC-41 and all post-treatment CTC cell lines. Taking into account the methylation profile obtained, we carried out a KEGG enrichment analysis and Panther pathway analysis, which revealed an enrichment of relevant pathways related to PI3K/AKT and MAPK signalling, metabolism, Wnt signalling, cadherin and integrin signalling, inflammation and apoptosis, among others (**Figure 21B**).

A



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

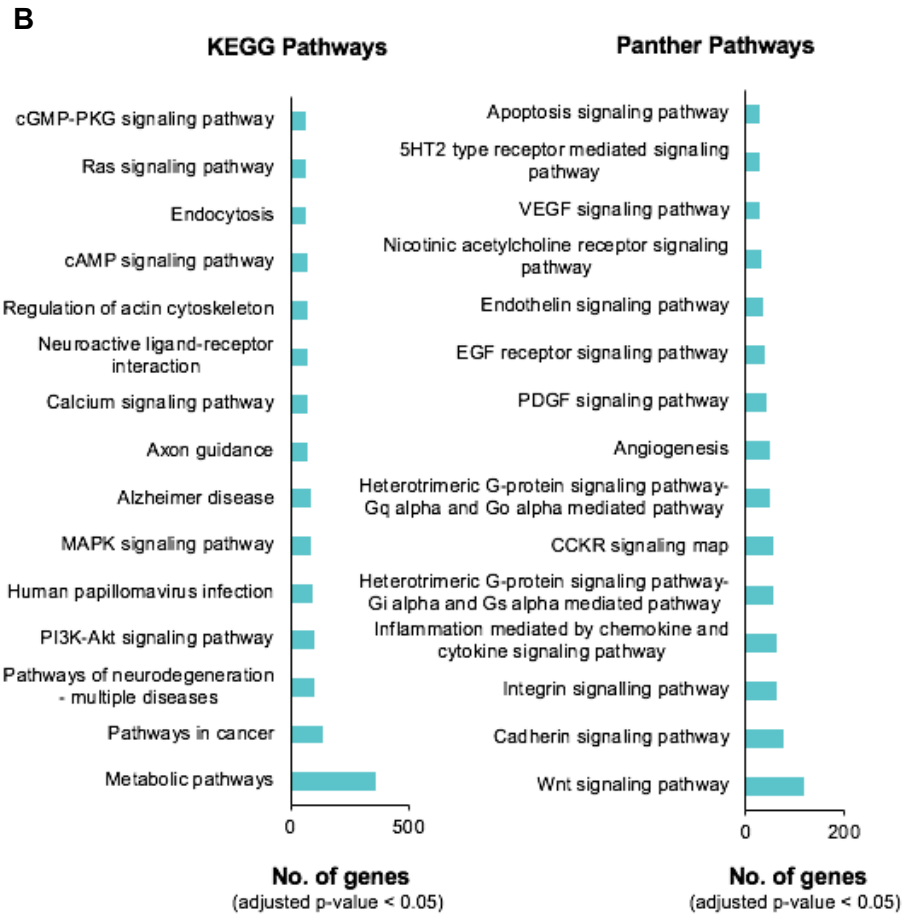


Figure 21. DNA methylation pattern of CTC-MCC-41 compared to post-treatment cell lines CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]. (A) Hierarchical clustering heatmap of the 10,000 most DMCpGs (FDR adjusted p-value <0.05) in CTC-MCC-41 respect to CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]. **(B)** GO analysis representing some of the most cancer relevant KEGG and Panther pathways based on the 10,000 most DMCpGs of the pre-treatment cell line compared to all post-treatment cell lines.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.2 Comparison of the DNA methylation profile among post-treatment CTC cell lines: CTC-MCC-41.4 respect CTC-MCC-41.5 [A-G] cells

The CTC cell line CTC-MCC-41.4 was stablished from CTCs obtained at the time of the second relapse at the end of second-line therapy, while the CTC-MCC-41.5 [A-G] cell lines were stablished from CTCs isolated after second-line treatment completion at the time of the third relapse and one week before patient´s death[139]. Therefore, we decided to compare the DNA methylation profile of CTC-MCC-41.4 with the CTC-MCC-41.5 [A-G] cells. This comparison only identified 4,475 significant DMCpGs (p-value <0.05; FDR adjusted p-value <0.05; Δβ-value >|0.20|) (**Figure 22A**). In line with this, the scatter plot analysis also revealed a similar methylation pattern between CTC-MCC-41.4 and CTC-MCC-41.5 [A-G] (**Figure 22B**).

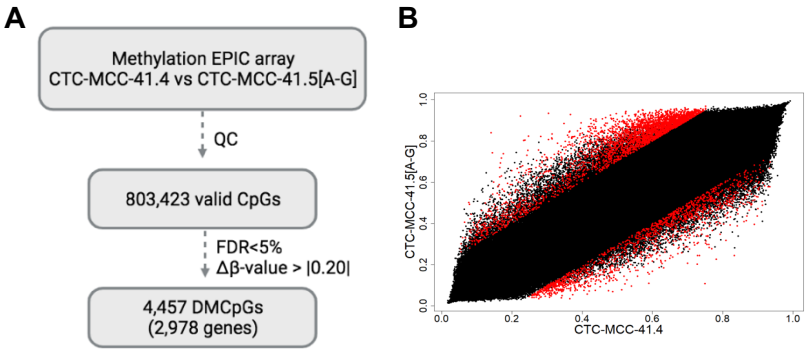
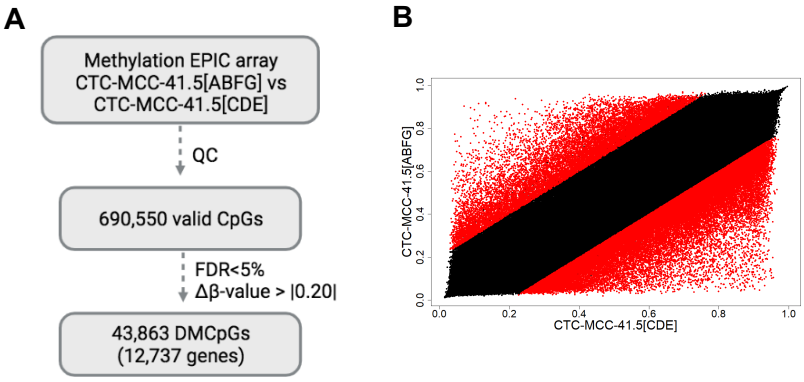


Figure 22. Genome-wide DNA methylation analysis of the post-treatment cell line CTC-MCC-41.4 respect to the post-treatment cell lines CTC-MCC-41.5 [A-G]. (A) Schematic flowchart used to identify significant DMCpGs in CTC-MCC-41.4 compared to CTC-MCC-41.5 [A-G]. **(B)** Scatter plot representing mean normalized levels of DNA methylation (β-values) in CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]. Dots in red show significant DMCpGs. QC, quality control; FDR, false discovery rate.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.3 Evaluation of the DNA methylation profile among post-treatment CTC lines: Comparison between CTC-MCC-41.5 [ABFG] and [CDE]

Recently, a previous work by Cayrefourq *et al.*[253] identified that the CTC-MCC-41.5A, CTC-MCC-41.5B, CTC-MCC-41.5F and CTC-MCC-41.5G (CTC-MCC-41.5 [ABFG]) cell lines have a different transcriptional program than CTC-MCC-41.5C, CTC-MCC-41.5D and CTC-MCC-41.5E (CTC-MCC-41.5 [CDE]). Based on this, to see if these two groups of CTC cell lines could also have a different methylation pattern, the DNA methylation profile of both groups of cell lines was evaluated. Thus, we identied that CTC-MCC-41.5 [ABFG] cells present 43,863 significant DMCpGs (p-value <0.05; FDR adjusted p-value <0.05; Δβ-value >|0.20|) in comparison with CTC-MCC-41.5 [CDE] cells (**Figure 23A**). The scatter plot also revealed a different methylation pattern between both groups of post-treatment cells (**Figure 23B**). Significantly, a hierarchical clustering analysis of the 10,000 most DMCpGs revealed a methylation profile that clearly differentiates CTC-MCC-41.5 [ABFG] from CTC-MCC-41.5 [CDE] cell lines (**Figure 23C**).



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

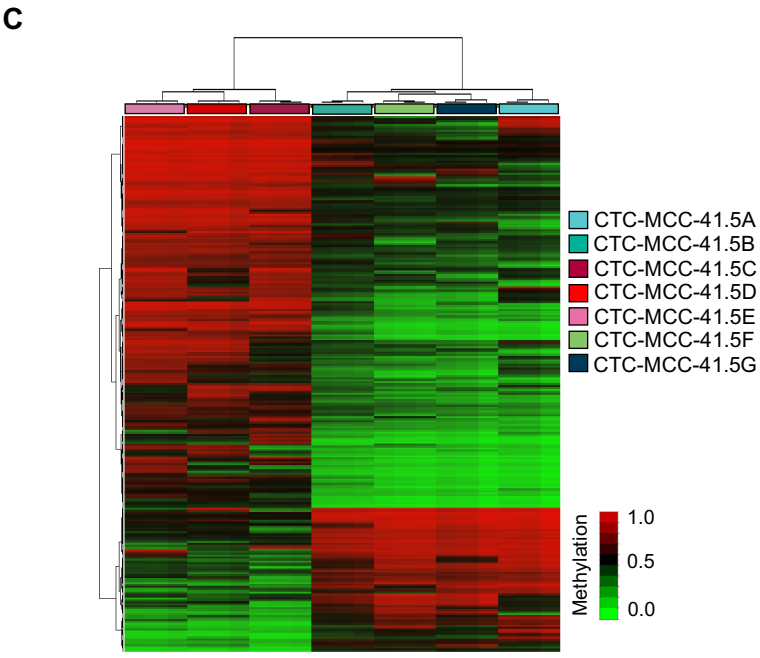


Figure 23. Genome-wide DNA methylation analysis of the post-treatment cell lines CTC-MCC-41.5[ABFG] respect to CTC-MCC-41.5[CDE]. (A) Schematic flowchart used to identify significant DMCpGs in CTC-MCC-41.5[ABFG] compared to CTC-MCC-41.5[CDE]. **(B)** Scatter plot representing mean normalized levels of DNA methylation (β-values) in CTC-MCC-41.5[ABFG] and CTC-MCC-41.5[CDE]. Dots in red show significant DMCpGs. **(C)** Hierarchical clustering heatmap of the 10,000 most DMCpGs (FDR adjusted p-value <0.05) in CTC-MCC-41[ABFG] respect to CTC-MCC-41[CDE]. QC, quality control; FDR, false discovery rate.

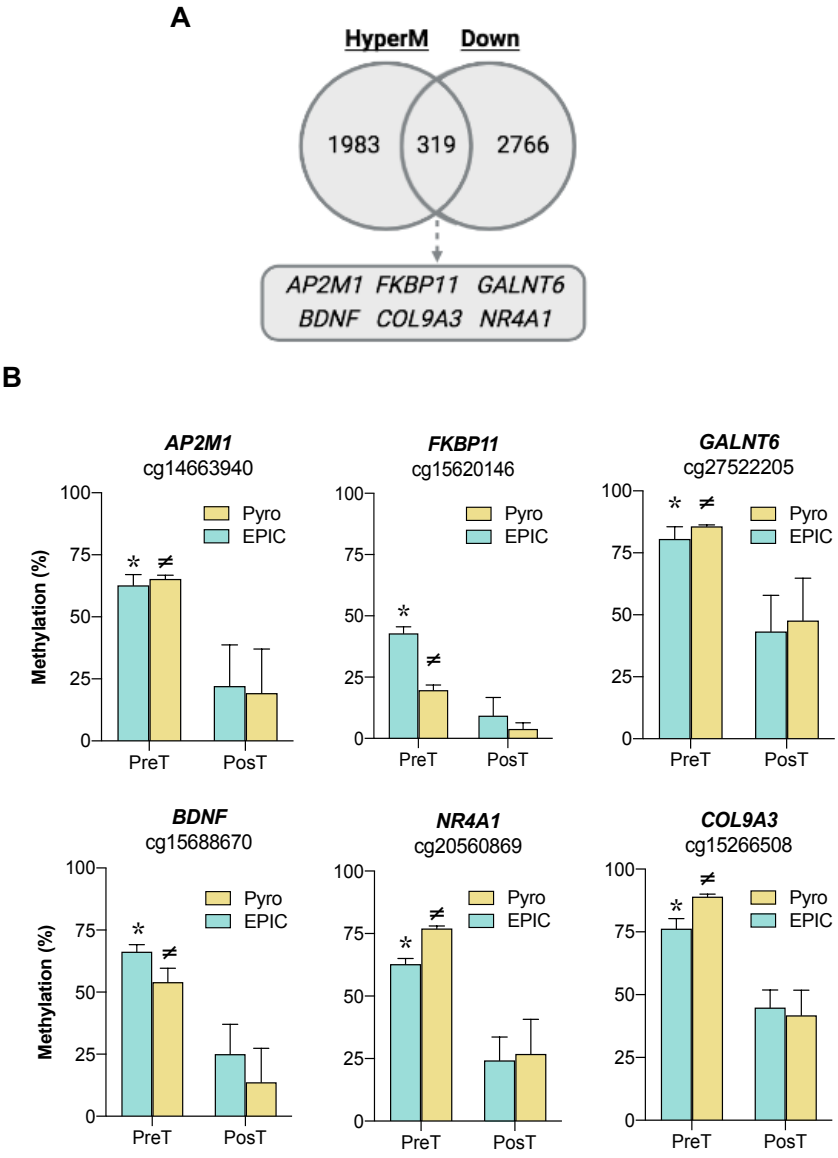
USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.4 Association between the DNA methylation profile and the transcriptional program of the pre-treatment and post-treatment CTC cell lines

The transcriptional profile of the pre-treatment CTC-MCC-41 and all the post-treatment cell lines (CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]) was analyzed in a previous work by Cayrefourq *et al.*[253] revealing differentially expressed genes (up- and down regulated) between these two groups of CTC cell lines. These data were used to test whether DNA methylation changes at CpGIs and shore regions of gene promoters were associated with gene expression of these colorectal CTC cell lines. In particular, since we have previously observed that CTC-MCC-41 is more hypermethylated than all the post-treatment cell lines, we focused our study in evaluating the association between the hypermethylated and downregulated genes found in CTC-MCC-41 (**Figure 24A**). Thus, Venn diagram analysis enabled the identification of 319 genes whose hypermethylation in the pre-treatment cell line CTC-MCC-41 was associated with their transcriptional silencing. Out of these genes, we selected six candidate genes (*AP2M1*, *FKBP11*, *GALNT6*, *BDNF*, *COL9A3* and *NR4A1*) based on the following criteria: i) High methylation (Δβ-value ≥ 0.30) and high expression difference (fold change ≥ 2.0) between CTC-MCC-41 and all the post-treatment cell lines; ii) Implication in relevant pathways (PI3K/AKT, MAPK, metabolism) previously identified in our GO analysis comparing CTC-MCC-41 and all the post-treatment cell lines.

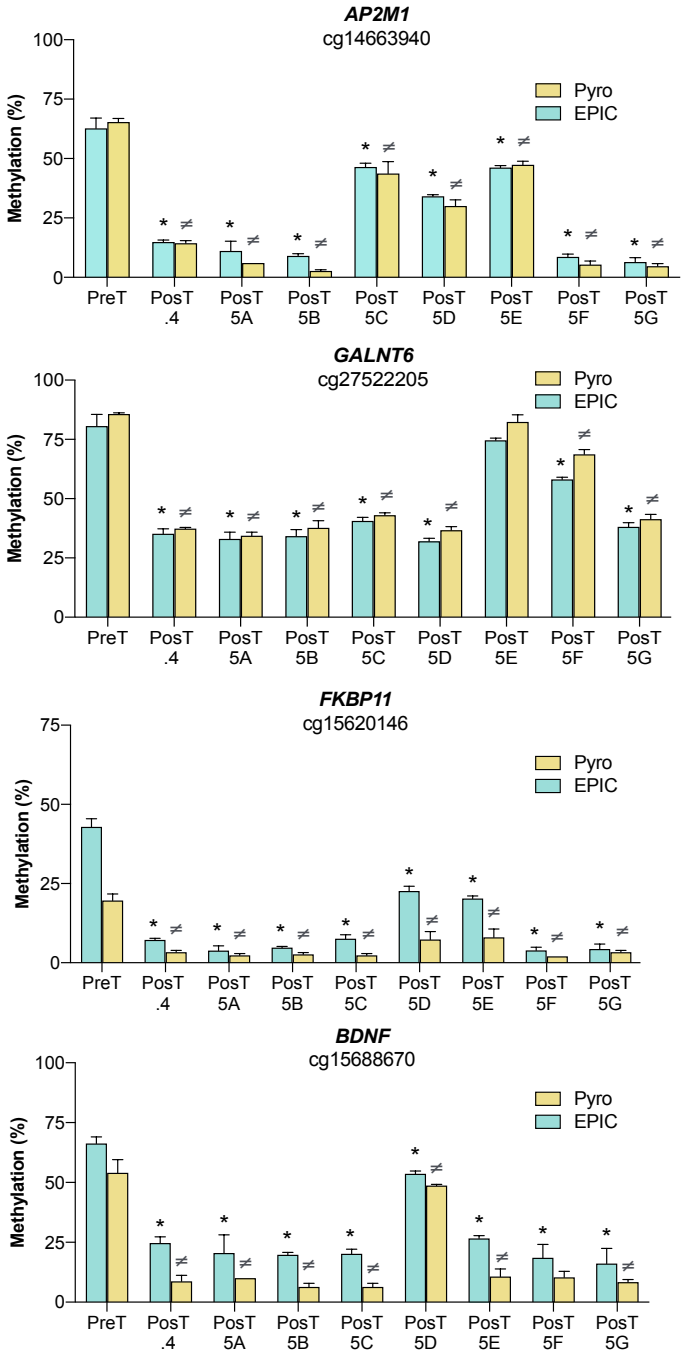
The DNA methylation levels of the six selected genes analyzed by EPIC array (**Supplementary Table 8**) were validated using bisulfite pyrosequencing. As expected, for all the CpGs and genes analyzed, pyrosequencing analysis showed higher methylation levels in the pre-treatment CTC-MCC-41 cell line than in the post-treatment cell lines (**Figure 24B**). Besides, in most cases, similar results were obtained when the DNA methylation of each CTC line was analyzed individually (**Figure 24C**).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

C



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

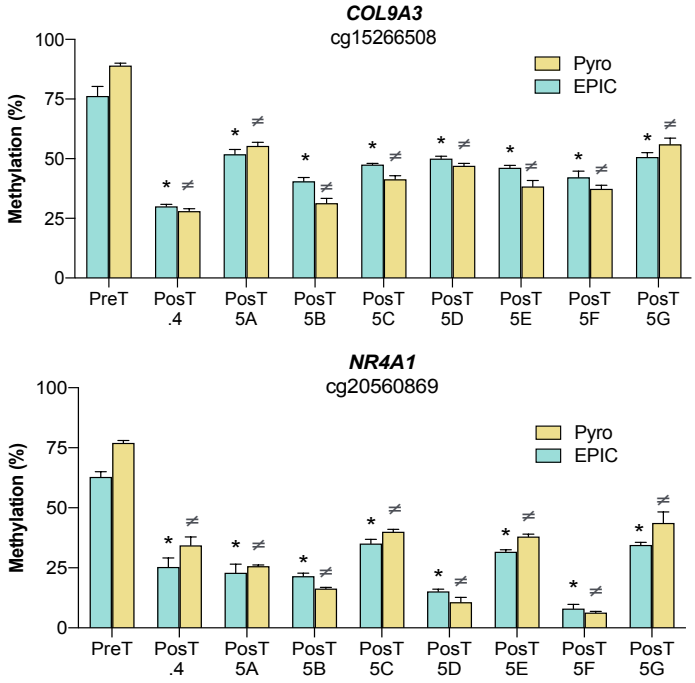
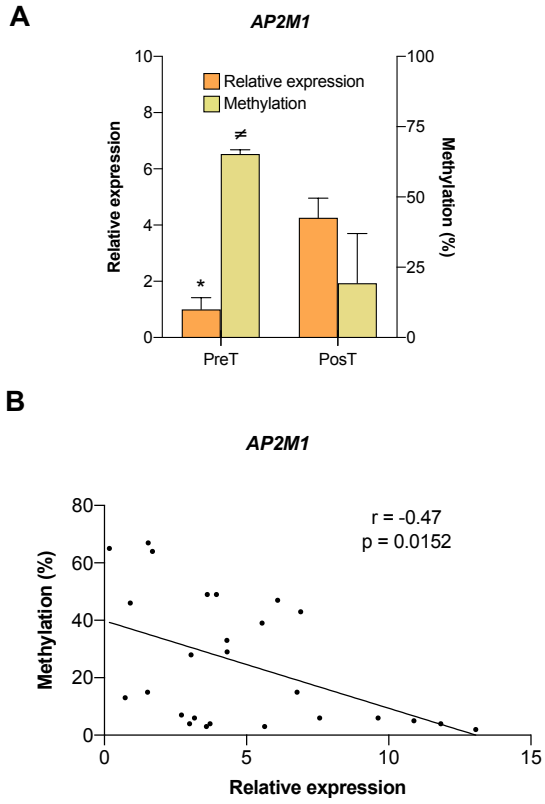


Figure 24. Impact of DNA methylation into the transcriptional program of the pre-treatment CTC-MCC-41 cell and all post-treatment cell lines. **(A)** Association between the hypermethylated-downregulated genes found in the pre-treatment cell line CTC-MCC-41 respect to all post-treatment cell lines CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]. HyperM, hypermethylated; Down, downregulated. **(B-C)** Validation of the methylation status of representative promoter CpGs from the selected genes **(B)** in the pre-treatment cell line CTC-MCC-41 compared to all post-treatment lines globally and **(C)** in the pretreatment cell line CTC-MCC-41 compared to each post-treatment line individually. EPIC array data are expressed as percentage (%) of methylation. *, p-value <0.05 between the pre-treatment cell line CTC-MCC-41 and post-treatment cell lines after analysis by EPIC array; ≠, p-value <0.05 between the pre-treatment cell line CTC-MCC-41 and post-treatment cell lines after analysis by pyrosequencing. Pyro, pyrosequencing; Post.4, Post-treatment cell line CTC-MCC-41.4; Post 5 [A-G], Post-treatment cell lines CTC-MCC-41.5 [A-G].

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Among the six selected genes, *AP2M1* was the gene that showed higher methylation differences between the pre- and post-treatment cells in the EPIC array analysis (**Supplementary Table 8**). Therefore, we decided to explore the association between the promoter DNA methylation levels of this gene and its transcriptional levels analyzed by pyrosequencing and RT-qPCR, respectively, in the pre- and post-treatment CTC lines. This analysis revealed a negative association between the methylation levels of *AP2M1* promoter and its transcriptional levels (**Figure 25A**). Of note, a negative association (r = -0.47; p = 0.0152) was also observed when we consider the paired promoter methylation and levels of each pre- and post-treatment cell line individually (**Figure 25B-C**).



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

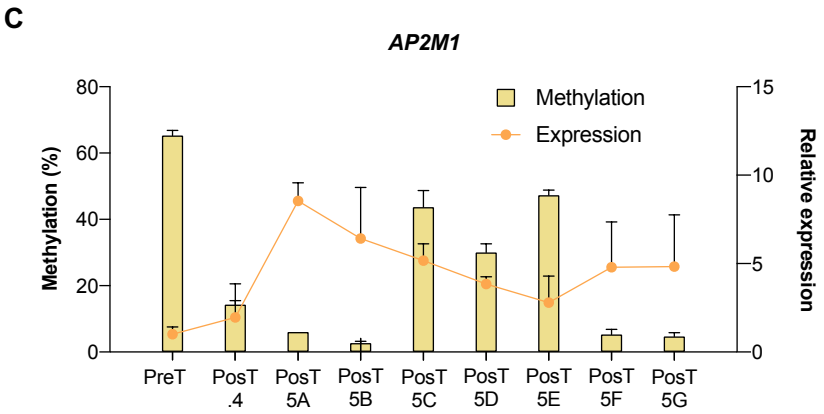


Figure 25. Expression and promoter methylation status of the *AP2M1* gene in the pre-treatment CTC-MCC-41 cells, and all post-treatment cell lines. (A) Levels of methylation and expression in the pre-treatment cell line CTC-MCC-41 compared to all post-treatment cell lines CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]. PreT, pre-treatment cell line; PosT, post-treatment cell lines. *, p-value <0.05 between the pre-treatment cell line CTC-MCC-41 and all post-treatment cell lines for relative expression; ≠, p-value <0.05 between the pre-treatment cell line CTC-MCC-41 and all post-treatment cell lines for methylation **(B)** Correlation between the methylation and expression levels of *AP2M1* considering the pre-treatment cell line CTC-MCC-41 and all the post-treatment cell lines. **(C)** Levels of methylation and expression in the pre-treatment cell line CTC-MCC-41 compared to all post-treatment lines individually. PreT, pre-treatment cell line; PosT.4, Post-treatment cell line CTC-MCC-41.4; PosT5 [A-G], post-treatment cell lines CTC-MCC-41.5 [A-G]. Methylation and expression values were determined from triplicates and are expressed as mean ± SD.

5. Discussion

In CRC treatment, great progress has been made in the last decades in the development and use of chemotherapeutic agents, molecularly targeted therapy and immunotherapy[256]. However, prognosis for advanced CRC is still poor, with a 5-year survival rate of 14–15% in patients with metastatic

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

disease[257], being drug resistance a major handicap for CRC treatment, contributing to the poor prognosis[242].

Although mechanisms that drive resistance are complex and not well established[258] it is known that drug resistance development implicates general gene expression reprogramming. In this sense, epigenetic regulation is one of the fundamental processes determining both primary and acquired drug resistance in CRC[256]. It has been reported, in fact, that deregulation of DNA methylation confers resistance to fluoropyrimidine-based treatments in CRC[259]. Understanding the mechanisms of treatment resistance and identifying new biomarkers associated with therapy resistance is necessary to improve the outcome of CRC patients[256]. In this sense, molecular characterization of CTCs and CTC cell lines has proven to be a relevant approach[253,260].

In this chapter, we compared the DNA methylome of the pre-treatment metastasis-competent cell line CTC-MCC-41 with eight resistant CTC lines (CTC-MCC-41.4 and CTC-MCC-41.5 [A-G]) stablished from the same patient after 5-fluorouracile-based treatment and progression of the disease. We show that the deregulation of DNA methylation in CTCs is a relevant mechanism associated with therapy resistance in CRC.

Results in this work clearly revealed that resistant CTCs have a different global DNA methylation profile than pre-treatment CTCs. Interestingly, we observed a generalized loss of methylation in resistant CTCs, suggesting that selective treatment pressure is able to reprogram the DNA methylome of CTCs leading to a more hypomethylated profile. These hypomethylated CpGs were mainly distributed in CpG-poor and body regions, which suggests that the epigenetic deregulation of this part of the genome may play an important role in the development of therapy resistance. In addition, the hypomethylated CpGs of resistant CTCs were also located in

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

important regulatory regions of the genome, including CpGIs and shore regions, promoters, TFBS and enhancer regions. In line with this, hypomethylation of gene promoters has been previously associated with acquiring drug resistance in several types of cancers due to their implications in multidrug resistance (MDR), proliferation, apoptotic signaling, mitochondrial function and DNA repair, among other mechanisms[261]. In particular, the promoter hypomethylation of *NRF2* and *NME2* has been linked to 5-FU chemoresistance in CRC by inducing cellular protection mechanisms and reducing apoptosis, respectively[259,262]. Moreover, it has been also described that hypomethylation of TFBS can activate transcription factors or prepare them for later activation during processes such as cancer progression[119], and that oncogenes and driver genes are characterized by the hypomethylation of their enhancer regions[263].

In addition, we identified a distinctive DNA methylation signature that clearly differentiated the resistant CTC lines from the pre-treatment CTC-MCC-41 cells. Of note, the genes involved in this signature were associated with relevant cancer pathways including PI3K/AKT, MAPK and metabolism; suggesting that these pathways could be relevant for drug resistance and proliferation of CTCs. This observation is in agreement with a recent work that identified the transcriptional deregulation of these pathways in these CTC lines[264]. The alteration of PI3K/AKT pathway has been reported in various cancer types, like CRC, playing a crucial role in cancer development including metastasis proliferation, survival and angiogenesis[265] and it has become a popular target for new drugs in cancer[264]. Similarly, MAPK signaling pathway promotes tumour growth, progression and drug resistance in CRC[265,266]. Besides, alterations of the metabolism have shown relevant implications in therapy resistance in CRC[267].

On the other hand, when we compared CTC-MCC-41.4 with CTC-MCC-41.5 [A-G] cells, we did not find many methylation differences. These results

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

are in line with the expression data already published for these resistant CTC cell lines, in which not many expression differences were found between both groups, suggesting that the clonal evolution of all these resistant CTC clones was already stablished during treatment pressure[253]. On the other hand, we did observe many methylation differences between the group of clones CTC-MCC-41.5 [ABFG] and CTC-MCC-41.5 [CDE], obtaining a methylation signature able to clearly differentiate both groups. In this sense, previous expression data of these cells showed a similar segregation of these groups based on their expression profile, which allowed the authors to suggest that the CTC-MCC-41.5 [ABFG] cells might be derived from CTCs released by liver metastases[253].

In this work, we have also observed that the reprogramming of the DNA methylation of CTC lines under selective therapy pressure is related to changes in the expression profile of these cells. This is the case of the genes *AP2M1*, *FKBP11*, *GALNT6*, *BDNF*, *COL9A3* and *NR4A1*, which showed a negative association between their DNA methylation and their transcriptional levels. All these genes showed hypomethylation and overexpression in all the resistant CTC lines (CTC-MCC-41.4 and CTC-MCC-41 [A-G]), suggesting that the selective pressure of treatment in CTCs induces the methylation loss and an enhanced expression of these genes, restoring their activity. In line with this, an increased activity of these genes has been previously associated with the development of mechanisms of resistance to treatment. In the case of *BDNF,* its overexpression has already been linked in CRC with enhanced dissemination, anoikis resistance and chemoresistance[268]. In neuroblastoma, expression of this gene protects cells from chemotherapy-induced apoptosis via PI3K and also increased metastasis via PI3K/AKT/mTOR and MAPK pathways[269,270]. Regarding *FKBP11*, the expression of this gene showed strong correlation with advanced disease and poor prognosis in renal cancer[271]. In oral squamous cell carcinoma, it has been described to promote

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

cell proliferation and tumourigenesis via p53-related pathways. Moreover, this gene could promote cell proliferation by regulating G2/M phase and apoptosis[272]. The gene *COL9A3* promotes cell EMT, invasion, and migration in gastric cancer[273]. Moreover, overexpression of *GALNT6* has already been associated in CRC with resistance to 5-FU and to the proliferation and migration of colon cancer cells[274]. This gene also participates, in breast cancer, in the activation of PI3K/AKT signaling pathway, involved in migration and invasion cancer cells[275]. Concerning *NR41A*, it is involved in cancer progression by promoting metastasis, inflammation, inhibition of apoptosis, proliferation, stemness and EMT. This gene inhibits the expression of Dicer to activate downstream Akt/mTORC1 signaling, inducing colon cancer EMT[276].

Besides, of especial interest is the gene *AP2M1,* which showed a very clear inverse relation between methylation and expression in the pre-treatment CTC-MCC-41 cell line and in each of the resistant cell lines. This gene encodes a subunit of the heterotetrameric coat assembly protein complex 2 (AP2) and is required for the activity of a vacuolar ATPase and may also play an important role in regulating the intracellular trafficking[277]. This *AP2M1* gene has been recently associated with chemotherapy resistance in several tumour types, including lymphoma[277], breast and colon cancer[278]. Chemotherapy is able to induce senescence producing growth arrest and avoiding tumour cell proliferation. However, recent studies have shown that this senescence is not always definitive and that persistent populations of tumour cells can use senescence as an adaptive pathway to restart proliferation and become more aggressive[279]. Thus, some tumour cells may escape from chemotherapy-induced senescence and reproliferate[278]. A better understanding of the cellular pathways underlying this escape from cell death is critical[280] to improve the outcome of mCRC patients. Importantly, it has been described that *AP2M1* participates in the transmission of soluble signals from

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

colorectal senescent cells promoting escape from senescence and inducing chemotherapy resistance[278], suggesting that the hypomethylation and overexpression of *AP2M1* observed in the resistant CTCs studied in our work might be a novel mechanism of resistance by which CTCs escape from chemotherapy-induced senescence.

All in all, in this chapter, we compared the DNA methylome of the pre-treatment cell line CTC-MCC-41 with eight resistant CTC cell lines obtained after 5-FU-based treatment, showing that selective pressure of treatment induces global changes in the DNA methylation profile of CTCs. In particular, resistant CTCs were characterized by a generalized loss of methylation leading to a more hypomethylated profile. Moreover, resistant CTCs showed a distinctive DNA methylation signature related to relevant cancer pathways, such as PI3K/AKT, MAPK and metabolic pathways. Several epigenetically deregulated genes associated with the resistance of CTCs were identified, mainly *AP2M1*, representing new potential biomarkers of resistance and therapeutic targets that could contribute to the development of new clinical tools and therapies for CRC

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

CHAPTER III

Epigenomic analysis of CTCs from mCRC patients in relation to therapy response and the progression of the disease

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Epigenomic analysis of CTCs from mCRC patients in relation to therapy response and the progression of the disease

1. Abstract

Molecular characterization of CTCs could contribute to a better understanding of the metastatic process and could have translation into the clinic as a valuable tool. In this chapter, the aim was to isolate CTCs from peripheral blood of mCRC patients and study their DNA methylation profile in relation to therapy response and progression of the disease. For this end, we first compared the CellSearch and the Parsortix systems for CTC enrichment. Then, we isolated pure CTCs from the enriched samples using the DEPArray, and WGBS was performed to profile their DNA methylation. Thus, we observed that the CellSearch showed higher CTC recovery rates, and the presence of ≥3 CTCs/7.5 mL of peripheral blood analyzed with CellSearch was associated with different clinicopathological characteristics and was also a prognostic independent factor associated with unfavourable PFS and OS in mCRC patients. Besides, characterization of genomic DNA revealed different methylation profiles of baseline CTCs from mCRC patients in association with response to FOLFOX in monotherapy or in combination with bevacizumab therapy. We also identified a methylation profile in the genomic DNA of CTCs associated with the development of resistance to FOLFOX and bevacizumab. In addition, mitochondrial DNA of CTCs also showed changes in the methylation profile in relation to therapy response and the development of resistances. These results reveal new potential therapeutic targets and epigenetic biomarkers able to predict therapy response and the development of resistance, which could assist in the clinical practice to improve the outcome of mCRC patients.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

2. Introduction

Standard treatment of mCRC is usually based on the use of fluoropyrimidines, either in monotherapy or combined with other agents to improve response rates. In this sense, the use of FOLFOX or FOLFOX with the anti-VEGF bevacizumab, are commonly used as a first-line therapy for mCRC patients[281]. However, in spite of treatment improvements, response rates are still low and a high number of patients do not benefit from therapy[282]. One of the great challenges continues to be patient selection before treatment initiation to obtain the greatest benefit from therapy. Besides, identifying biomarkers and mechanisms of resistance able to monitor therapy effectiveness during the patients´ follow-up is also challenging[283].

Treatment effectiveness in the clinical routine is generally assessed using imaging methods that cannot be repeatedly performed due to radiation overexposure to the patient[284]. In this sense, liquid biopsy is a minimally invasive technique that provides information on the cancer burden in real time and reveals the evolution and heterogeneity of the disease over time[285]. Moreover, CTC quantification is used as prognostic marker for disease outcome, being the CellSearch system the only clinically validated, FDA-cleared test for enumerating CTCs in peripheral blood of CRC patients[284]. This system is epitope-dependent and isolates CTCs based on the expression of the epithelial marker EpCAM. However, CTCs are very heterogeneous and not always express epithelial markers, making their detection and isolation extremely difficult and challenging[285,286].

In line with this, new approaches for CTC isolation and detection have been recently developed, such as the Parsortix system, that isolates these cells based on their size and deformability or the DEPArray, that isolates pure CTCs based on their dielectrophorelectric properties respect to blood cells[284,286]. Further beyond detection and enumeration, molecular

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

characterization of CTCs may provide useful information of the biology of the metastasis, but also could have translation into the clinic as a tool for real time monitoring of therapies or the identification of biomarkers, therapeutic targets and resistance mechanisms[284,285]. In this sense, the DNA methylation profiling of CTCs in patients with mCRC in relation to therapy response and progression is completely unexplored.

In this chapter we aim to identify aberrantly methylated CpGs and new potential biomarkers in CTCs of mCRC patients associated with the response and resistance to therapy with FOLFOX in monotherapy or combined with the anti-VEGF bevacizumab. For this, we first compared the performance of the CellSearch and Parsortix systems, to select the most appropriate for CTC detection and enrichment in our mCRC patient cohort. Then, after evaluating the clinical impact of the enumeration of CTCs in the mCRC patients analyzed, we isolated pure CTCs using the DEPArray system to characterize, using WGBS, their DNA methylome in three clinical situations: i) CTCs obtained before starting first-line therapy with FOLFOX in a responder and a non-responder patient to treatment; ii) CTCs obtained before starting first-line therapy with FOLFOX and bevacizumab in a responder and non-responder patient to treatment; and iii) CTCs obtained from the same patient before starting first-line therapy with FOLFOX and bevacizumab and after treatment initiation at the moment of progression.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3. Materials and methods

3.1 Patient cohort

A total of 73 mCRC patients were recruited for this study, between April 2019 and October 2021, in the Medical Oncology Departments at the University Clinical Hospital of Santiago de Compostela and the Ramon y Cajal Hospital of Madrid. This study was approved by the Ethics Committee of Galicia (reference number 2019/017) and conducted in accordance with the Declaration of Helsinki. Patient samples were obtained after signing informed consent and samples were encoded and anonymized after withdrawal. Inclusion and exclusion criteria are collected in **Supplementary Table 9**. All patients were followed-up, and their clinical information was collected, until exitus or the end of the study. Clinicopathological characteristics of patients are described in **Table 7**. Response to treatment was evaluated with computed tomography scan during follow-up and patients were classified, according to the results of the first computed tomography after starting therapy (~3 months) and based on RECIST 1.1 criteria, as responders (stable disease, partial response or complete response) or non-responders (progressive disease) to treatment. PFS was defined as the time from the date of initial treatment until the date of disease progression, death or the last follow-up if progression or death had not occurred. OS was defined as the time from the date of initial treatments until death or the last follow-up.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Table 7. Clinical characteristics of mCRC patients of the study.

Characteristics	Patients *(N = 73)*	
	No.	%
Age (38-85 yrs)		
<65	34	47
≥65	39	53
Gender		
Female	24	33
Male	49	67
Histology		
Adenocarcinoma	67	92
Mucinous carcinoma	6	8
Stage		
IV	73	100
Metastatic sites		
≤1	37	51
>1	36	49
First-line of therapy		
FOLFOX	31	43
FOLFOX + bevacizumab	22	30
Others	20	27
CEA at diagnosis		
≤5 ng/mL	13	18
>5 ng/mL	59	82
***KRAS* tissue**		
KRAS wild-type	22	30
KRAS mutated	50	68
Unknown	1	1

Yrs, years.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.1.1 Blood sample collection

For each patient, two 7.5 mL CellSave preservative blood tubes (Menarini, Silicon Biosystems) were collected by phlebotomy at the moment of the mCRC diagnosis and before the initiation of treatment and, if taken case, at the moment of progression after initiating first line therapy.

3.2 Enumeration and enrichment of CTCs

3.2.1 The Parsortix Cell Separation System

One CellSave preservative tube (Menarini, Silicon Biosystems, Italy) containing 7.5 mL of blood was processed for each patient, within the next three days after collection, using the Parsortix Cell Separation System (Angle Inc, UK). Blood samples were processed through this microfluidic device using 99 mbar pressure and a single-use Parsortix GEN3D6.5 Cell Separation Cassette (Angle Inc, UK) with a size of 6.5 µm, following manufacturer´s guidelines. Blood cells flow through the cassette and CTCs remain trapped inside it due to their larger size and deformability. After CTCs separation, samples were fixed inside the cassette using 4% paraformaldehyde and immunofluorescence staining was performed with the selected antibodies (**Table 8**) and the nuclear dye Hoechst 33342 with NucBlue Live ReadyProbes (Invitrogen, USA), according to manufacturer´s guidelines. Stained cassettes were analyzed for fluorescence using the microscope Leica DMi8 (Leica Microsystems, Germany) and the Leica Application Suite X (LAS X) software (Leica Microsystems, Germany), identifying CTCs as $CKs^{+}/DAPI^{+}/CD45^{-}$ for cells expressing the epithelial marker cytokeratin, or $VIM^{+}/DAPI^{+}/CD45^{-}$ cells for cells expressing the mesenchimal marker vimentin.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Table 8. Antibodies used for immunofluorescence staining in the Parsortix Cell Separation System.

Target	Antibody	Dilution
CK	Anti-Cytokeratins Alexa Fluor 488 (ExBio)	1:50
VIM	Vimentin (D21H3) XP® Rabbit mAb (Alexa Fluor® 647 Conjugate) (Cell Signaling)	1:100
CD45	Anti-Hu CD45 PE (ExBio)	1:50

CK, citokeratin; VIM, vimentin.

3.2.2 The CellSearch system

One or two CellSave preservative tubes (Menarini, Silicon Biosystems, Italy) containing 7.5 mL of blood were processed for each patient, within the next three days after collection, using the CellSearch System (Menarini, Silicon Biosystems, Italy). Blood samples were processed firstly with the CellSearch CellTracks AutoPrep System (Menarini, Silicon Biosystems, Italy), where the solid blood components from plasma are incubated using ferrofluid nanoparticles coated with antibodies with different targets according to the kit. The kit used for all samples was the CellSearch Circulating Tumour Cell Kit (Menarini, Silicon Biosystems, Italy), this kit includes ferrofluid nanoparticles coated with antibodies that target the epithelial cell adhesion molecule EpCAM to enrich epithelial cells and also a mixture of antibodies that target CKs 8/18/19 conjugated to phycoerythrin (PE), CD45 conjugated to allophycocyanin (APC) and the nuclear dye 4′,6-diamidino-2-phenylindole (DAPI) to fluorescently label the cells. Once the $EpCAM^+$ cells were stained and inside the cassette, samples were processed using the CellSearch CellTracks Analyzer II System (Menarini, Silicon Biosystems, Italy). This is an automated system that scans the

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

sample cartridge and displays the candidate fluorescent images for an operator´s review, identifying CTCs as $CD45^{-}/CK^{+}/DAPI^{+}$ cells.

3.3 Isolation of CTCs

3.3.1 The DEPArray system

Isolation of pure CTCs was performed in collaboration with the Laboratory of Rare Circulating Human Cells of the University Institute of Clinical Research in the University Hospital of Montpellier, led by Dr. Catherine Alix-Panabières, using the DEPArray System (Menarini, Silicon Biosystems, Italy). Samples with a CTC count of ≥10 CTCs after enrichment with the CellSearch system (Menarini, Silicon Biosystems, Italy) were selected as candidates for CTC isolation with the DEPArray system.

The stained cell suspension recovered from the CellSearch cartridge (Menarini, Silicon Biosystems, Italy) was prepared following the manufacturer´s instructions to be loaded into the DEParray cartridge (Menarini, Silicon Biosystems, Italy) and processed with the DEPArray. This cartridge combines microfluidics with microelectronics to manipulate single CTCs.

The main chamber of the cartridge, which is covered by microelectrodes that allow the generation of dielectrophoretic cages able to capture a cell due to the dielectrophoretic force, is scanned and generates images of the caged cells, allowing to identifying CTCs as CK^{+} (PE)/ $CD45^{-}$ (APC)/ $DAPI^{+}$ cells. After CTCs selection, the strategy of recovery was designed, and cells moved to the collection tube by motion of the positive dielectrophoretic cages. Individual CTCs or clustered CTCs were recovered individually or as pools in several tubes for the downstream analysis.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.4 DNA methylation analysis of CTCs

3.4.1 DNA isolation and bisulfite treatment

Single or pooled pure CTCs recovered from the DEPArray system, were treated with the EZ DNA Methylation-Direct kit (Zymo Research, USA) following manufacturer's instructions. This kit can accommodate cells as direct input and combines DNA isolation and bisulfite conversion, reducing the chance of material loss.

The isolated DNA of CTC samples were modified using a ProFlex PCR System (Thermo Fisher Scientific, USA) under the incubation conditions of 98°C for 8 min. and 64°C for 3.5 h and eluted in a volume of 8 µL.

3.4.2 Whole-genome bisulfite sequencing analysis (WGBS)

3.4.2.1 Library preparation and sequencing

All the elution volume of the bisulfite-modified DNA was used as input to perform the WGBS libraries using the Pico Methyl-Seq Library Prep Kit (Zymo Research, USA) following manufacturers´s recommendations. This kit allows the performance of libraries with very low quantities as input, and briefly, samples undergo two amplification rounds and a third amplification round with an index primer. Between steps, DNA is purified and concentrated using the DNA Clean & Concentrator-5 kit (Zymo Research, USA). All the thermocycling steps were performed using a ProFlex PCR System (Thermo Fisher Scientific, USA) and following the guideline instructions.

Finally, libraries were purified by the AMPure XP system (Beckman Coulter, USA) to eliminate primer dimers and they were analyzed as quality control for size distribution using the Agilent 4200 TapeStation system (Agilent Technologies, USA) with the D100 ScreenTape and Reagents (Agilent Technologies, USA). Libraries were quantified using the NEBNext Library Quant Kit for Illumina (New England Biolabs, USA) following manufacturer´s

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

instructions, and sequenced, with 151 bp paired-end reads, in a NextSeq500 or HiSeqXTen system (Illumina, USA).

3.4.2.2 Bioinformatic analysis

After WGBS, the resulting raw sequence data (FASTQ files) were evaluated using FastQC version 0.11.9. The FASTQ sequencing files were trimmed by removing adapter sequences, poor quality reads and bases until 50 bp length with Trim Galore! version 0.6.5[287]. Trimmed sequences were mapped to the human reference genome (hg38) in single-end mode using Bismark version 0.22.3[288] with the alignment tool Bowtie2 version 2.4.5[289], resulting in an average of 32% mapping efficiency. SAM files from the previous two rounds of alignment in single-end mode were merged using samtools version 1.16.1[290]. Duplicated reads were removed from each sample using deduplicate_bismark function and coverage and single CpG methylation status of all cytosines was extracted using bismark_methylation_extractor function from Bismark[288]. The resulting coverage files were processed in R environment version 4.1.1[291] to obtain differences in methylation levels between samples. DNA methylation levels for each CpG site was determined by counting the methylated and unmethylated reads. The sequencing depth threshold used to call the DNA methylation level for any cytosine was 1x for single CTCs and 3-5x for pool CTCs. Unsupervised hierarchical clustering heatmaps of methylation values and circos plot were generated using the ComplexHeatmap package[292] and the circlize package[293], respectively.

3.5 Statistical analysis

The Spearman correlation coefficients were calculated to assess the correlation between the CellSearch and Parsortix systems, and between the number of CTCs with CEA levels and cfDNA. The association of ≥3

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

CTCs/7.5mL with clinicopathological parameters was evaluated with a chi square test. The relationship between ≥3 CTCs/7.5mL and survival were performed with Kaplan–Meier analysis. The univariate and multivariate Cox regression analyses were performed using the survival package in R[294]. The GraphPad Prism 8.0 software (GraphPad Software, USA) and the R statistical environment version 4.2.0 were used for statistical analysis and graphical representation.

4. Results

4.1 Comparison of the efficiency between the CellSearch and Parsortix systems for the detection and enumeration of CTCs of mCRC patients

To compare the efficiency of the CellSearch and Parsortix systems for the detection and enumeration of CTCs, we selected 46 baseline blood samples of mCRC patients to be analyzed using both technologies. CTCs were identified in CellSearch (**Figure 26**) and Parsortix as $CKs^+/DAPI^+/CD45^-$ and also, in the case of the Parsortix system, as $VIM^+/DAPI^+/CD45^-$ cells. After CTC enumeration, we compared both technologies using the linear regression analysis, which showed high agreement between both systems ($R^2 = 0.91$) for the detection of CTCs (**Figure 27A**). However, the number of detected CTCs was higher with the CellSearch than the Parsortix system (**Figure 27B**). In addition, CellSearch showed higher sensitivity, detecting ≥1 CTC/7.5 mL in 46% (21/46) of the patients analyzed, while Parsortix only the 17% (8/46). Regarding the detection of VIM^+ CTCs with the Parsortix, we only detected VIM^+ CTCs in 3 of the 8 samples that contained CTCs.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

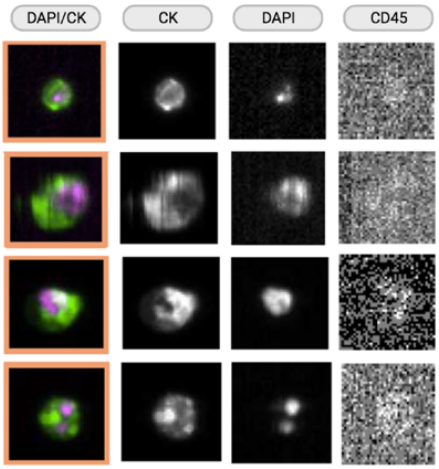


Figure 26. Detection of CTCs in peripheral blood samples of mCRC using the CellSearch system. Representative images of CTCs detected with the CellSearch system in patients with mCRC. CTCs were identified as DAPI^{+}, CK^{+} and CD45^{-}.

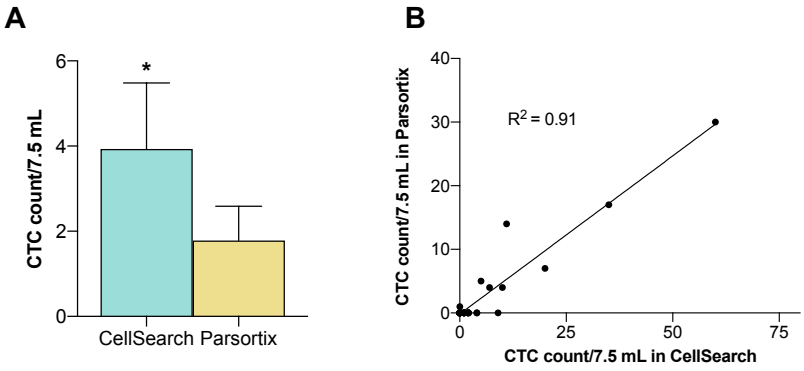


Figure 27. Comparison of the number of CTCs detected using the Parsortix and CellSearch systems in peripheral blood samples of mCRC. **(A)** Number of CTCs/7.5 mL of peripheral blood detected using both enumeration systems. **(B)** Correlation in the number of CTCs detected in both systems. *, p-value <0.05 between CTC counts using Parsortix and CellSearch systems.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.2 Correlation between clinicopathological characteristics of mCRC patients and CTCs enumeration

Since the CellSearch showed higher percentage of CTC detection in blood samples of mCRC, we selected this system to analyze all the mCRC patients of our cohort. Out of the 73 mCRC patients at baseline enrolled in the study, 53% (39/73) were positive for the detection of ≥1CTC/7.5 mL and 32% (23/73) of the patients had peripheral blood counts of ≥3CTCs/7.5 mL. The median number of CTCs per 7.5 mL peripheral blood in the cohort was 1 and ranged from 0 to 83 CTCs.

Based on previous data[295], we used a threshold of ≥3 CTCs/7.5 mL to evaluate the association between the enumeration of CTCs and relevant clinicopathological characteristics in our cohort (**Table 9**). Thus, we observed that the presence of ≥3 CTCs/7.5 mL at baseline was positively associated with the number of metastasis, the presence of metastasis in the lymph nodes and negatively associated with the response to therapy (all p-value <0.05). However, no significant association was found between the presence of ≥3 CTCs/7.5 mL with other clinicopathological variables such as histology, primary tumour location, *KRAS* status or CEA levels amongst others.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Table 9. Clinicopathologic characteristics of the cohort of mCRC patients in relation with their CTC counts.

Characteristics	Baseline CTC count CellSearch (N)		p-value
	<3CTCs/7.5mL	≥3 CTCs/7.5mL	
Age (yrs)			
<65	23	11	0.88
≥65	27	12	
Primary tumour location			
Right colon	17	7	0.76
Left colon/rectum	33	16	
Lymph node afectation			
Yes	12	12	**0.02**
No	38	11	
Number of metastasis locations			
≤1	32	5	**<0.01**
>1	18	18	
Therapy response			
Responder	43	13	**<0.01**
Non-responder	5	9	
***KRAS* status in tissue**			
Wild-type	12	10	
Mutated	38	12	0.07
Unknown		1	
CEA levels			
≤5 ng/mL	11	2	0.16
>5 ng/mL	38	21	

Yrs, years.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

It is important to note that, although the presence of ≥3 CTCs/7.5mL did not show association with the serum levels of CEA, we were able to detect a positive correlation between the number of CTCs and CEA levels (r = 0.26; p = 0.0266) (**Figure 28A**). In addition, we also observed a positive association (r = 0.55; p <0.0001) between the levels of cfDNA in plasma and the number of CTCs present in peripheral blood of mCRC patients (**Figure 28B**).

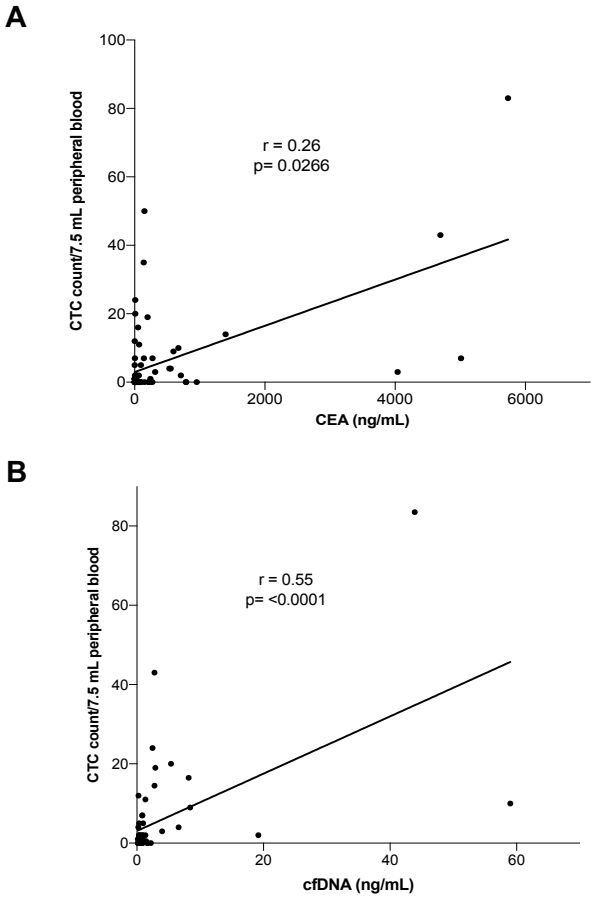


Figure 28. Association of CTC counts at baseline with (A) CEA and (B) cfDNA levels at the moment of diagnosis in mCRC patients.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.3 Evaluation of prognostic value of CTC enumeration in mCRC patients

During the follow-up period, the 73 CRC patients of our cohort were evaluated for disease progression. Out of these patients, 71% (52/73) patients showed progression of the disease, and 56% (41/73) patients died from the disease. The median follow-up time was 11 months (range 1–31 months). A univariate and multivariate cox regression analysis of the association between clinical variables with OS and PFS was performed. Results of the univariate analysis showed that, among the variables analyzed, the presence of ≥3 CTCs/7.5 mL ($p < 0.05$) were significantly associated with unfavorable PFS. The multivariate analysis revealed that counts of ≥3 CTCs/7.5 mL ($p < 0.05$) was an independent prognostic factor for poor PFS. Regarding the OS of all the patients, the presence of ≥3 CTCs/7.5 mL ($p < 0.05$), and *KRAS* mutations in tumour tissue ($p < 0.05$) were significantly associated with unfavourable OS in the univariate analysis and were independent prognostic factors for adverse OS in the multivariant analysis (**Table 10**).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Table 10. Univariate and multivariate analysis for PFS and OS in the mCRC patients cohort.

Variable	Univariate				Multivariate		
	N	HR	95% CI	*p*	HR	95% CI	*p*
PFS							
Tumour location	73	1.32	0.74,2.36	0.35	1.32	0.67,2.58	0.43
No. metastases	73	0.64	0.37,1.11	0.11	0.94	0.49,1.79	0.84
Gender	73	0.96	0.53,1.74	0.89	1.01	0.53,1.94	0.97
Age	73	1.31	0.75,2.28	0.34	1.52	0.84,2.74	0.16
CEA (ng/mL)	72	0.54	0.27,1.06	0.090	0.48	0.22,1.06	0.078
KRAS in tissue	72	1.54	0.81,2.92	0.17	1.38	0.67,2.84	0.37
≥3 CTCs/7.5mL	73	1.93	1.06,3.52	**0.038**	2.70	1.28,5.71	**0.010**
OS							
Tumour location	73	1.79	0.88,3.65	0.12	2.01	0.86,4.70	0.11
No. metastases	73	0.65	0.32,1.32	0.23	1.20	0.51,2.84	0.68
Gender	73	0.71	0.35,1.46	0.37	0.91	0.41,2.02	0.82
Age	73	1.33	0.66,2.68	0.42	1.21	0.55,2.65	0.64
CEA (ng/mL)	72	1.71	0.65,4.53	0.25	1.82	0.63,5.31	0.25
KRAS in tissue	72	2.70	1.03,7.08	**0.025**	2.82	1.03,7.72	**0.027**
≥3 CTCs/7.5mL	73	2.62	1.29,5.30	**0.010**	3.85	1.54,9.62	**0.004**

PFS, Progression free survival; OS, Overall survival; HR, Hazard Ratio; CI, Confidence Interval; *p*, p-value.

Besides, Kaplan–Meier survival curves **(Figure 29**) showed that mCRC patients with ≥3 CTCs/7.5 mL at baseline had significantly (Logrank p-value <0.05) shorter PFS and OS than patients with <3 CTCs/7.5 mL. The median of PFS was 8.8 and 6.3 months in the group of <3 CTCs/7.5 mL and ≥3 CTCs/7.5 mL, respectively. On the other hand, the median of OS in the group of ≥3 CTCs/7.5 mL was 9.2 months, whereas in the group of <3 CTCs/7.5 mL was of 23.3 months.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

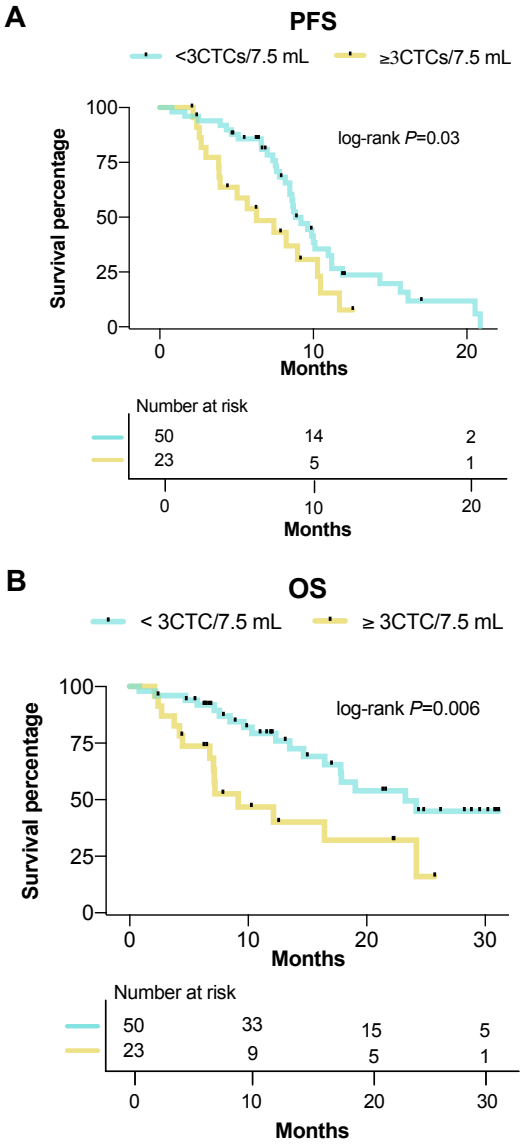


Figure 29. Impact of the number of CTCs on mCRC patient's survival. Kaplan-Meier survival curves of CTCs at baseline for **(A)** progression free survival (PFS) and **(B)** overall survival (OS) of mCRC patients with <3 CTCs/7.5 mL and ≥3 CTCs/7.5 mL of peripheral blood. OS, overall survival; PFS, progression free survival.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.4 Analysis of the DNA methylome of CTCs of mCRC patients with WGBS

After CTC enrichment with the CellSearch of the peripheral blood samples of our mCRC patients, these were processed with the DEPArray system to obtain pure CTCs (**Figure 30**) for their methylome analysis. We characterized with WGBS the DNA methylome portrait of CTCs obtained from 5 mCRC patients in relation to therapy response and the progression of the disease.

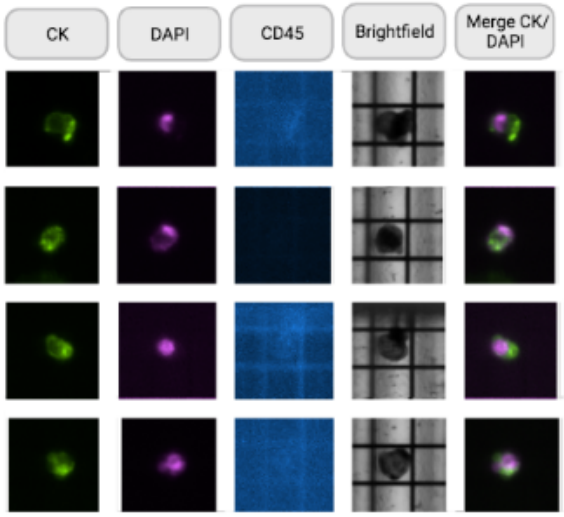


Figure 30. Detection of CTCs from CellSearch enriched samples in the DEPArray system. Representative images of CTCs detected with the DEPArray in patients with mCRC. CTCs were identified as $DAPI^{+}$, CK^{+} and $CD45^{-}$

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.4.1 Analysis of the DNA methylation profile of baseline CTCs in relation to response to FOLFOX treatment

The DNA methylome of baseline CTCs from one responder and one non-responder patient to first-line treatment with FOLFOX in monotherapy was compared. More precisely, 1 CTC from the responder patient and a pool of 7 CTCs from the non-responder patient were analysed by WGBS. This analysis revealed a total of 10,697 DMCpGs (methylation difference >20%) between the baseline CTCs of the responder and the non-responder patient. These DMCpGs were widely distributed across the chromosomes of the genome presenting, the responder patient, a slightly higher number of hypermethylated CpGs (5,646 CpGs; 54% of all DMCpGs) than hypomethylated CpGs (4,880 CpGs; 46% of all DMCpGs) in comparison to the non-responder (**Figure 31A**). Regarding the CpG context (**Figure 31B**), most of the DMCpGs were located in open sea, while the remaining 18% were distributed between CpGIs, shore and shelf regions. According to gene location (**Figure 31B**), most of the methylation differences were distributed in intragenic regions outside gene promoters, and only a minor part of the CpGs were located in enhancer regions. Interestingly, a relevant part of the DMCpGs (26%) were in long non-coding RNA regions (**Figure 31B**).

The hierarchical clustering analysis of the 10,697 DMCpGs identified, which corresponded with 5,717 genes, showed a clearly different DNA methylation signature between the CTCs of the responder and non-responder patient (**Figure 32A**). The top 20 DMCpGs in promoter CpGIs of this signature are indicated in **Supplementary Table 10**. Besides, GO analysis (**Figure 32B**) revealed that the genes of this epigenetic signature were involved in relevant cancer pathways, including mainly metabolic, PI3K/AKT, MAPK, Wnt, Integrin and Cadherin pathways.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

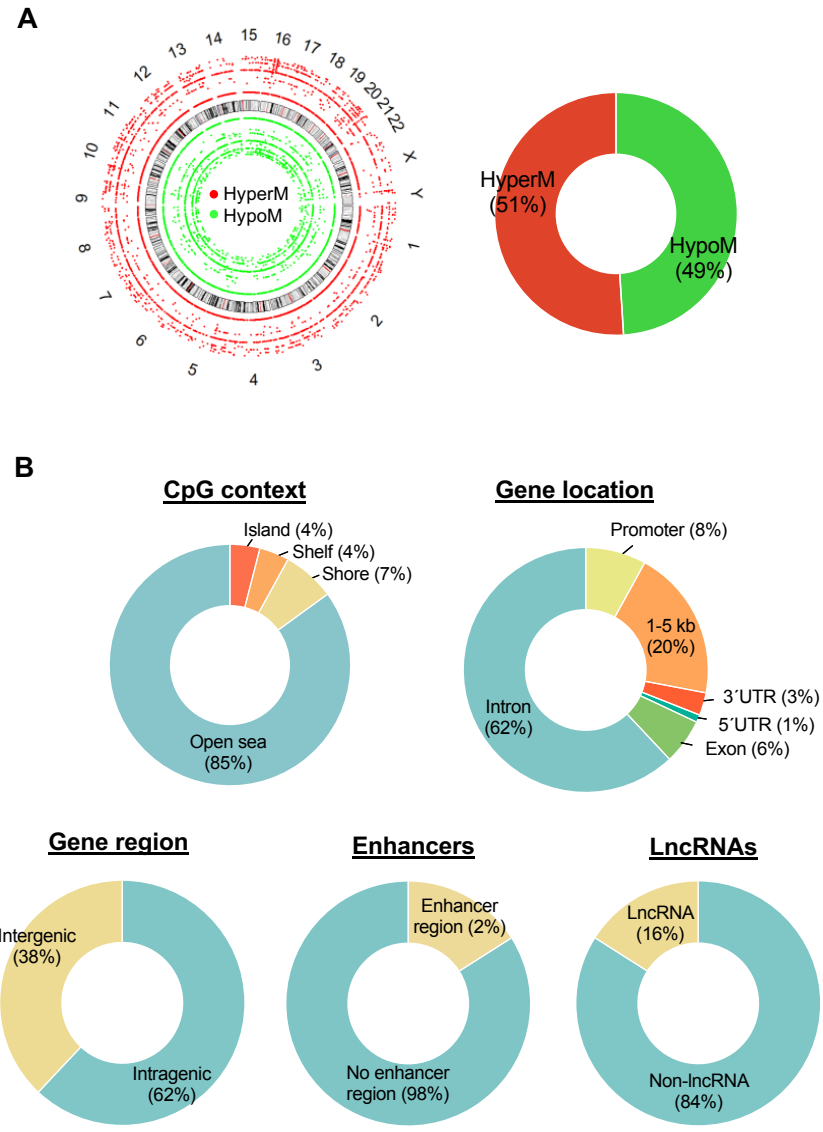
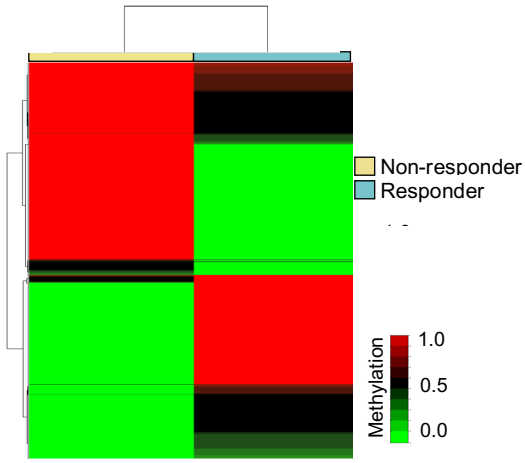


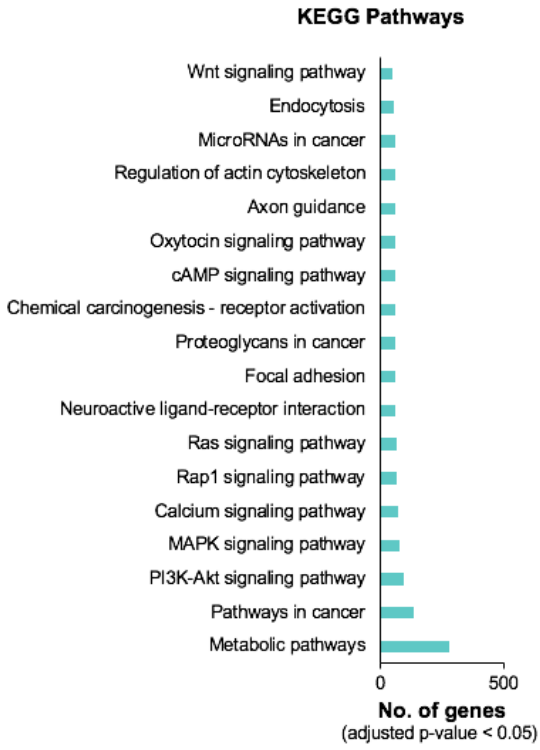
Figure 31. Description of the 10,697 DMCpGs found in the baseline CTCs of the responder and the non-responder patient to FOLFOX treatment. Distribution of the DMCpGs according to **(A)** chromosome location and methylation status, and **(B)** CpG context, gene location, gene region, enhancers and lncRNAs. LncRNAs, long non-coding RNAs.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

A



B



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

C

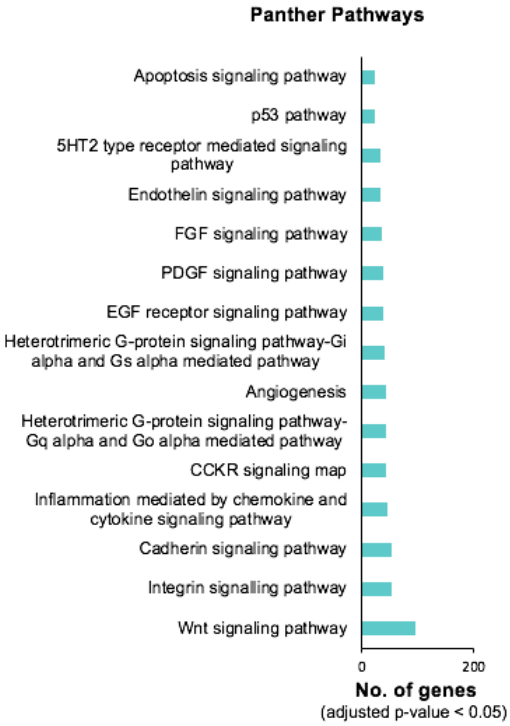


Figure 32. DNA methylation pattern of the responder and non-responder patient to FOLFOX. (A) Hierarchical clustering heatmap of the 10,697 DMCpGs (methylation difference >20%) in the responder patient respect to the non-responder patient. **(B-C)** GO analysis representing some of the most cancer relevant KEGG and Panther pathways based on the 10,697 DMCpGs between the CTCs of the responder and the non-responder patient to FOLFOX treatment.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Of note, in addition to the methylation differences previously described in the genomic DNA of CTCs, WGBS analysis also allowed the identification of 171 DMCpGs (methylation difference >20%) in the mitochondrial DNA (mtDNA) of the CTCs of the responder and non-responder patient. In particular, the mtDNA in the CTCs of the responder patient showed more hypermethylated CpGs than the non-responder (**Figure 33**).

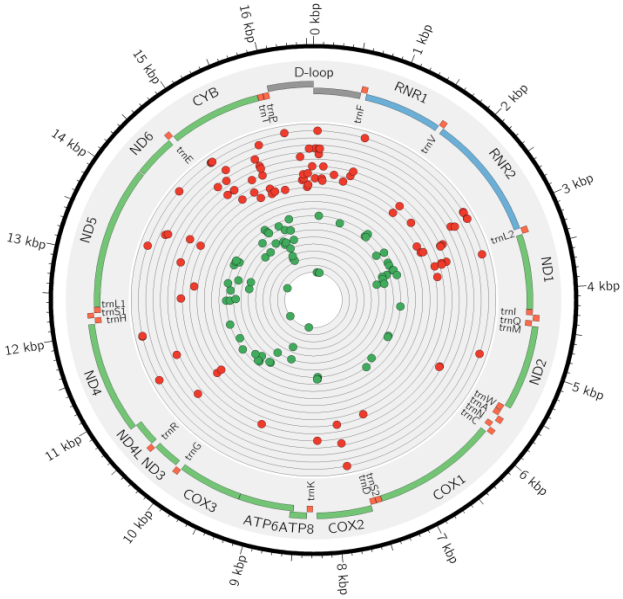


Figure 33. Circos plot showing the methylation differences in mitochondrial DNA between the responder and non-responder patient to FOLFOX. Gene names are listed around the outside of the plot and red and green dots indicate hypermethylated and hypomethylated CpGs, respectively, in the responder patient compared to the non-responder.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.4.2 Evaluation of the DNA methylation profile of baseline CTCs in relation with response to FOLFOX and bevacizumab treatment

We analysed the DNA methylome of baseline CTCs from one responder and one non-responder patient to the first-line treatment based on the combination of FOLFOX and bevacizumab. Thus, 1 CTC from the responder patient and a pool of 37 CTCs from the non-responder patient were analysed using WGBS. Sequencing data revealed a total of 3,458 DMCpGs (methylation difference >20%) between the baseline CTCs of the responder and non-responder patient. These DMCpGs were widely located along the chromosomes of the genome, having the responder patient a very similar number of hypermethylated CpGs (1,709 CpGs; 51% of all DMCpGs) than hypomethylated CpGs (1,644 CpGs; 49% of all DMCpGs) in comparison to the non-responder (**Figure 34A**). In relation to the CpG context (**Figure 34B**), most of the DMCpGs (85%) were distributed in open sea, and the remaining ones were located in CpGIs, shore and shelf regions. Regarding gene location (**Figure 34B**), most of the methylation differences were distributed in intragenic regions and outside gene promoters. Only a minor part of the DMCpGs were located in enhancer regions, while a relevant part of the methylation differences (16%) was distributed in long non-coding RNA regions (**Figure 34B**).

By means of hierarchical clustering analysis, the 3,458 DMCpGs, which corresponded with 1,994 genes, revealed a DNA methylation signature that clearly segregated the CTCs of the responder and the non-responder patient (**Figure 35A**). **Supplementary Table 11** shows the top 20 DMCpGs of this signature. In addition, by GO analysis (**Figure 35B**), we observed that the genes of this epigenetic signature participate in relevant cancer pathways related to metabolism, angiogenesis, inflammation, and PI3K/AKT, MAPK, Wnt and Cadherin signaling pathways.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

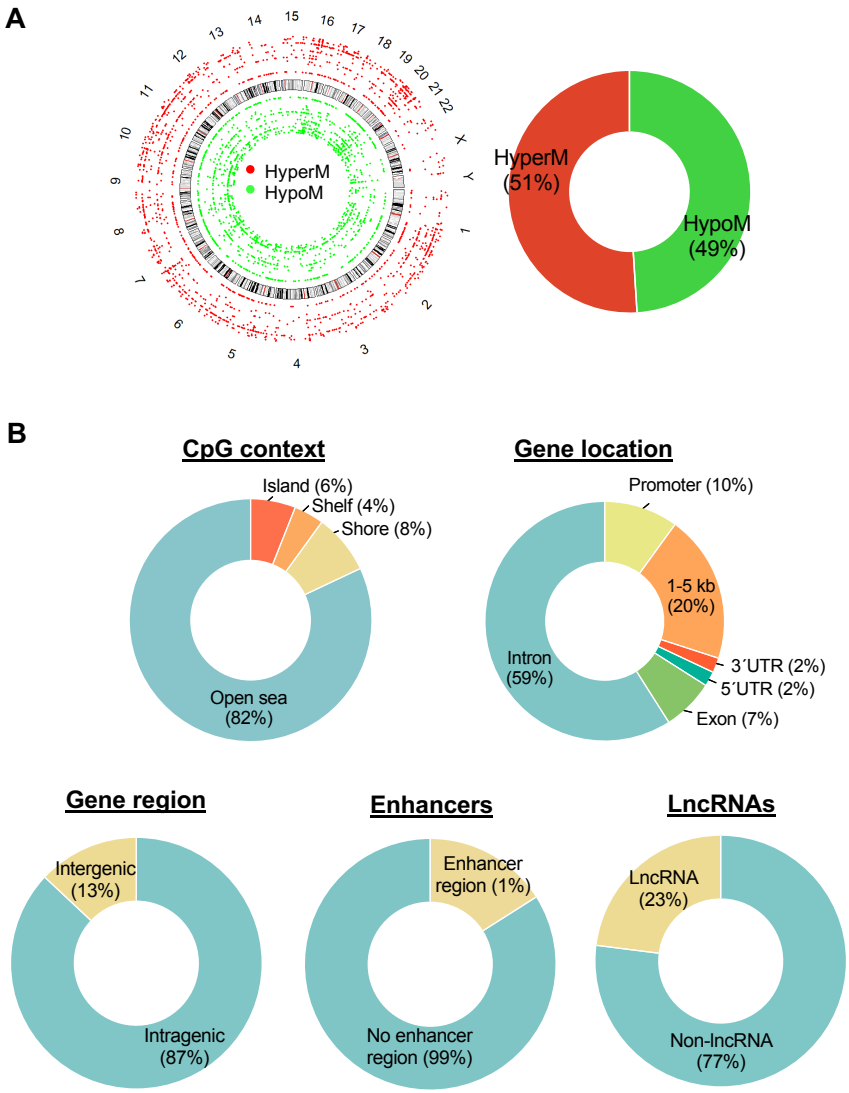
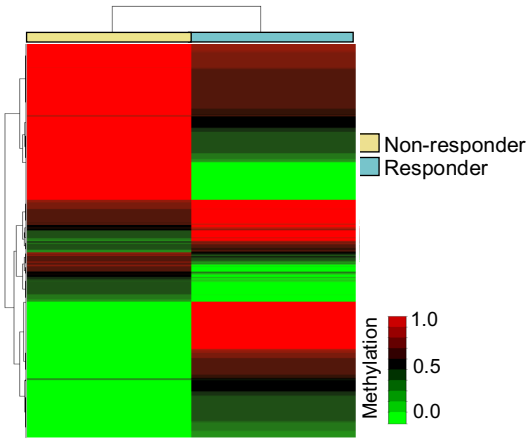


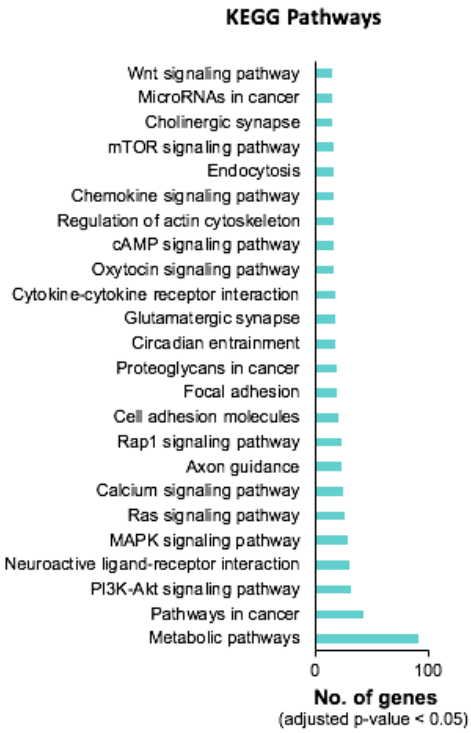
Figure 34. Description of the 3,458 DMCpGs found in the baseline CTCs of the responder and the non-responder patient to FOLFOX and bevacizumab treatment. Distribution of the DMCpGs according to **(A)** chromosome location and methylation status, and **(B)** CpG context, gene location, gene region, enhancers and lncRNAs. LncRNAs, long non-coding RNAs.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

A



B



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

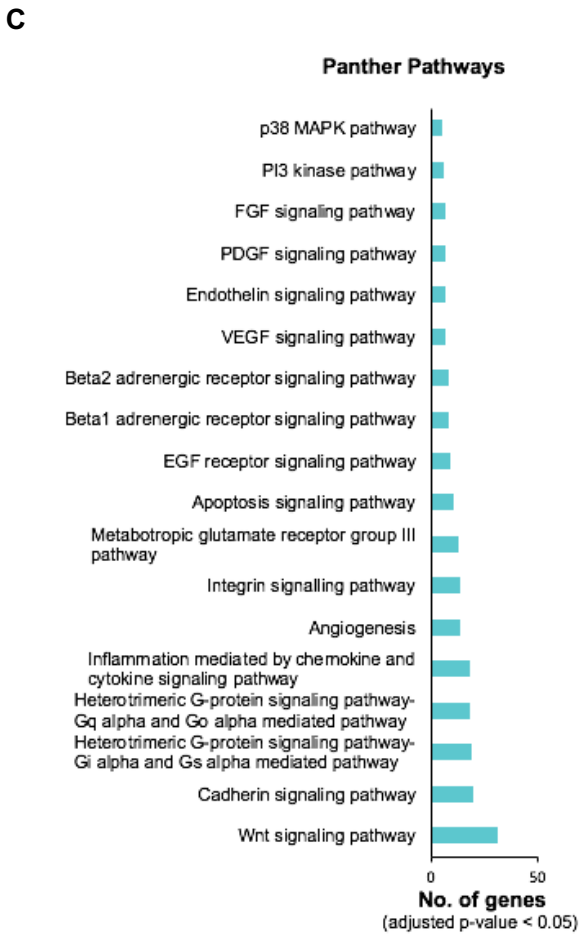


Figure 35. DNA methylation pattern of the responder and non-responder patient to FOLFOX and bevacizumab. (A) Hierarchical clustering heatmap of the 3,458 DMCpGs (methylation difference >20%) in the responder patient respect to the non-responder patient. **(B-C)** GO analysis representing some of the most cancer relevant KEGG and Panther pathways based on the 3,458 DMCpGs between the CTCs of the responder and the non-responder patient to FOLFOX and bevacizumab treatment.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Moreover, this WGBS analysis, besides revealing relevant methylation differences in the genomic DNA of CTCs, also revealed methylation differences in the mtDNA higher than 20% in 105 CpGs between the responder and non-responder patient. Most of these DMCpGs were hypermethylated in the CTCs of the responder patient respect to the non-responder patient (**Figure 36**).

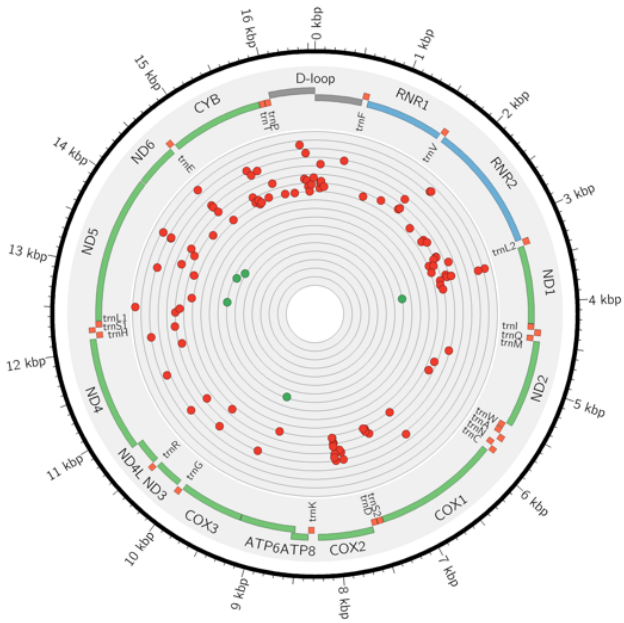


Figure 36. Circos plot showing the methylation differences in mitochondrial DNA between the responder and non-responder patient to FOLFOX and bevacizumab. Gene names are listed around the outside of the plot and red and green dots indicate hypermethylated and hypomethylated CpGs, respectively, in the responder patient compared to the non-responder.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.4.3 Characterization of the DNA methylation profile of CTCs at baseline and after the progression to FOLFOX and bevacizumab

The DNA methylome of CTCs obtained from one mCRC patient at the time of diagnosis before treatment initiation, and at the moment of progression after initiating first-line treatment with the combination of FOLFOX and bevacizumab was compared. In particular, we analysed a pool of 37 CTCs at baseline and a pool of 22 CTCs at the progression of the disease. This comparison revealed a total of 11,717 DMCpGs (methylation difference >20%) between the CTCs at baseline and at progression. These DMCpGs were widely distributed across the chromosomes of the genome, presenting baseline CTCs a slightly higher number of hypermethylated CpGs (6,324 CpGs; 54% of all DMCpGs) than hypomethylated CpGs (5,337 CpGs; 46% of all DMCpGs) in comparison to the CTCs at progression (**Figure 37A**). Regarding CpG context (**Figure 37B**), most of the DMCpGs (86%) were located in open sea, while the remaining 14% were distributed between CpGIs, shore and shelf regions. According to gene location (**Figure 37B**), most of the methylation differences were distributed in intragenic regions outside gene promoters. Only a minor part of the CpGs were located in enhancer regions. Interestingly, a relevant part of the DMCpGs (16%) were distributed in long non-coding RNA regions (**Figure 37B**).

Besides, the hierarchical clustering analysis of the 11,717 DMCpGs identified, which corresponded with 5,631 genes, showed a clearly different DNA methylation signature between CTCs at baseline and at progression (**Figure 38A**). The top 20 DMCpGs in promoter CpGIs of this signature are indicated in **Supplementary Table 12**. The GO analysis (**Figure 38B**) disclosed that the genes of this epigenetic signature were involved in relevant cancer processes, including pathways related to metabolism, inflammation or angiogenesis, and PI3K/AKT, Wnt, Integrin and Cadherin signaling.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

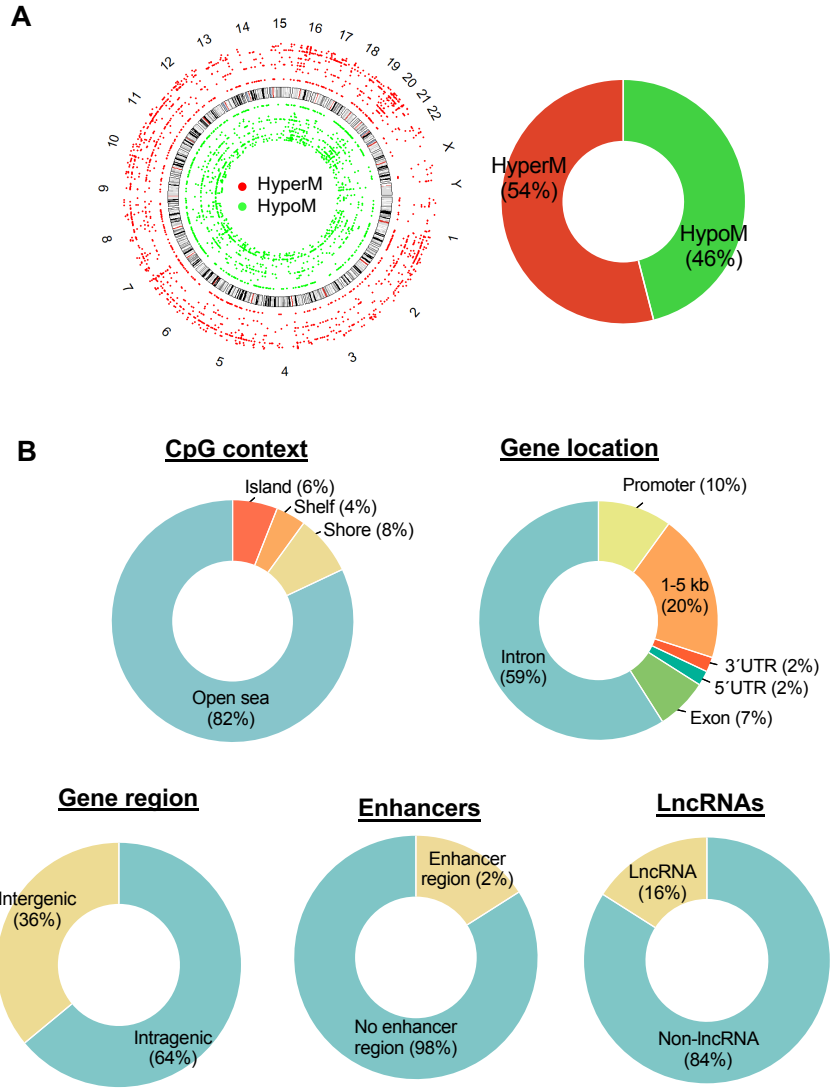
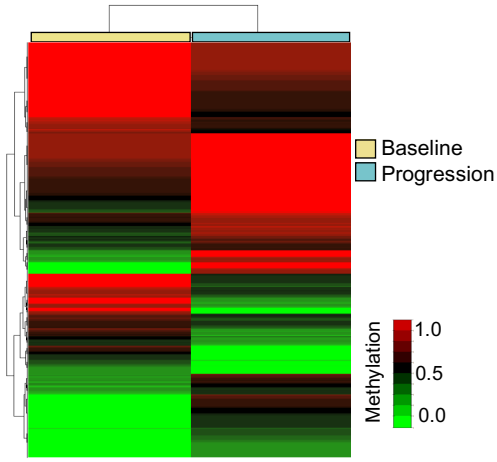


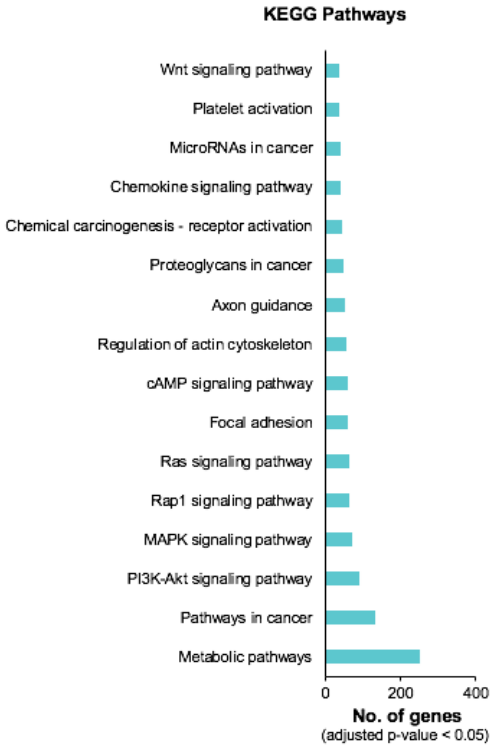
Figure 37. Description of the 11,717 DMCpGs found in the CTCs at baseline and at progression after FOLFOX and bevacizumab treatment initiation. Distribution of the DMCpGs according to **(A)** chromosome location and methylation status, and **(B)** CpG context, gene location, gene region, enhancers and lncRNAs. LncRNAs, long non-coding RNAs.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

A



B



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

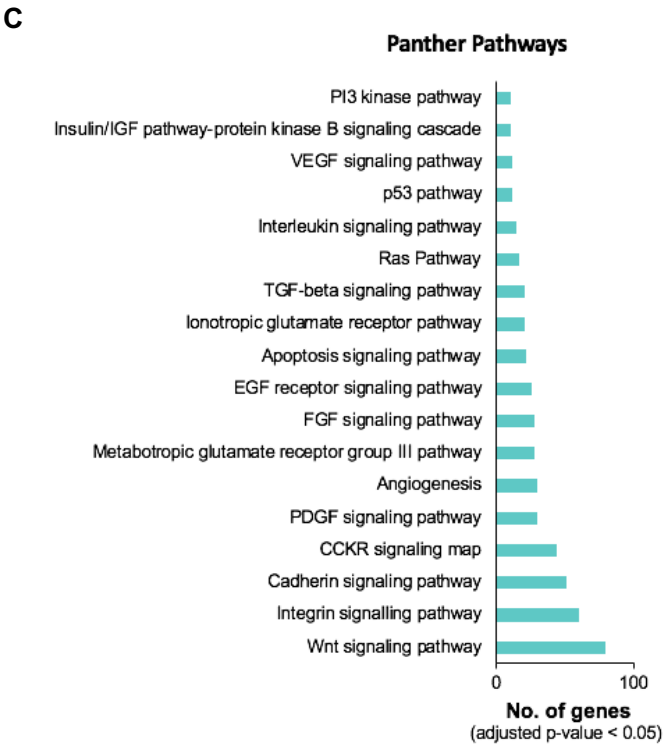


Figure 38. DNA methylation pattern of the CTC at baseline and at progression after FOLFOX and bevacizumab treatment initiation of one mCRC patient. (A) Hierarchical clustering heatmap of the 11,717 DMCpGs (methylation difference >20%) in the CTCs at baseline respect to the CTCs after treatment initiation at the moment of progression. **(B-C)** GO analysis representing some of the most cancer relevant KEGG and Panther pathways based on the 11,717 DMCpGs between the CTCs of one patient at baseline and at the moment of progression after treatment initiation with FOLFOX and bevacizumab.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Relevantly, WGBS analysis also discovered 56 DMCpGs (methylation difference >20%) between the mtDNA of the CTCs at baseline and at the time of progression. Among these DMCpGs of mtDNA, the number of hypermethylated CpGs was higher at baseline than at the time of progression of the disease (**Figure 39**).

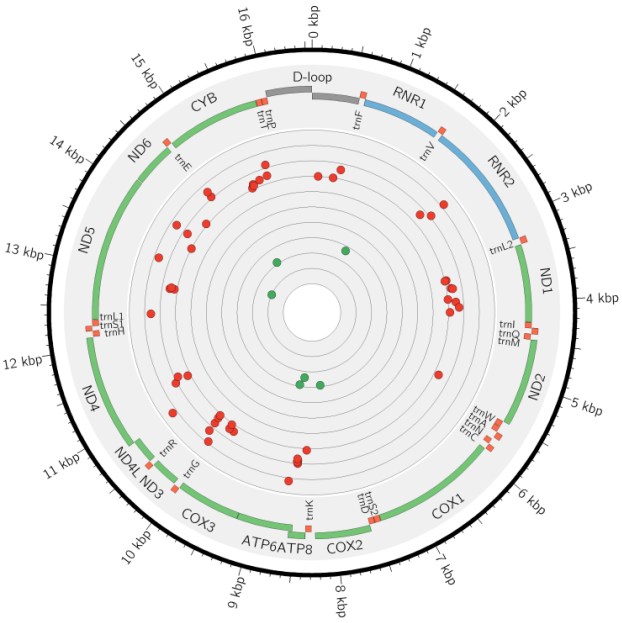


Figure 39. Circos plot showing the methylation differences in mitochondrial DNA between the CTCs at baseline and the CTCs after treatment initiation at the moment of progression of one mCRC patient treated with FOLFOX and bevacizumab. Gene names are listed around the outside of the plot and red and green dots indicate hypermethylated and hypomethylated CpGs, respectively, in the CTCs baseline compared to the resistant CTCs.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

5. Discussion

Both isolation and characterization of CTCs are key challenges for their clinical utility[296]. Scarcity and heterogeneity of these CTCs makes very challenging their isolation from patient blood samples with high purity and efficiency, being also difficult to stablish common criteria to target for their isolation[297]. Current methods for CTC isolation are mainly based on the differential characteristics between CTCs and blood cells[298] and despite the emergence of novel technologies for CTC isolation, the CellSearch technology is the only method approved for monitoring patients with metastatic colon cancer[296,299]. This CellSearch system is based on the recognition of the EpCAM antibody on the surface of CTCs for their isolation[298]. However, in 35 % of patients, even with advanced metastatic disease, CTCs are not detectable, probably due to the EMT process[296]. During the EMT, epithelial cells change their phenotype to mesenchymal-like cells and lose the expression of characteristic epithelial markers[300]. In fact, studies have shown CTCs can be present in multiple phenotypes and subpopulations: epithelial, mesenchymal, and a hybrid epithelial–mesenchymal state[120]. Therefore, CTC isolation methods based on the expression of these epithelial markers may miss other cell subpopulations[296]. In this sense, an outbreak of new microfluidic-based technologies, such as the Parsortix system, has emerged to address this problem[301]. The Parsortix is a microfluidic device, that relies on cells size and deformability, overlooking the epitope dependency of the CellSearch technology[105].

The emergence of different approaches to detect and isolate CTCs from peripheral blood, highlights the need to find the most suitable method for CTC analysis[296]. In line with this, our work revealed that the CellSearch system has higher rates of CTC detection in the peripheral blood of mCRC patients than Parsortix, indicating that CellSearch is a better approach for the enumeration and enrichment of CTCs in advanced CRC cancer. The lower

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

sensitivity showed by Parsortix in our cohort could be influenced by the fact that CTCs of CRC patients have a low size in comparison with CTCs of other common tumour types, such as breast, prostate or bladder cancer[302], which could make it more difficult to retain this type of colorectal cells in this microfluidic device.

A wide range of studies have revealed that CTCs can explain relevant clinical features, such as tumour invasion and lymph node metastasis in patients with CRC[303]. In this sense, we observed that the presence of at least 3 baseline CTCs in 7.5 mL of peripheral blood of mCRC patients could indicate lymph node affectation and the existence of more than one distant metastasis. Importantly, this CTC threshold at baseline also showed utility to identify those patients who could benefit from therapy. Moreover, baseline CTC counts increased along with cfDNA and serum CEA levels, which are biomarkers that have already shown association with the clinical outcome of cancer patients[303–305]. Besides, our study showed that the presence of at least 3 CTCs in 7.5 mL of peripheral blood is an independent prognostic factor of PFS and OS, which is in accordance with previous published data[112,306–308]. Altogether, these results confirm the importance of CTC baseline enumeration to evaluate and predict the outcome of mCRC patients.

In addition to CTC enumeration, recent studies have explored the molecular characteristics of these cells to better understand the molecular biology of metastasis[309], and to identify new valuable clinical tools for the management of cancer patients[296], including in relation to the response to therapy. In line with this, we identified a distinctive methylation signature in the genomic DNA of CTCs at baseline with the potential to select those patients who who could benefit from the treatment with FOLFOX. Moreover, we also revealed a methylation signature in the genomic DNA of CTCs at baseline characteristic of the response to treatment with FOLFOX combined with bevacizumab. The identification of these two epigenetic signatures

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

suggests that aberrant DNA methylation of baseline CTCs may be useful as a biomarker to predict therapy response in mCRC patients. In line with this, other authors have previously found that alterations in DNA methylation can predict therapy response. This is, for example, the case of the aberrant methylation of *MGMT* in tumour tissue samples of gliomas, which has shown utility to select those patients who will benefit from therapy with alkylating agents[310]. Similarly, the hypermethylation of *TFAP2E* in CRC tumour tissues has been associated with clinical non-responsiveness to chemotherapy of patients[311].

In this work we also identified a methylation signature in the genomic DNA of CTCs associated with the progression of the disease after treatment initiation with FOLFOX combined with bevacizumab in mCRC patients, suggesting that this treatment is able to induce the reprogramming of DNA methylation in CTCs leading to therapy resistance. In this sense, epigenetic reprogramming of cancer cells has been postulated as a crucial mechanism associated with the acquisition of therapy resistance[312]. Therefore, liquid biopsy makes possible to monitor the emergence of resistant clones[313], making the analysis of DNA methylation in CTCs a good approach for early detection of resistance in mCRC patients. Additionally, increasing knowledge about the mechanisms that underly drug resistance in CRC may provide clinicians the opportunity to tailor treatment strategies accordingly[313].

Interestingly, the epigenetic signatures associated with baseline therapy response and development of resistance identified in this work were associated with relevant cancer pathways such as metabolic alterations, PI3K/AKT and Wnt signaling, among others. Epigenetic alteration of these pathways in CTCs reinforces their relevant role in the development of therapy resistance and suggests the possibility that the aberrantly methylated genes found in relation with these pathways could serve as potential therapeutic

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

targets. In this scenario, epidrugs could represent an attractive option to be used in CRC as a co-treatment along with conventional therapies[312].

Intriguingly, most of the genomic methylation changes found in the epigenetic signatures identified in this work were located in CpG-poor locations and in intragenic regions outside promoters. Traditionally, in cancer DNA methylation has been studied in association with promoter CpGIs of coding regions[314]. However, in last years the development of `omic´ technologies has started to reveal the importance of methylation in other less explored parts of the genome[315]. In this sense, genome-scale mapping of methylation outside promoter CpGIs of coding genes might also contain relevant information in CTCs to evaluate therapy response of CRC patients. Besides, we found a relevant percentage of methylation changes in CTCs located in genomic regions of lncRNAs. These non-coding molecules have a relevant regulatory role in cells, are modulated by methylation, and have shown their utility in assessing response to therapy and detection of CRC[137,228]. This suggests that the study of methylation in non-coding regions of the CTCs, such as lncRNAs, might be useful to evaluate therapy response and resistance, and to design new epidrugs for CRC.

In addition to genomic DNA, we also detected relevant methylation changes in mtDNA of CTCs related to initial treatment response and resistance after treatment of CRC patients. The mtDNA is a circular genome that encodes 13 proteins, 22 tRNAs, and 12S and 16S rRNAs, which play a fundamental role in the mitochondria and cell metabolism[316]. Importantly, several works have identified that molecular alterations of mtDNA may be significant for cancer disease[317]. In fact, a recent work has revealed the association of low mtDNA copy number in chemotherapy resistance of gastrointestinal tumours[318]. In addition, methylation of mtDNA has shown to be deregulated in cancer[319]. However, the aberrant methylation of mtDNA in CTCs represents a completely unexplored field. It is recognized that the

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

interplay between mitochondria metabolism and epigenetic mechanisms has a critical role in tumour initiation and evolution[320], highlighting the importance of studying the aberrantly methylated mtDNA to identify new predictive and resistance biomarkers, and new therapeutic targets for CRC. In this sense, it has been recently proposed that targeting mitochondrial metabolism and epigenetics are promising therapeutic strategies for tumour treatment[320].

Of note, this work has stablished a workflow to characterize the DNA methylome of CTCs isolated from peripheral blood of cancer patients. CTC enrichment platforms, such as CellSearch, need an additional step to isolate pure CTCs for molecular analysis, since contaminating DNA from leukocytes is an important obstacle for precise and efficient molecular characterization of CTCs[107]. In this sense, the DEPArray allows isolation of pure, single, rare cells from a heterogeneous sample[107]. Different studies have already showed successful molecular characterization of CTCs isolated with this technology[131,321–325]. Thus, in our approach, blood samples of CRC patients were first used for CTC enrichment with CellSearch system, then single or pooled pure CTCs were isolated with DEPArray and, finally, the DNA methylome of these CTCs was analyzed with WGBS. Our work provided a useful strategy for the epigenomic characterization of CTCs in cancer patients, which could be implemented for future studies with CTCs in several research and clinical settings. Importantly, this strategy was able to analyze the DNA methylome of single CTCs, which is of great importance for the in-depth study of CTC heterogeneity and the understanding of the development of cancer and its progression[326].

Although the DNA methylation analysis of this work has been evaluated in only five selected patients of mCRC, the results obtained pave the way to further explore, in more detail, the genomic and mitochondrial DNA methylome of CTCs in relation to therapy response and the development of resistance not only in CRC but also in other tumour types.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

In summary, in this chapter we demonstrated that the CellSearch system has higher performance than Parsortix for the enrichment and enumeration of CTCs in peripheral blood of mCRC. We also confirmed that the number of baseline CTCs detected by CellSearch is associated with relevant clinicopathological characteristics and is an independent prognostic factor of worse PFS and OS in mCRC. Moreover, we also stablished a workflow useful to isolate pure CTCs and characterize their DNA methylome in mCRC patients. This workflow revealed that the DNA methylation profile of baseline CTCs could be useful to select the patients who could potentially benefit from the treatment with FOLFOX in monotherapy or in combination with bevacizumab. The selective pressure of treatment with FOLFOX and bevacizumab was also able to induce changes in the DNA methylation profile of CTCs associated with the development of therapy resistance. In addition, relevant methylation changes in mtDNA of CTCs were detected in association with treatment response and resistance. These results could unveil new potential therapeutic targets and epigenetic biomarkers of CTCs able to predict therapy response and the development of resistance, which could assist in the clinical practice to improve the outcome of mCRC patients

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

OVERALL DISCUSSION

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

OVERALL DISCUSSION

In the management of patients with advanced CRC there is the need for biomarkers able to predict treatment response, as well as to identify new mechanisms of resistance and therapeutic targets that allows to achieve a more personalized medicine. In this sense, liquid biopsy, defined as the study of blood components to determine the properties of solid tumours, has revolutionized the field of clinical oncology.

It is especially interesting the analysis of CTCs due to the great versatility of studies that these cells enable to perform. Of note, CTCs have gained increased interest as diagnostic, prognostic and predictive biomarkers in a great variety of cancer types. Clinical utility of these cells goes beyond simple enumeration and it is now evident that their characterization at molecular level is equally relevant nor only to shed some light on the metastatic process or the mechanisms of resistance to therapy, but also for the dynamic monitoring of tumour progression over time. However, these types of CTC analyses have been limited as a result of the technical difficulties in their isolation due to its scarcity and their heterogeneous biological features.

Most studies in cancer are done using primary or metastatic cancer tissues or cell lines derived from tumours. However, taking into account that CTCs are the main drivers of metastasis, there is a major part of the metastatic process that has been overlooked. This occurs due to the lack of methods and models to study these CTCs, which difficults their detection and characterization. In line with this, the development of models to study these rare cells is of special importance to discover new relevant information of the metastatic cascade and also to identify new mechanisms and biomarkers that could have translation into the clinic to improve patient management. Recent

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

advances in cell culture technologies have opened a new avenue to develop permanent cell lines derived from CTCs.

Thus, in this work, we performed a genome-wide characterization of the methylome of human colorectal cancer-derived CTC cell lines and CTCs directly isolated from patients with mCRC. Along the two first chapters, the methylome of nine cells was studied, revealing differential methylation patterns in relation with colorectal primary and metastatic tumour cell lines, but also in relation to response and resistance to CRC therapy. Moreover, the study of these cell lines has allowed the detection of particular aberrantly methylated genes that could be potential predictive and resistance biomarkers that may help to guide CRC treatment. Therefore, this CTC cell line profiling has revealed the importance of CTC models to get information that is relevant for the clinic. Besides CTC cell lines, we were also able to isolate CTCs derived from CRC patients and study their methylation profile, revealing the importance of genomic and mitochondrial DNA to predict therapy response and detect the development of resistances. This type of epigenetic approach is specially challenging since when studying the methylome of single CTCs or of a limited number of cells, bisulfite conversion of DNA employs harsh reaction conditions, causing degradation and loss of DNA and reduced sequence quality, which results in loss of final data yield.

All in all, results in this thesis reveal new information about the biology of these rare cells, as well as novel aberrantly DNA methylated genes related to the formation of metastasis. Moreover, this study enabled to identify new therapeutic targets, resistance mechanisms and potential epigenetic biomarkers that could benefit mCRC patients in the clinical practice. However, further work is needed to validate the aberrantly methylated genes identified in a large cohort of mCRC patients.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

CONCLUSIONS

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CONCLUSIONS

The work conducted in this doctoral thesis allowed us to conclude that:

1. The colorectal metastasis-competent cell line CTC-MCC-41 displays a global different DNA methylation program than primary and metastatic colorectal tumour cell lines.

2. The transcriptional program of CTC-MCC-41 cells is regulated by DNA methylation. The identified epigenetically regulated genes represent potential biomarkers and therapeutic targets for metastasis-competent CTCs.

3. The evaluation of the DNA methylation profile in CTC cell lines resistant to 5-fluorouracile-based treatment reveals that selective pressure of therapy is able to change the methylation pattern of CTCs leading to a more hypomethylated profile.

4. The DNA methylation signature of CTC cell lines resistant to 5-fluorouracile-based treatment is enriched in genes involved in relevant cancer pathways, such as PI3K/AKT, MAPK and metabolism. Among the identified epigenetically regulated genes, *AP2M1* represents a new potential biomarker of resistance and a therapeutic target for CRC.

5. In mCRC patients, CellSearch showed higher performance than Parsortix for the enrichment and enumeration of CTCs from peripheral blood. The baseline detection of ≥3 CTCs by CellSearch was associated with metastasis and worse therapy response and was an independent prognostic factor of poor PFS and OS, reinforcing the clinical value of CTC enumeration in mCRC.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

6. The DNA methylation profile of CTCs from mCRC patients before starting therapy was able to classify the patients according to the response to FOLFOX in monotherapy and in combination with bevacizumab, indicating that the epigenetic portrait of CTCs could be a predictive biomarker of therapy response.

7. The treatment of mCRC patients with FOLFOX and bevacizumab was able to induce changes in the DNA methylation profile of CTCs associated with the development of resistance, which could be useful as epigenetic biomarker to monitor therapy response.

8. In mCRC patients, the mtDNA methylation profile of CTCs is associated with baseline therapy response and the development of resistance, supporting the possibility to be a biomarker for precision oncology.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

SUPPLEMENTARY MATERIAL

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Top 20 differentially methylated CpGs identified in CTC-MCC-41 in comparison to HT29 cells.

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg12558358	3	23244824	*UBE2E2;UBE2E2*	1stExon;5'UTR	Island	-0.980
cg15816080	7	65670215	*TPST1*	TSS200	Island	-0.978
cg02478789	19	9649310	*ZNF426*	TSS200	Island	0.976
cg21028251	3	133614589	*RAB6B;RAB6B*	1stExon;5'UTR	Island	-0.975
cg15361065	3	55522301	*WNT5A*	TSS1500	Island	0.974
cg11777419	14	104604401	*KIF26A*	TSS1500	Island	-0.974
cg10975049	10	72432542	*ADAMTS14;ADAMTS14*	TSS200;TSS200	Island	0.974
cg24091001	2	172949870	*DLX1;DLX1*	TSS1500;TSS1500	Island	-0.973
cg18202521	1	19600702	*AKR7L;AKR7L*	TSS200;TSS200	Island	0.973
cg21578090	17	4488018	*SMTNL2;SMTNL2*	1stExon;5'UTR	Island	-0.973
cg17192454	2	74726651	*LBX2;LBX2*	1stExon;Body	Island	0.973
cg09247193	17	47075095	*IGF2BP1;IGF2BP1;IGF2BP1;IGF2BP1*	5'UTR;5'UTR;1stExon;1stExon	Island	-0.971
cg12034979	12	133532882	*ZNF605;ZNF605;ZNF605*	1stExon;TSS200;5'UTR	Island	0.970
cg27047283	X	135333567	*MAP7D3;MAP7D3*	1stExon;5'UTR	Island	-0.970
cg12087304	4	122873018	*TRPC3*	TSS200	Island	-0.970
cg16112120	3	133614580	*RAB6B;RAB6B*	1stExon;5'UTR	Island	-0.970
cg23755113	7	65670237	*TPST1*	TSS200	Island	-0.970
cg24027179	17	8534165	*MYH10*	TSS200	Island	-0.969
cg12798157	1	19600719	*AKR7L;AKR7L*	TSS200;TSS200	Island	0.969
cg20591405	3	133614578	*RAB6B;RAB6B*	5'UTR;1stExon	Island	-0.968

[1]Identification of the CpG according to EPIC array; [2]Chromosome; [3]Δβ-values (β-value CTC-MCC-41 - β-value HT29). All CpGs in the table are significant (FDR adjusted p-value <0.05) and are arranged according to their absolute Δβ-value.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 2. Top 20 differentially methylated CpGs from CpG islands or shore regions of gene promoters identified in CTC-MCC-41 in comparison to HT29 cells.

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg19127343	19	57106620	*ZNF71*	TSS200	Island	-0.979
cg11111131	X	105855108	*CXorf57*	TSS200	Island	-0.977
cg19398192	X	105855052	*CXorf57*	TSS200	Island	-0.977
cg23216956	19	57106627	*ZNF71*	TSS200	Island	-0.973
cg25230111	19	56159817	*CCDC106*	5'UTR	Island	0.971
cg24447438	18	21718979	*CABYR;CABYR;CABYR;CABYR;CABYR;CABYR;CABYR;CABYR;CABYR;CABYR*	1stExon;1stExon;1stExon;5'UTR;5'UTR;5'UTR;TSS1500;1stExon;5'UTR;5'UTR;1stExon	Island	-0.970
cg11739758	17	46670663	*LOC404266;LOC404266;LOC404266;HOXB5;LOC404266*	Body;Body;Body;1stExon;Body	Island	0.970
cg11757144	7	91510093	*MTERF*	TSS200	Island	0.969
cg07030010	6	74405929	*CD109;CD109;CD109;CD109;CD109;CD109;LOC101928489*	5'UTR;5'UTR;5'UTR;1stExon;1stExon;1stExon;TSS200	Island	-0.969
cg21159778	9	117266918	*DFNB31;DFNB31*	TSS1500;1stExon	Island	-0.969
cg27586390	X	105855176	*CXorf57;CXorf57*	1stExon;5'UTR	Island	-0.969
cg18496937	22	19702076	*SEPT5;SEPT5*	1stExon;5'UTR	Shore	-0.967
cg04210573	X	71351784	*RGAG4;NHSL2*	TSS200;Body	Island	-0.967
cg07735777	7	140715095	*MRPS33;MRPS33*	TSS1500;TSS1500	Shore	0.966
cg00018261	X	117957689	*ZCCHC12*	TSS200	Island	-0.965

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 2. Continuation

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg25249518	6	119399928	*FAM184A;FAM184A*	5'UTR;TSS200	Island	-0.965
cg19452316	1	206680966	*RASSF5;RASSF5*	1stExon;1stExon	Island	-0.964
cg03899721	8	8748792	*MFHAS1*	1stExon	Island	0.964
cg11218091	10	124638976	*LOC399815;FAM24B*	TSS200;5'UTR	Island	-0.962
cg05168842	11	61584319	*FADS1*	1stExon	Island	-0.962

[1]Identification of the CpG according to EPIC array; [2]Chromosome; [3]Δβ-values (β-value CTC-MCC-41 - β-value HT29). All CpGs in the table are significant (FDR adjusted p-value <0.05) and are arranged according to their absolute Δβ-value.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 3. Top 20 differentially methylated CpGs identified in CTC-MCC-41 in comparison to COLO205 cells.

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg14052728	2	95941719	*PROM2;PROM2;PROM2*	Body;Body;Body		0.981
cg12558358	3	23244824	*UBE2E2;UBE2E2*	1stExon;5'UTR	Island	-0.980
cg15816080	7	65670215	*TPST1*	TSS200	Island	-0.978
cg02478789	19	9649310	*ZNF426*	TSS200	Island	0.976
cg21028251	3	133614589	*RAB6B;RAB6B*	1stExon;5'UTR	Island	-0.975
cg08935899	11	6651712	*DCHS1*	Body	Island	-0.975
cg26472802	21	45713719	*AIRE;AIRE*	Body;Body	Island	0.975
cg15361065	3	55522301	*WNT5A*	TSS1500	Island	0.974
cg11777419	14	104604401	*KIF26A*	TSS1500	Island	-0.974
cg10975049	10	72432542	*ADAMTS14;ADAMTS14*	TSS200;TSS200	Island	0.974
cg24091001	2	172949870	*DLX1;DLX1*	TSS1500;TSS1500	Island	-0.973
cg18202521	1	19600702	*AKR7L;AKR7L*	TSS200;TSS200	Island	0.973
cg21578090	17	4488018	*SMTNL2;SMTNL2*	1stExon;5'UTR	Island	-0.973
cg17192454	2	74726651	*LBX2;LBX2*	1stExon;Body	Island	0.973
cg04138112	1	167090618	*DUSP27*	Body	Island	0.972
cg09247193	17	47075095	*IGF2BP1;IGF2BP1;IGF2BP1;IGF2BP1*	5'UTR;5'UTR;1stExon;1stExon	Island	-0.971

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 3. Continuation

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg09744766	16	53407421			Island	0.971
cg12034979	12	133532882	*ZNF605;ZNF605;ZNF605*	1stExon;TSS200;5' UTR	Island	0.970
cg27047283	X	135333567	*MAP7D3;MAP7D3*	1stExon;5' UTR	Island	-0.970
cg12087304	4	122873018	*TRPC3*	TSS200	Island	-0.970

[1]Identification of the CpG according to EPIC array; [2]Chromosome; [3]Δβ-values (β-value CTC-MCC-41 - β-value COLO205). All CpGs in the table are significant (FDR adjusted p-value <0.05) and are arranged according to their absolute Δβ-value.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 4. Top 20 differentially methylated CpGs from CpG islands or shore regions of gene promoters identified in CTC-MCC-41 in comparison to COLO205 cells.

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg12558358	3	23244824	*UBE2E2;UBE2E2*	1stExon;5'UTR	Island	-0.980
cg15816080	7	65670215	*TPST1*	TSS200	Island	-0.978
cg02478789	19	9649310	*ZNF426*	TSS200	Island	0.976
cg21028251	3	133614589	*RAB6B;RAB6B*	1stExon;5'UTR	Island	-0.975
cg15361065	3	55522301	*WNT5A*	TSS1500	Island	0.974
cg11777419	14	104604401	*KIF26A*	TSS1500	Island	-0.974
cg10975049	10	72432542	*ADAMTS14;ADAMTS14*	TSS200;TSS200	Island	0.974
cg24091001	2	172949870	*DLX1;DLX1*	TSS1500;TSS1500	Island	-0.973
cg18202521	1	19600702	*AKR7L;AKR7L*	TSS200;TSS200	Island	0.973
cg21578090	17	4488018	*SMTNL2;SMTNL2*	1stExon;5'UTR	Island	-0.973
cg17192454	2	74726651	*LBX2;LBX2*	1stExon;Body	Island	0.973
cg09247193	17	47075095	*IGF2BP1;IGF2BP1;IGF2BP1;IGF2BP1*	5'UTR;5'UTR;1stExon;1stExon	Island	-0.971
cg12034979	12	133532882	*ZNF605;ZNF605;ZNF605*	1stExon;TSS200;5'UTR	Island	0.970

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Suplementary Table 4. Continuation

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg27047283	X	135333567	*MAP7D3;MAP7D3*	1stExon;5'UTR	Island	-0.970
cg12087304	4	122873018	*TRPC3*	TSS200	Island	-0.970
cg16112120	3	133614580	*RAB6B;RAB6B*	1stExon;5'UTR	Island	-0.970
cg23755113	7	65670237	*TPST1*	TSS200	Island	-0.970
cg24027179	17	8534165	*MYH10*	TSS200	Island	-0.969
cg12798157	1	19600719	*AKR7L;AKR7L*	TSS200;TSS200	Island	0.969
cg20591405	3	133614578	*RAB6B;RAB6B*	5'UTR;1stExon	Island	-0.968

[1]Identification of the CpG according to EPIC array; [2]Chromosome; [3]Δβ-values (β-value CTC-MCC-41 - β-value COLO205). All CpGs in the table are significant (FDR adjusted p-value <0.05) and are arranged according to their absolute Δβ-value.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 5. Promoter hypermethylated genes obtained by genome-wide DNA methylation analysis with EPIC array in CTC-MCC-41 respect to HT29 cells selected for validation.

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg12621514	10	54073047	*DKK1*	TSS1500	Shore	0.518
cg25158147	10	54073150	*DKK1*	TSS1500	Shore	0.313
cg11988964	10	54073189	*DKK1*	TSS1500	Shore	0.377
cg02302582	10	54073281	*DKK1*	TSS1500	Shore	0.635
cg11931116	10	54074040	*DKK1;DKK1*	5'UTR;1stExon	Island	0.703
cg09445939	10	54074079	*DKK1;DKK1*	5'UTR;1stExon	Island	0.576
cg07684796	**10**	**54074209**	***DKK1***	**1stExon**	**Island**	**0.750**
cg27591349	10	54074250	*DKK1*	1stExon	Island	0.670
cg26677741	10	54074260	*DKK1;PRKG1-AS1*	1stExon;TSS1500	Island	0.439
cg02836487	3	128206457	*GATA2;GATA2;GATA2*	5'UTR;5'UTR;5'UTR	Island	0.262
cg00847029	3	128206844	*GATA2;GATA2;GATA2*	5'UTR;5'UTR;TSS200	Island	0.436
cg23520930	3	128206967	*GATA2;GATA2;GATA2*	5'UTR;5'UTR;TSS1500	Island	0.452
cg09024124	3	128207255	*GATA2;GATA2;GATA2;GATA2*	1stExon;5'UTR;5'UTR;TSS1500	Island	0.597
cg02980693	**3**	**128208970**	***GATA2***	**5'UTR**	**Island**	**0.897**
cg13808674	3	128209319	*GATA2*	5'UTR	Island	0.750
cg04101969	13	20806477	*GJB6;GJB6;GJB6;GJB6;GJB6;GJB6*	5'UTR;1stExon;1stExon;5'UTR;5'UTR;1stExon;TSS1500	Island	0.453
cg14467488	**13**	**20806663**	***GJB6;GJB6;GJB6***	**TSS200;TSS200;TSS1500**	**Island**	**0.305**

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementry Table 5. Continuation

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg03460558	X	128657473	*SMARCA1;SMARCA1*	TSS200;TSS200	Shore	0.457
cg08617833	**X**	**128657697**	***SMARCA1;SMARCA1***	**TSS1500;TSS1500**	**Shore**	**0.294**
cg26010110	**X**	**128657699**	***SMARCA1;SMARCA1***	**TSS1500;TSS1500**	**Shore**	**0.258**
cg20736997	4	157891938	*PDGFC*	1stExon	Shore	0.547
cg26777861	4	157892013	*PDGFC*	1stExon	Shore	0.796
cg09848659	4	157892087	*PDGFC;PDGFC*	5'UTR;1stExon	Shore	0.725
cg02278887	**4**	**157892159**	***PDGFC;PDGFC;PDGFC***	**1stExon;5'UTR;Body**	**Shore**	**0.753**
cg03376310	**4**	**157892187**	***PDGFC;PDGFC;PDGFC***	**1stExon;5'UTR;Body**	**Shore**	**0.758**
cg27110994	4	157892489	*PDGFC;PDGFC*	5'UTR;1stExon	Shore	0.234
cg24634420	4	157892748	*PDGFC*	TSS1500	Island	0.435
cg22451233	4	157892765	*PDGFC*	TSS1500	Island	0.220
cg13153190	4	157893231	*PDGFC*	TSS1500	Island	0.354
cg04419157	4	157893306	*PDGFC*	TSS1500	Shore	0.299
cg07537734	4	157893541	*PDGFC*	TSS1500	Shore	0.289
cg25147026	1	186650441	*PTGS2*	TSS1500	Shore	0.335
cg16101346	**1**	**186650479**	***PTGS2***	**TSS1500**	**Shore**	**0.369**
cg14598116	1	218518317	*TGFB2;TGFB2;TGFB2-AS1*	TSS1500;TSS1500;Body	Shore	0.418
cg06899755	**1**	**218520325**	***TGFB2;TGFB2***	**1stExon;1stExon**	**Island**	**0.319**

[1]Identification of the CpG according to EPIC array; [2]Chromosome; [3]Δβ-values (β-value CTC-MCC-41 - β-value HT29). All CpGs in the table are significant (FDR adjusted p-value <0.05) and are arranged according to the alphabetical order of the genes and the genomic position. CpGs selected for validation are indicated in bold.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 6. Promoter hypomethylated genes obtained by genome-wide DNA methylation analysis with EPIC array in CTC-MCC-41 respect to HT29 cells selected for validation.

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg12270822	7	87229722	*ABCB1*	5'UTR	Island	-0.300
cg16772035	7	87229880	*ABCB1*	5'UTR	Island	-0.266
cg09105881	**7**	**87230060**	***ABCB1***	**5'UTR**	**Island**	**-0.275**
cg16288248	7	87230087	*ABCB1*	5'UTR	Island	-0.227
cg04599912	7	87257030	*RUNDC3B;RUNDC3B;RUNDC3B;ABCB1*	TSS1500;TSS1500;TSS1500;5'UTR	Island	-0.350
cg04710402	7	87257074	*RUNDC3B;RUNDC3B;ABCB1;RUNDC3B*	TSS1500;TSS1500;5'UTR;TSS1500	Island	-0.283
cg00031165	7	87258360	*RUNDC3B;ABCB1;RUNDC3B;RUNDC3B*	Body;5'UTR;Body;Body	Island	-0.535
cg12040846	7	87258443	*RUNDC3B;ABCB1;RUNDC3B;RUNDC3B*	Body;5'UTR;Body;Body	Island	-0.256
cg03874092	**2**	**60780415**	***BCL11A;BCL11A;BCL11A;BCL11A;BCL11A***	**1stExon;1stExon;1stExon;5'UTR;5'UTR;5'UTR**	**Shore**	**-0.490**
cg03754165	**2**	**60780427**	***BCL11A;BCL11A;BCL11A;BCL11A;BCL11A***	**1stExon;1stExon;1stExon;5'UTR;5'UTR;5'UTR**	**Shore**	**-0.526**
cg15091323	2	60780474	*BCL11A;BCL11A;BCL11A;BCL11A;BCL11A*	1stExon;1stExon;1stExon;5'UTR;5'UTR;5'UTR	Shore	-0.266
cg26516759	**20**	**55840973**	***BMP7***	**1stExon**	**Island**	**-0.750**
cg20340302	20	55841149	*BMP7*	1stExon	Island	-0.955

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 6. Continuation

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg20292547	20	55841342	*BMP7;BMP7*	1stExon;5'UTR	Island	-0.380
cg11670519	20	55841506	*BMP7;BMP7*	1stExon;5'UTR	Island	-0.451
cg23482891	20	55841841	*BMP7*	TSS200	Shore	-0.782
cg14839404	20	55841869	*BMP7*	TSS200	Shore	-0.358
cg15069906	20	55841888	*BMP7*	TSS200	Shore	-0.596
cg03573446	20	55841893	*BMP7*	TSS200	Shore	-0.555
cg22047703	20	55841896	*BMP7*	TSS200	Shore	-0.578
cg10941747	20	55841898	*BMP7*	TSS200	Shore	-0.695
cg12518442	20	55841906	*BMP7*	TSS200	Shore	-0.730
cg00309582	20	55841925	*BMP7*	TSS1500	Shore	-0.749
cg18759209	20	55842071	*BMP7*	TSS1500	Shore	-0.457
cg14782437	12	4381731	*CCND2*	TSS1500	Island	-0.244
cg25425078	**12**	**4381777**	***CCND2***	**TSS1500**	**Island**	**-0.294**
cg15993083	**12**	**4381788**	***CCND2***	**TSS1500**	**Island**	**-0.347**
cg13575161	**12**	**4381792**	***CCND2***	**TSS1500**	**Island**	**-0.263**
cg10091994	**12**	**4381803**	***CCND2***	**TSS1500**	**Island**	**-0.498**
cg22500428	12	4381882	*CCND2*	TSS1500	Island	-0.603
cg05667158	12	4381890	*CCND2*	TSS1500	Island	-0.450
cg19582614	12	4381894	*CCND2*	TSS1500	Island	-0.301
cg08553284	12	4381997	*CCND2*	TSS1500	Island	-0.250
cg00814733	12	4382051	*CCND2*	TSS1500	Island	-0.387
cg17296482	12	4382188	*CCND2*	TSS1500	Island	-0.332
cg26910092	**2**	**216300599**	***FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1***	**5'UTR;1stExon;1stExon;5'UTR;1stExon;5'UTR;5'UTR;1stExon;1stExon;5'UTR;5'UTR;5'UTR;1stExon;1stExon**	**Island**	**-0.413**

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 6. Continuation

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg14190674	2	216300738	*FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1;FN1*	5'UTR;1stExon;1stExon;5'UTR;1stExon;5'UTR;5'UTR;1stExon;1stExon;5'UTR;5'UTR;5'UTR;1stExon;1stExon	Island	-0.299
cg22511947	2	216301558	*FN1;FN1;FN1;FN1;FN1;FN1;FN1*	TSS1500;TSS1500;TSS1500;TSS1500;TSS1500;TSS1500;TSS1500	Shore	-0.309
cg21075278	**11**	**68452087**	***GAL;GAL***	**1stExon;5'UTR**	**Island**	**-0.388**
cg00226931	**11**	**68452097**	***GAL;GAL***	**1stExon;5'UTR**	**Island**	**-0.569**
cg11582456	5	115908931	*SEMA6A*	5'UTR	Island	-0.282
cg07100508	**5**	**115909921**	***SEMA6A;SEMA6A***	**1stExon;5'UTR**	**Island**	**-0.406**

[1]Identification of the CpG according to EPIC array; [2]Chromosome; [3]Δβ-values (β-value CTC-MCC-41 - β-value HT29). All CpGs in the table are significant (FDR adjusted p-value <0.05) and are arranged according to the alphabetical order of the genes and the genomic position. CpGs selected for validation are indicated in bold.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 7. Top 20 differentially methylated CpGs identified in CTC-MCC-41 in comparison to all post-treatment CTC cell lines.

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg01448115	12	130422172			Open Sea	0.826
cg17196033	19	35168591	*ZNF302;ZNF302;ZNF302*	1stExon;5'UTR;Body	Island	0.817
cg23352597	X	46435742	*CHST7*	3'UTR	Shore	0.812
cg20506380	1	158612442	*SPTA1*	Body	Open Sea	-0.801
cg15356425	11	128822408			Open Sea	0.799
cg05147578	22	18559240	*PEX26*	TSS1500	Shore	0.796
cg14753356	6	30720108			Open Sea	0.795
cg01201472	12	12764255	*CREBL2*	TSS1500	Shore	0.789
cg22032510	X	49649233			Open Sea	-0.788
cg15979885	13	113244675			Shelf	0.787
cg23366050	18	159793	*USP14*	Body	Shore	0.786
cg06384491	X	57163888	*SPIN2A;SPIN2A*	5'UTR;1stExon	Shore	-0.780
cg10675561	1	120052814	*HSD3B1*	Body	Open Sea	0.779
cg05158109	X	40592695	*MED14*	Body	Shore	0.777
cg11646050	X	16736350	*SYAP1*	TSS1500	Shore	-0.774
cg18241219	X	16463998			Open Sea	-0.774
cg17057190	19	2747608			Open Sea	0.770
cg05056349	2	207629824	*FASTKD2;MDH1B*	TSS1500;Body	Shore	-0.765
cg02298765	19	11450741	*RAB3D*	TSS1500	Shore	0.763
cg04517512	1	218651707	*MIR548F3*	Body	Open Sea	0.763

[1]Identification of the CpG according to EPIC array; [2]Chromosome; [3]Δβ-values (β-value pre-treatment cell line CTC-MCC-41 - β-value all post-treatment cell lines). CpGs in the table are arranged according to their absolute Δβ-value.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 8. DNA methylation levels of the significantly DMCpGs of the six selected genes hypermethylated and downregulated in CTC-MCC-41 in comparison to all post-treatment CTC cell lines analyzed by EPIC array.

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg14663940	**3**	**183894397**	***AP2M1***	**5'UTR**	**Shore**	**0.406**
cg00260810	3	183894469	*AP2M1*	5'UTR	Shore	0.537
cg15914769	11	27722774	*BDNF;BDNF;BDNF;BDNF*	TSS200;TSS1500;Body;5'UTR	Shore	0.341
cg15688670	**11**	**27723190**	***BDNF;BDNF;BDNF;BDNF***	**TSS1500;Body;5'UTR;TSS200**	**Shore**	**0.413**
cg23497217	11	27723214	*BDNF;BDNF;BDNF;BDNF*	TSS1500;Body;5'UTR;TSS200	Shore	0.201
cg06816235	11	27742219	*BDNF;BDNF;BDNF;BDNF*	Body;5'UTR;TSS1500;1stExon	Island	0.368
cg16709385	20	61447581	*COL9A3*	TSS1500	Shore	-0.201
cg16461246	20	61447614	*COL9A3*	TSS1500	Shore	-0.228
cg12084392	20	61447623	*COL9A3*	TSS1500	Shore	-0.201
cg15266508	**20**	**61447742**	***COL9A3***	**TSS1500**	**Island**	**0.314**
cg16653692	20	61447749	*COL9A3*	TSS1500	Island	0.282
cg00289503	20	61470028	*COL9A3*	Body	Shore	-0.256
cg05883907	12	49318487	*FKBP11*	Body;Body;5'UTR	Island	0.430
cg23692087	12	49318557	*FKBP11*	Body;Body;5'UTR;1stExon	Island	0.224
cg22021756	12	49318740	*FKBP11*	Body;Body;TSS200	Island	0.210

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 8. Continuation

TargetID[1]	Chr[2]	Position	Gene name	Gene region	CpG context	Δβ[3]
cg19190217	12	49318749	*FKBP11*	Body;Body;TSS200	Island	0.292
cg15620146	**12**	**49318784**	***FKBP11***	**Body;Body;TSS200**	**Island**	**0.336**
cg01612057	12	49318847	*FKBP11*	Body;Body;TSS200	Island	0.335
cg23699916	12	49319263	*FKBP11*	1stExon;TSS1500;5'UTR;1stExon;5'UTR	Island	0.357
cg05999245	12	49319428	*FKBP11*	TSS200;TSS1500;TSS200	Island	0.241
cg27522205	**12**	**51784426**	***GALNT6***	**5'UTR**	**Shore**	**0.374**
cg25360385	12	51786547	*GALNT6*	TSS1500	Shore	0.233
cg12541478	12	52444020	*NR4A1*	TSS1500	Shore	0.373
cg17242930	12	52434102	*NR4A1;NR4A1*	Body;5'UTR	Shelf	0.508
cg20560869	**12**	**52447054**	***NR4A1;NR4A1***	**5'UTR;Body**	**Shore**	**0.386**

[1]Identification of the CpG according to EPIC array; [2]Chromosome; [3]Δβ-values (β-value pre-treatment cell line CTC-MCC-41 - β-value all post-treatment cell lines). All CpGs in the table are significant (FDR adjusted p-value <0.05) and are arranged according to the alphabetical order of the genes and the genomic position. CpGs selected for validation are indicated in bold.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 9. Patients inclusion and exclusion criteria of the study.

Inclusion criteria	Exclusion criteria
>18 years	Diagnosis of other tumours except basal cell carcinoma of the skin or in situ of cervix
Presence of measurable metastatic disease by computed tomography scan before initiation first-line therapy	
Eastern Oncology Cooperative Group (ECOG) performance status (PS) score equal or less than 2	
Not having received chemotherapy for metastatic disease and, in case of having received adjuvant treatment, it should have been completed within the previous 6 months	Psychiatric conditions of the patient that make it difficult to understand informed consent.
Adequate organ function and blood counts	

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 10. Top 20 differentially methylated CpGs in promoter CpGIs of baseline CTCs between a responder and a non-responder patient to FOLFOX.

Chr[1]	Position	Gene name	Gene region	CpG context	ΔMeth R vs NR (%)[2]
1	19645070	*NBL1*	Promoter	Island	80
9	35111536	*FAM214B*	Promoter	Island	-75
11	47426447	*PSMC3*	Promoter	Island	71
14	21098287	*TMEM253*	Promoter	Island	-67
11	77589763	*AQP11*	Promoter	Island	60
11	77589773	*AQP11*	Promoter	Island	60
1	19645095	*NBL1*	Promoter	Island	50
8	21788842	*GFRA2*	Promoter	Island	50
X	24054553	*EIF2S3*	Promoter	Island	50
10	22316494	*COMMD3*	Promoter	Island	50
8	22551821	*SORBS3*	Promoter	Island	-50
8	22551824	*SORBS3*	Promoter	Island	-50
8	22551836	*SORBS3*	Promoter	Island	-50
1	154328515	*ATP8B2*	Promoter	Island	36
17	63523187	*KCNH6*	Promoter	Island	33
14	21098294	*TMEM253*	Promoter	Island	33
2	218985536	*FEV*	Promoter	Island	33
14	21098302	*TMEM253*	Promoter	Island	-33
2	218985525	*FEV*	Promoter	Island	-33
5	1883741	*IRX4*	Promoter	Island	-33

[1]Chromosome; [2]ΔMethylation (Methylation of responder – Methylation of non-Responder). CpGs in the table are arranged according to their absolute ΔMethylation value.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 11. Top 20 differentially methylated CpGs in promoter CpGIs of baseline CTCs between a responder and a non-responder patient to FOLFOX and bevacizumab.

Chr[1]	Position	Gene name	Gene region	CpG context	ΔMeth R vs NR (%)[2]
14	21098287	*TMEM253*	Promoter	Island	100
1	162023717	*OLFML2B*	Promoter	Island	100
10	22316477	*COMMD3*	Promoter	Island	100
16	18790499	*RPS15A*	Promoter	Island	100
6	71420367	*LINC00472*	Promoter	Island	-80
11	62678990	*UBXN1*	Promoter	Island	-60
14	21098294	*TMEM253*	Promoter	Island	50
20	64084272	*OPRL1*	Promoter	Island	-50
20	64084283	*OPRL1*	Promoter	Island	-50
14	21098273	*TMEM253*	Promoter	Island	-50
7	30771306	*MINDY4*	Promoter	Island	-50
17	48543052	*HOXB-AS1*	Promoter	Island	-46
17	42677038	*PLEKHH3*	Promoter	Island	-40
20	64084251	*OPRL1*	Promoter	Island	-33
20	64084254	*OPRL1*	Promoter	Island	-33
6	10412412	*TFAP2A*	Promoter	Island	33
17	36937514	*LHX1-DT*	Promoter	Island	33
8	24913391	*NEFM*	Promoter	Island	-28
19	44758227	*BCL3*	Promoter	Island	25
18	50282526	*MBD1*	Promoter	Island	-25

[1]Chromosome; [2]ΔMethylation (Methylation of responder – Methylation of non-Responder). CpGs in the table are arranged according to their absolute ΔMethylation value.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Supplementary Table 12. Top 20 differentially methylated CpGs in promoter CpGIs of CTCs between baseline and the time of progression after treatment with FOLFOX and bevacizumab.

Chr[1]	Position	Gene name	Gene region	CpG context	ΔMeth R vs NR (%)[2]
10	129466708	*MGMT*	Promoter	Island	-100
7	91880885	*MTERF1*	Promoter	Island	83
16	31569482	*YBX3P1*	Promoter	Island	80
16	31569486	*YBX3P1*	Promoter	Island	80
16	86510034	*FENDRR*	Promoter	Island	80
16	86510051	*FENDRR*	Promoter	Island	80
2	138780956	*NXPH2*	Promoter	Island	80
2	138780960	*NXPH2*	Promoter	Island	80
11	65111608	*TM7SF2*	Promoter	Island	-67
8	143829879	*PUF60*	Promoter	Island	-67
7	24284816	*NPY*	Promoter	Island	60
7	91880871	*MTERF1*	Promoter	Island	60
7	91880873	*MTERF1*	Promoter	Island	60
8	69832946	*SLCO5A1*	Promoter	Island	60
10	125896739	*DHX32*	Promoter	Island	-59
20	64063270	*TCEA2*	Promoter	Island	57
20	29080270	*FRG1DP*	Promoter	Island	56
21	9325009	*LOC101930100*	Promoter	Island	52
10	129466713	*MGMT*	Promoter	Island	50
17	68600654	*FAM20A*	Promoter	Island	50

[1]Chromosome; [2]ΔMethylation (Methylation at baseline – Methylation at progression of disease). CpGs in the table are arranged according to their absolute ΔMethylation value.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

REFERENCES

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

REFERENCES

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209–249 (2021).
2. Ferlay, J. et al. (2020). Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. https://gco.iarc.fr/today, accessed [14th November 2022].
3. Hull, R., Francies, F. Z., Oyomno, M. & Dlamini, Z. Colorectal Cancer Genetics, Incidence and Risk Factors: In Search for Targeted Therapies. *Cancer Manag Res* **Volume 12**, 9869–9882 (2020).
4. Arnold, M. *et al.* Global patterns and trends in colorectal cancer incidence and mortality. *Gut* **66**, 683–691 (2017).
5. Center, M. M., Jemal, A., Smith, R. A. & Ward, E. Worldwide Variations in Colorectal Cancer. *CA Cancer J Clin* **59**, 366–378 (2009).
6. Haggar, F. & Boushey, R. Colorectal Cancer Epidemiology: Incidence, Mortality, Survival, and Risk Factors. *Clin Colon Rectal Surg* **22**, 191–197 (2009).
7. Pitchumoni, C. S. & Broder, A. Epidemiology of colorectal cancer. in *Colorectal Neoplasia and the Colorectal Microbiome* 5–33 (Elsevier, 2020). doi:10.1016/B978-0-12-819672-4.00002-7.
8. Steven A Frank. *Dynamics of Cancer: Incidence, Inheritance, and Evolution*. (2007).
9. Simon, K. Colorectal cancer development and advances in screening. *Clin Interv Aging* **11**, 967–76 (2016).
10. Meseeha, M. & Attia, M. *Colon Polyps*. (2022).
11. Yamane, L. Serrated pathway in colorectal carcinogenesis. *World J Gastroenterol* **20**, 2634 (2014).
12. Tariq, K. & Ghias, K. Colorectal cancer carcinogenesis: a review of mechanisms. *Cancer Biol Med* **13**, 120–35 (2016).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

13. Mármol, I., Sánchez-de-Diego, C., Pradilla Dieste, A., Cerrada, E. & Rodriguez Yoldi, M. Colorectal Carcinoma: A General Overview and Future Perspectives in Colorectal Cancer. *Int J Mol Sci* **18**, 197 (2017).

14. Galon, J. *et al.* Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. *J Pathol* **232**, 199–209 (2014).

15. Harada, S. & Morlote, D. Molecular Pathology of Colorectal Cancer. *Adv Anat Pathol* **27**, 20–26 (2020).

16. Binefa, G. Colorectal cancer: From prevention to personalized medicine. *World J Gastroenterol* **20**, 6786 (2014).

17. Gallois, C., Pernot, S., Zaanan, A. & Taieb, J. Colorectal Cancer: Why Does Side Matter? *Drugs* **78**, 789–798 (2018).

18. Compton, C. C. Pathology Report in Colon Cancer: What Is Prognostically Important? *Digestive Diseases* **17**, 67–79 (1999).

19. Suthananthan, A. E., Bhandari, M. & Platell, C. Influence of primary site on metastatic distribution and survival in stage IV colorectal cancer. *ANZ J Surg* **88**, 445–449 (2018).

20. Alexandrescu, S. T. *et al.* Influence of the Primary Tumor Location on the Pattern of Synchronous Metastatic Spread in Patients with Stage IV Colorectal Carcinoma, According to the 8 th Edition of the AJCC Staging System. *Journal of Gastrointestinal and Liver Diseases* **29**, 561–568 (2020).

21. Gomez, D. Anatomical distribution of colorectal cancer over a 10 year period in a district general hospital: is there a true 'rightward shift'? *Postgrad Med J* **80**, 667–669 (2004).

22. Ogino, S. & Goel, A. Molecular Classification and Correlates in Colorectal Cancer. *The Journal of Molecular Diagnostics* **10**, 13–27 (2008).

23. Sadanandam, A. *et al.* A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat Med* **19**, 619–625 (2013).

24. de Sousa E Melo, F. *et al.* Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nat Med* **19**, 614–618 (2013).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

25. Blanco-Calvo, M., Concha, Á., Figueroa, A., Garrido, F. & Valladares-Ayerbes, M. Colorectal Cancer Classification and Cell Heterogeneity: A Systems Oncology Approach. *Int J Mol Sci* **16**, 13610–13632 (2015).

26. Rashid, M., Vishwakarma, R. K., Deeb, A. M., Hussein, M. A. & Aziz, M. A. Molecular classification of colorectal cancer using the gene expression profile of tumor samples. *Exp Biol Med* **244**, 1005–1016 (2019).

27. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med* **21**, 1350–1356 (2015).

28. Menter, D. G. *et al.* Back to the Colorectal Cancer Consensus Molecular Subtype Future. *Curr Gastroenterol Rep* **21**, 5 (2019).

29. Akkoca, A. N. *et al.* TNM and Modified Dukes staging along with the demographic characteristics of patients with colorectal carcinoma. *Int J Clin Exp Med* **7**, 2828–35 (2014).

30. Li Destri, G. *et al.* Evaluation of the prognostic accuracy of Astler-Coller's and Jass' classifications of colorectal cancer. *Tumori* **87**, 127–9 (2001).

31. Amin, M. B. *et al.* The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin* **67**, 93–99 (2017).

32. Amin, M. B. et al. AJCC Cancer Staging Manual. (2017). doi:10.1001/jama.2010.1525.

33. Rosen, R. D. & Sapra, A. *TNM Classification*. (2022).

34. Troiani, T. *et al.* Molecular Aspects of Resistance to Biological and Non-Biological Drugs and Strategies to Overcome Resistance in Colorectal Cancer. *Curr Med Chem* **21**, 1639–1653 (2014).

35. van Cutsem, E., Cervantes, A., Nordlinger, B. & Arnold, D. Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* **25**, iii1–iii9 (2014).

36. Davis, L. E. The evolution of biomarkers to guide the treatment of metastatic colorectal cancer. *Am J Manag Care* **24**, S107–S117 (2018).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

37. Cervantes, A. *et al.* Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Annals of Oncology* (2022). doi:10.1016/j.annonc.2022.10.003.

38. ALLEN, W., COYLE, V. & JOHNSTON, P. Predicting the outcome of chemotherapy for colorectal cancer. *Curr Opin Pharmacol* **6**, 332–336 (2006).

39. Hernando-Cubero, J., Matos-García, I., Alonso-Orduña, V. & Capdevila, J. The Role of Fluoropirimidines in Gastrointestinal Tumours: from the Bench to the Bed. *J Gastrointest Cancer* **48**, 135–147 (2017).

40. Piawah, S. & Venook, A. P. Targeted therapy for colorectal cancer metastases: A review of current methods of molecularly targeted therapy and the use of tumor biomarkers in the treatment of metastatic colorectal cancer. *Cancer* **125**, 4139–4147 (2019).

41. Kuipers, E. J. *et al.* Colorectal cancer. *Nat Rev Dis Primers* **1**, 15065 (2015).

42. Payandeh, Z. *et al.* PD-1/PD-L1-dependent immune response in colorectal cancer. *J Cell Physiol* **235**, 5461–5475 (2020).

43. Venkatesan, S., Swanton, C., Taylor, B. S. & Costello, J. F. Treatment-Induced Mutagenesis and Selective Pressures Sculpt Cancer Evolution. *Cold Spring Harb Perspect Med* **7**, (2017).

44. Lippert, T., Ruoff, H.-J. & Volm, M. Intrinsic and Acquired Drug Resistance in Malignant Tumors. *Arzneimittelforschung* **58**, 261–264 (2011).

45. van der Jeught, K., Xu, H.-C., Li, Y.-J., Lu, X.-B. & Ji, G. Drug resistance and new therapies in colorectal cancer. *World J Gastroenterol* **24**, 3834–3848 (2018).

46. Hammond, W. A., Swaika, A. & Mody, K. Pharmacologic resistance in colorectal cancer: a review. *Ther Adv Med Oncol* **8**, 57–84 (2016).

47. Gottesman, M. M., Fojo, T. & Bates, S. E. Multidrug resistance in cancer: role of ATP-dependent transporters. *Nat Rev Cancer* **2**, 48–58 (2002).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

48. Salonga, D. *et al.* Colorectal tumors responding to 5-fluorouracil have low gene expression levels of dihydropyrimidine dehydrogenase, thymidylate synthase, and thymidine phosphorylase. *Clin Cancer Res* **6**, 1322–7 (2000).

49. Johnston, P. G. *et al.* Thymidylate synthase gene and protein expression correlate and are associated with response to 5-fluorouracil in human colorectal and gastric tumors. *Cancer Res* **55**, 1407–12 (1995).

50. Schwartz, P. M., Moir, R. D., Hyde, C. M., Turek, P. J. & Handschumacher, R. E. Role of uridine phosphorylase in the anabolism of 5-fluorouracil. *Biochem Pharmacol* **34**, 3585–9 (1985).

51. Houghton, J. A. & Houghton, P. J. Elucidation of pathways of 5-fluorouracil metabolism in xenografts of human colorectal adenocarcinoma. *Eur J Cancer Clin Oncol* **19**, 807–15 (1983).

52. Xu, Y. & Villalona-Calero, M. A. Irinotecan: mechanisms of tumor resistance and novel strategies for modulating its activity. *Ann Oncol* **13**, 1841–51 (2002).

53. Palshof, J. A. *et al.* Topoisomerase I copy number alterations as biomarker for irinotecan efficacy in metastatic colorectal cancer. *BMC Cancer* **17**, 48 (2017).

54. Makondi, P. T., Chu, C.-M., Wei, P.-L. & Chang, Y.-J. Prediction of novel target genes and pathways involved in irinotecan-resistant colorectal cancer. *PLoS One* **12**, e0180616 (2017).

55. Gnoni, A. *et al.* Pharmacokinetic and metabolism determinants of fluoropyrimidines and oxaliplatin activity in treatment of colorectal patients. *Curr Drug Metab* **12**, 918–31 (2011).

56. Yan, D. *et al.* WBSCR22 confers oxaliplatin resistance in human colorectal cancer. *Sci Rep* **7**, 15443 (2017).

57. Mao, L. *et al.* Transforming growth factor-β1 contributes to oxaliplatin resistance in colorectal cancer via epithelial to mesenchymal transition. *Oncol Lett* **14**, 647–654 (2017).

58. Tejpar, S., Prenen, H. & Mazzone, M. Overcoming resistance to antiangiogenic therapies. *Oncologist* **17**, 1039–50 (2012).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

59. Longley, D. & Johnston, P. Molecular mechanisms of drug resistance. *J Pathol* **205**, 275–292 (2005).

60. Pineda, E. *et al.* Dynamic soluble changes in sVEGFR1, HGF, and VEGF promote chemotherapy and bevacizumab resistance: A prospective translational study in the BECOX (GEMCAD 09-01) trial. *Tumour Biol* **39**, 1010428317705509 (2017).

61. Carbone, C. *et al.* Homeobox B9 Mediates Resistance to Anti-VEGF Therapy in Colorectal Cancer Patients. *Clin Cancer Res* **23**, 4312–4322 (2017).

62. Davies, J. M. & Goldberg, R. M. Treatment of Metastatic Colorectal Cancer. *Semin Oncol* **38**, 552–560 (2011).

63. Peeters, M. *et al.* Analysis of KRAS/NRAS Mutations in a Phase III Study of Panitumumab with FOLFIRI Compared with FOLFIRI Alone as Second-line Treatment for Metastatic Colorectal Cancer. *Clin Cancer Res* **21**, 5469–79 (2015).

64. Kim, T. W. *et al.* A phase 3 trial evaluating panitumumab plus best supportive care vs best supportive care in chemorefractory wild-type KRAS or RAS metastatic colorectal cancer. *Br J Cancer* **115**, 1206–1214 (2016).

65. de Roock, W. *et al.* Association of KRAS p.G13D mutation with outcome in patients with chemotherapy-refractory metastatic colorectal cancer treated with cetuximab. *JAMA* **304**, 1812–20 (2010).

66. Karapetis, C. S. *et al.* PIK3CA, BRAF, and PTEN status and benefit from cetuximab in the treatment of advanced colorectal cancer--results from NCIC CTG/AGITG CO.17. *Clin Cancer Res* **20**, 744–53 (2014).

67. De Roock, W. *et al.* Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol* **11**, 753–62 (2010).

68. Chiang, A. C. & Massagué, J. Molecular Basis of Metastasis. *New England Journal of Medicine* **359**, 2814–2823 (2008).

69. Gupta, G. P. & Massagué, J. Cancer Metastasis: Building a Framework. *Cell* **127**, 679–695 (2006).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

70. Steeg, P. S. Tumor metastasis: mechanistic insights and clinical challenges. *Nat Med* **12**, 895–904 (2006).

71. Nandy, S. B. & Lakshmanaswamy, R. Cancer Stem Cells and Metastasis. in 137–176 (2017). doi:10.1016/bs.pmbts.2017.07.007.

72. Valastyan, S. & Weinberg, R. A. Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell* **147**, 275–292 (2011).

73. Friedl, P. & Wolf, K. Tumour-cell invasion and migration: diversity and escape mechanisms. *Nat Rev Cancer* **3**, 362–374 (2003).

74. Liu, Q.-L., Luo, M., Huang, C., Chen, H.-N. & Zhou, Z.-G. Epigenetic Regulation of Epithelial to Mesenchymal Transition in the Cancer Metastatic Cascade: Implications for Cancer Therapy. *Front Oncol* **11**, (2021).

75. Dmello, R. S., To, S. Q. & Chand, A. L. Therapeutic Targeting of the Tumour Microenvironment in Metastatic Colorectal Cancer. *Int J Mol Sci* **22**, 2067 (2021).

76. Sonoshita, M. *et al.* Suppression of Colon Cancer Metastasis by Aes through Inhibition of Notch Signaling. *Cancer Cell* **19**, 125–137 (2011).

77. Zhang, X. *et al.* Human colorectal cancer-derived mesenchymal stem cells promote colorectal cancer progression through IL-6/JAK2/STAT3 signaling. *Cell Death Dis* **9**, 25 (2018).

78. Li, S. *et al.* Shear stress promotes anoikis resistance of cancer cells via caveolin-1-dependent extrinsic and intrinsic apoptotic pathways. *J Cell Physiol* **234**, 3730–3743 (2019).

79. Adeshakin, F. O. *et al.* Mechanisms for Modulating Anoikis Resistance in Cancer and the Relevance of Metabolic Reprogramming. *Front Oncol* **11**, (2021).

80. Joyce, J. A. & Pollard, J. W. Microenvironmental regulation of metastasis. *Nat Rev Cancer* **9**, 239–252 (2009).

81. Welch, D. R. & Hurst, D. R. Defining the Hallmarks of Metastasis. *Cancer Res* **79**, 3011–3027 (2019).

82. Dmello, R. S., To, S. Q. & Chand, A. L. Therapeutic Targeting of the Tumour Microenvironment in Metastatic Colorectal Cancer. *Int J Mol Sci* **22**, 2067 (2021).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

83. Psaila, B. & Lyden, D. The metastatic niche: adapting the foreign soil. *Nat Rev Cancer* **9**, 285–293 (2009).

84. Liu, Y. & Cao, X. Characteristics and Significance of the Pre-metastatic Niche. *Cancer Cell* **30**, 668–681 (2016).

85. Chambers, A. F., Groom, A. C. & MacDonald, I. C. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* **2**, 563–572 (2002).

86. Kilgour, E., Rothwell, D. G., Brady, G. & Dive, C. Liquid Biopsy-Based Biomarkers of Treatment Response and Resistance. *Cancer Cell* **37**, 485–495 (2020).

87. Poulet, G., Massias, J. & Taly, V. Liquid Biopsy: General Concepts. *Acta Cytol* **63**, 449–455 (2019).

88. de Rubis, G., Rajeev Krishnan, S. & Bebawy, M. Liquid Biopsies in Cancer Diagnosis, Monitoring, and Prognosis. *Trends Pharmacol Sci* **40**, 172–186 (2019).

89. von Bubnoff, N. Liquid Biopsy: Approaches to Dynamic Genotyping in Cancer. *Oncol Res Treat* **40**, 409–416 (2017).

90. Pantel, K. & Alix-Panabières, C. Circulating tumour cells and cell-free DNA in gastrointestinal cancer. *Nat Rev Gastroenterol Hepatol* **14**, 73–74 (2017).

91. Ashworth, T. A case of cancer in which cells similar to those in the tumors were seen in blood after death. Aust Med J. 14, 146–147 (1869).

92. Castro-Giner, F. & Aceto, N. Tracking cancer progression: from circulating tumor cells to metastasis. *Genome Med* **12**, 31 (2020).

93. Palmirotta, R. *et al.* Liquid biopsy of cancer: a multimodal diagnostic tool in clinical oncology. *Ther Adv Med Oncol* **10**, 175883591879463 (2018).

94. Kim, M.-Y. *et al.* Tumor Self-Seeding by Circulating Cancer Cells. *Cell* **139**, 1315–1326 (2009).

95. Parkinson, D. R. *et al.* Considerations in the development of circulating tumor cell technology for clinical use. *J Transl Med* **10**, 138 (2012).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

96. Diamantopoulou, Z., Castro-Giner, F. & Aceto, N. Circulating tumor cells: Ready for translation? *Journal of Experimental Medicine* **217**, (2020).

97. Aceto, N. *et al.* Circulating Tumor Cell Clusters Are Oligoclonal Precursors of Breast Cancer Metastasis. *Cell* **158**, 1110–1122 (2014).

98. Mader, S. & Pantel, K. Liquid Biopsy: Current Status and Future Perspectives. *Oncol Res Treat* **40**, 404–408 (2017).

99. Alix-Panabières, C. & Pantel, K. Circulating Tumor Cells: Liquid Biopsy of Cancer. *Clin Chem* **59**, 110–118 (2013).

100. Alix-Panabières, C. & Pantel, K. Clinical Applications of Circulating Tumor Cells and Circulating Tumor DNA as Liquid Biopsy. *Cancer Discov* **6**, 479–491 (2016).

101. Yap, T. A., Lorente, D., Omlin, A., Olmos, D. & de Bono, J. S. Circulating Tumor Cells: A Multifunctional Biomarker. *Clinical Cancer Research* **20**, 2553–2568 (2014).

102. Lowes, L. E. & Allan, A. L. Circulating Tumor Cells and Implications of the Epithelial-to-Mesenchymal Transition. *Adv Clin Chem* **83**, 121–181 (2018).

103. Toss, A., Mu, Z., Fernandez, S. & Cristofanilli, M. CTC enumeration and characterization: moving toward personalized medicine. *Ann Transl Med* **2**, 108 (2014).

104. Nagrath, S. *et al.* Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* **450**, 1235–1239 (2007).

105. Miller, M. C., Robinson, P. S., Wagner, C. & O'Shannessy, D. J. The Parsortix™ Cell Separation System-A versatile liquid biopsy platform. *Cytometry A* **93**, 1234–1239 (2018).

106. Dhar, M. *et al.* High efficiency vortex trapping of circulating tumor cells. *Biomicrofluidics* **9**, 064116 (2015).

107. di Trapani, M., Manaresi, N. & Medoro, G. DEPArray™ system: An automatic image-based sorter for isolation of pure circulating tumor cells. *Cytometry A* **93**, 1260–1266 (2018).

108. Chinen, L. T. D. *et al.* Circulating Tumor Cells as Cancer Biomarkers in the Clinic. in 1–41 (2017). doi:10.1007/978-3-319-55947-6_1.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

109. Abalde-Cela, S., Piairo, P. & Diéguez, L. The Significance of Circulating Tumour Cells in the Clinic. *Acta Cytol* **63**, 466–478 (2019).

110. Cristofanilli, M. *et al.* Circulating Tumor Cells, Disease Progression, and Survival in Metastatic Breast Cancer. *New England Journal of Medicine* **351**, 781–791 (2004).

111. de Bono, J. S. *et al.* Circulating Tumor Cells Predict Survival Benefit from Treatment in Metastatic Castration-Resistant Prostate Cancer. *Clinical Cancer Research* **14**, 6302–6309 (2008).

112. Cohen, S. J. *et al.* Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. *J Clin Oncol* **26**, 3213–21 (2008).

113. Lindsay, C. R. *et al.* EPAC-lung: pooled analysis of circulating tumour cells in advanced non-small cell lung cancer. *Eur J Cancer* **117**, 60–68 (2019).

114. Foy, V., Fernandez-Gutierrez, F., Faivre-Finn, C., Dive, C. & Blackhall, F. The clinical utility of circulating tumour cells in patients with small cell lung cancer. *Transl Lung Cancer Res* **6**, 409–417 (2017).

115. Bidard, F. C. *et al.* Circulating tumor cells in locally advanced pancreatic adenocarcinoma: the ancillary CirCe 07 study to the LAP 07 trial. *Annals of Oncology* **24**, 2057–2061 (2013).

116. Vasseur, A., Kiavue, N., Bidard, F., Pierga, J. & Cabel, L. Clinical utility of circulating tumor cells: an update. *Mol Oncol* **15**, 1647–1666 (2021).

117. Wang, C. *et al.* Longitudinally collected CTCs and CTC-clusters and clinical outcomes of metastatic breast cancer. *Breast Cancer Res Treat* **161**, 83–94 (2017).

118. Murlidhar, V. *et al.* Poor Prognosis Indicated by Venous Circulating Tumor Cell Clusters in Early-Stage Lung Cancers. *Cancer Res* **77**, 5194–5206 (2017).

119. Gkountela, S. *et al.* Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding. *Cell* **176**, 98-112.e14 (2019).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

120. Lowes, L. E. & Allan, A. L. Circulating Tumor Cells and Implications of the Epithelial-to-Mesenchymal Transition. in 121–181 (2018). doi:10.1016/bs.acc.2017.10.004.

121. Wei, T. *et al.* Vimentin-positive circulating tumor cells as a biomarker for diagnosis and treatment monitoring in patients with pancreatic cancer. *Cancer Lett* **452**, 237–243 (2019).

122. Krol, I. *et al.* Detection of clustered circulating tumour cells in early breast cancer. *Br J Cancer* **125**, 23–27 (2021).

123. Hou, J.-M. *et al.* Clinical Significance and Molecular Characteristics of Circulating Tumor Cells and Circulating Tumor Microemboli in Patients With Small-Cell Lung Cancer. *Journal of Clinical Oncology* **30**, 525–532 (2012).

124. Goldkorn, A. *et al.* Circulating Tumor Cell Counts Are Prognostic of Overall Survival in SWOG S0421: A Phase III Trial of Docetaxel With or Without Atrasentan for Metastatic Castration-Resistant Prostate Cancer. *Journal of Clinical Oncology* **32**, 1136–1142 (2014).

125. Pearl, M. L. *et al.* Treatment monitoring of patients with epithelial ovarian cancer using invasive circulating tumor cells (iCTCs). *Gynecol Oncol* **137**, 229–238 (2015).

126. Wang, C. *et al.* Prognostic value of HER2 status on circulating tumor cells in advanced-stage breast cancer patients with HER2-negative tumors. *Breast Cancer Res Treat* **181**, 679–689 (2020).

127. Papadaki, M. A. *et al.* Clinical Relevance of Immune Checkpoints on Circulating Tumor Cells in Breast Cancer. *Cancers (Basel)* **12**, 376 (2020).

128. Lin, D. *et al.* Circulating tumor cells: biology and clinical significance. *Signal Transduct Target Ther* **6**, 404 (2021).

129. Hinz, S. *et al.* Detection of circulating tumor cells with CK20 RT-PCR is an independent negative prognostic marker in colon cancer patients - a prospective study. *BMC Cancer* **17**, 53 (2017).

130. Hoshimoto, S. *et al.* Association Between Circulating Tumor Cells and Prognosis in Patients With Stage III Melanoma With Sentinel Lymph Node Metastasis in a Phase III International Multicenter Trial. *Journal of Clinical Oncology* **30**, 3819–3826 (2012).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

131. Carter, L. *et al.* Molecular analysis of circulating tumor cells identifies distinct copy-number profiles in patients with chemosensitive and chemorefractory small-cell lung cancer. *Nat Med* **23**, 114–119 (2017).

132. El-Heliebi, A. *et al.* In Situ Detection and Quantification of AR-V7, AR-FL, PSA, and KRAS Point Mutations in Circulating Tumor Cells. *Clin Chem* **64**, 536–546 (2018).

133. Nicolazzo, C. *et al.* Baseline CD44v6-positive circulating tumor cells to predict first-line treatment failure in patients with metastatic colorectal cancer. *Oncotarget* **11**, 4115–4122 (2020).

134. Aranda, E. *et al.* FOLFOXIRI plus bevacizumab versus FOLFOX plus bevacizumab for patients with metastatic colorectal cancer and ≥3 circulating tumour cells: the randomised phase III VISNÚ-1 trial. *ESMO Open* **5**, e000944 (2020).

135. Pantel, K. & Alix-Panabières, C. Real-time Liquid Biopsy in Cancer Patients: Fact or Fiction? *Cancer Res* **73**, 6384–6388 (2013).

136. Katt, M. E., Placone, A. L., Wong, A. D., Xu, Z. S. & Searson, P. C. In Vitro Tumor Models: Advantages, Disadvantages, Variables, and Selecting the Right Platform. *Front Bioeng Biotechnol* **4**, (2016).

137. Diaz-Lagares, A. *et al.* Epigenetic inactivation of the p53-induced long noncoding RNA TP53 target 1 in human cancer. *Proc Natl Acad Sci U S A* **113**, E7535–E7544 (2016).

138. Guo, T. Culture of Circulating Tumor Cells - Holy Grail and Big Challenge. *Int J Cancer Clin Res* **3**, (2016).

139. Soler, A. *et al.* Autologous cell lines from circulating colon cancer cells captured from sequential liquid biopsies as model to study therapy-driven tumor changes. *Sci Rep* **8**, 15931 (2018).

140. Cayrefourcq, L. *et al.* Establishment and characterization of a cell line from human circulating colon cancer cells. *Cancer Res* **75**, 892–901 (2015).

141. Berger, S. L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. An operational definition of epigenetics. *Genes Dev* **23**, 781–3 (2009).

142. Tammen, S. A., Friso, S. & Choi, S.-W. Epigenetics: The link between nature and nurture. *Mol Aspects Med* **34**, 753–764 (2013).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

143. Kanwal, R., Gupta, K. & Gupta, S. Cancer epigenetics: an introduction. *Methods Mol Biol* **1238**, 3–25 (2015).

144. Skinner, M. K., Manikkam, M. & Guerrero-Bosagna, C. Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends in Endocrinology & Metabolism* **21**, 214–222 (2010).

145. Portela, A. & Esteller, M. Epigenetic modifications and human disease. *Nat Biotechnol* **28**, 1057–1068 (2010).

146. Kronfol, M. M., Dozmorov, M. G., Huang, R., Slattum, P. W. & McClay, J. L. The role of epigenomics in personalized medicine. *Expert Rev Precis Med Drug Dev* **2**, 33–45 (2017).

147. Dupont, C., Armant, D. & Brenner, C. Epigenetics: Definition, Mechanisms and Clinical Perspective. *Semin Reprod Med* **27**, 351–357 (2009).

148. Reik, W., Dean, W. & Walter, J. Epigenetic Reprogramming in Mammalian Development. *Science (1979)* **293**, 1089–1093 (2001).

149. Egger, G., Liang, G., Aparicio, A. & Jones, P. A. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429**, 457–463 (2004).

150. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).

151. Qu, Y. *et al.* Differential methylation in CN-AML preferentially targets non-CGI regions and is dictated by DNMT3A mutational status and associated with predominant hypomethylation of HOX genes. *Epigenetics* **9**, 1108–19 (2014).

152. Dawson, M. A. & Kouzarides, T. Cancer Epigenetics: From Mechanism to Therapy. *Cell* **150**, 12–27 (2012).

153. Zhao, S. G. *et al.* The DNA methylation landscape of advanced prostate cancer. *Nat Genet* **52**, 778–789 (2020).

154. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat Rev Genet* **14**, 204–220 (2013).

155. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–5 (2009).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

156. Xu, T. & Gao, H. Hydroxymethylation and tumors: can 5-hydroxymethylation be used as a marker for tumor diagnosis and treatment? *Hum Genomics* **14**, 15 (2020).

157. Allfrey, V. G., Faulkner, R. & Mirsky, A. E. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proceedings of the National Academy of Sciences* **51**, 786–794 (1964).

158. Putiri, E. L. & Robertson, K. D. Epigenetic mechanisms and genome stability. *Clin Epigenetics* **2**, 299–314 (2011).

159. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128**, 693–705 (2007).

160. Weber, C. M. & Henikoff, S. Histone variants: dynamic punctuation in transcription. *Genes Dev* **28**, 672–682 (2014).

161. Bai, L. & Morozov, A. v. Gene regulation by nucleosome positioning. *Trends in Genetics* **26**, 476–483 (2010).

162. Ribeiro-Silva, C., Vermeulen, W. & Lans, H. SWI/SNF: Complex complexes in genome stability and cancer. *DNA Repair (Amst)* **77**, 87–95 (2019).

163. Panni, S., Lovering, R. C., Porras, P. & Orchard, S. Non-coding RNA regulatory networks. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1863**, 194417 (2020).

164. Gibney, E. R. & Nolan, C. M. Epigenetics and gene expression. *Heredity (Edinb)* **105**, 4–13 (2010).

165. Hombach, S. & Kretz, M. Non-coding RNAs: Classification, Biology and Functioning. in 3–17 (2016). doi:10.1007/978-3-319-42059-2_1.

166. Bure, I. v., Nemtsova, M. v. & Kuznetsova, E. B. Histone Modifications and Non-Coding RNAs: Mutual Epigenetic Regulation and Role in Pathogenesis. *Int J Mol Sci* **23**, 5801 (2022).

167. Golbabapour, S., Abdulla, M. A. & Hajrezaei, M. A Concise Review on Epigenetic Regulation: Insight into Molecular Mechanisms. *Int J Mol Sci* **12**, 8661–8694 (2011).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

168. McCabe, M. T., Brandes, J. C. & Vertino, P. M. Cancer DNA methylation: molecular mechanisms and clinical implications. *Clin Cancer Res* **15**, 3927–37 (2009).

169. Shin, S. *et al.* MicroRNAs that respond to histone deacetylase inhibitor SAHA and p53 in HCT116 human colon carcinoma cells. *Int J Oncol* **35**, 1343–52 (2009).

170. Léveillé, N. *et al.* Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a lncRNA. *Nat Commun* **6**, 6520 (2015).

171. Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–6 (2010).

172. Watanabe, Y. & Maekawa, M. Methylation of DNA in Cancer. in 145–167 (2010). doi:10.1016/S0065-2423(10)52006-7.

173. Klutstein, M., Nejman, D., Greenfield, R. & Cedar, H. DNA Methylation in Cancer and Aging. *Cancer Res* **76**, 3446–3450 (2016).

174. Tse, J. W. T., Jenkins, L. J., Chionh, F. & Mariadason, J. M. Aberrant DNA Methylation in Colorectal Cancer: What Should We Target? *Trends Cancer* **3**, 698–712 (2017).

175. Wild, L. & Flanagan, J. M. Genome-wide hypomethylation in cancer may be a passive consequence of transformation. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1806**, 50–57 (2010).

176. Feinberg, A. P. & Vogelstein, B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301**, 89–92 (1983).

177. Feinberg, A. P. & Vogelstein, B. Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun* **111**, 47–54 (1983).

178. Feinberg, A. P. & Tycko, B. The history of cancer epigenetics. *Nat Rev Cancer* **4**, 143–53 (2004).

179. Eden, A., Gaudet, F., Waghmare, A. & Jaenisch, R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* **300**, 455 (2003).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

180. Kulis, M. & Esteller, M. DNA methylation and cancer. *Adv Genet* **70**, 27–56 (2010).

181. Kondo, Y. & Issa, J.-P. J. DNA methylation profiling in cancer. *Expert Rev Mol Med* **12**, e23 (2010).

182. Rainier, S. *et al.* Relaxation of imprinted genes in human cancer. *Nature* **362**, 747–9 (1993).

183. Frühwald, M. C. DNA Methylation Patterns in Cancer. *American Journal of PharmacoGenomics* **3**, 245–260 (2003).

184. Heyn, H. & Esteller, M. DNA methylation profiling in the clinic: applications and challenges. *Nat Rev Genet* **13**, 679–692 (2012).

185. Costa-Pinheiro, P., Montezuma, D., Henrique, R. & Jerónimo, C. Diagnostic and prognostic epigenetic biomarkers in cancer. *Epigenomics* **7**, 1003–1015 (2015).

186. Pan, Y., Liu, G., Zhou, F., Su, B. & Li, Y. DNA methylation profiles in cancer diagnosis and therapeutics. *Clin Exp Med* **18**, 1–14 (2018).

187. Chen, W.-D. *et al.* Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene. *J Natl Cancer Inst* **97**, 1124–32 (2005).

188. Grützmann, R. *et al.* Sensitive Detection of Colorectal Cancer in Peripheral Blood by Septin 9 DNA Methylation Assay. *PLoS One* **3**, e3759 (2008).

189. Imperiale, T. F. *et al.* Multitarget stool DNA testing for colorectal-cancer screening. *N Engl J Med* **370**, 1287–97 (2014).

190. Nilsson, T. K., Löf-Öhlin, Z. M. & Sun, X.-F. DNA methylation of the p14ARF, RASSF1A and APC1A genes as an independent prognostic factor in colorectal cancer patients. *Int J Oncol* **42**, 127–33 (2013).

191. Fukushige, S. & Horii, A. DNA Methylation in Cancer: A Gene Silencing Mechanism and the Clinical Potential of Its Biomarkers. *Tohoku J Exp Med* **229**, 173–185 (2013).

192. Miranda Furtado, C. L. *et al.* Epidrugs: targeting epigenetic marks in cancer treatment. *Epigenetics* **14**, 1164–1176 (2019).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

193. Hunter, P. The second coming of epigenetic drugs: a more strategic and broader research framework could boost the development of new drugs to modify epigenetic factors and gene expression. *EMBO Rep* **16**, 276–9 (2015).

194. Ganesan, A., Arimondo, P. B., Rots, M. G., Jeronimo, C. & Berdasco, M. The timeline of epigenetic drug discovery: from reality to dreams. *Clin Epigenetics* **11**, 174 (2019).

195. Majchrzak-Celińska, A., Warych, A. & Szoszkiewicz, M. Novel Approaches to Epigenetic Therapies: From Drug Combinations to Epigenetic Editing. *Genes (Basel)* **12**, 208 (2021).

196. Raynal, N. J. *et al.* Abstract 380: Discovery of new epigenetic drugs among FDA-approved drug libraries. *Cancer Res* **74**, 380–380 (2014).

197. Jung, G., Hernández-Illán, E., Moreira, L., Balaguer, F. & Goel, A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. *Nat Rev Gastroenterol Hepatol* **17**, 111–130 (2020).

198. Lyberopoulou, A. *et al.* Identification of Methylation Profiles of Cancer-related Genes in Circulating Tumor Cells Population. *Anticancer Res* **37**, 1105–1112 (2017).

199. Chimonidou, M. *et al.* DNA methylation of tumor suppressor and metastasis suppressor genes in circulating tumor cells. *Clin Chem* **57**, 1169–77 (2011).

200. Rodriguez-Casanova, A. *et al.* Epigenetic Landscape of Liquid Biopsy in Colorectal Cancer. *Front Cell Dev Biol* **9**, 622459 (2021).

201. Shirahata, A. *et al.* Vimentin methylation as a marker for advanced colorectal carcinoma. *Anticancer Res* **29**, 279–81 (2009).

202. Loboda, A. *et al.* EMT is the dominant program in human colon cancer. *BMC Med Genomics* **4**, 9 (2011).

203. Pixberg, C. F. *et al.* Analysis of DNA methylation in single circulating tumor cells. *Oncogene* **36**, 3223–3231 (2017).

204. Barbazán, J. *et al.* Molecular characterization of circulating tumor cells in human metastatic colorectal cancer. *PLoS One* **7**, e40476 (2012).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

205. Barbazán, J. *et al.* Liver Metastasis Is Facilitated by the Adherence of Circulating Tumor Cells to Vascular Fibronectin Deposits. *Cancer Res* **77**, 3431–3441 (2017).

206. Eslami-S, Z., Cortés-Hernández, L. E., Thomas, F., Pantel, K. & Alix-Panabières, C. Functional analysis of circulating tumour cells: the KEY to understand the biology of the metastatic cascade. *Br J Cancer* **127**, 800–810 (2022).

207. Dujon, A. M. *et al.* Is There One Key Step in the Metastatic Cascade? *Cancers (Basel)* **13**, (2021).

208. Koch, C. *et al.* Characterization of circulating breast cancer cells with tumorigenic and metastatic capacity. *EMBO Mol Med* **12**, e11908 (2020).

209. Alix-Panabières, C. *et al.* Molecular Portrait of Metastasis-Competent Circulating Tumor Cells in Colon Cancer Reveals the Crucial Role of Genes Regulating Energy Metabolism and DNA Repair. *Clin Chem* **63**, 700–713 (2017).

210. Carmona, F. J. *et al.* Epigenetic disruption of cadherin-11 in human cancer metastasis. *J Pathol* **228**, 230–40 (2012).

211. Shinozaki, M. *et al.* Distinct hypermethylation profile of primary breast cancer is associated with sentinel lymph node metastasis. *Clin Cancer Res* **11**, 2156–62 (2005).

212. Pixberg, C. F., Schulz, W. A., Stoecklein, N. H. & Neves, R. P. L. Characterization of DNA Methylation in Circulating Tumor Cells. *Genes (Basel)* **6**, 1053–75 (2015).

213. Galm, O. & Herman, J. G. Methylation-specific polymerase chain reaction. *Methods Mol Med* **113**, 279–91 (2005).

214. Moran, S., Arribas, C. & Esteller, M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* **8**, 389–99 (2016).

215. Tabas-Madrid, D., Nogales-Cadenas, R. & Pascual-Montano, A. GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res* **40**, W478-83 (2012).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

216. Harrington, C. T., Lin, E. I., Olson, M. T. & Eshleman, J. R. Fundamentals of pyrosequencing. *Arch Pathol Lab Med* **137**, 1296–303 (2013).

217. Untergasser, A. *et al.* Primer3--new capabilities and interfaces. *Nucleic Acids Res* **40**, e115 (2012).

218. Cottrell, S. *et al.* Discovery and validation of 3 novel DNA methylation markers of prostate cancer prognosis. *J Urol* **177**, 1753–8 (2007).

219. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–8 (2001).

220. Baylin, S. B. & Chen, W. Y. Aberrant gene silencing in tumor progression: implications for control of cancer. *Cold Spring Harb Symp Quant Biol* **70**, 427–33 (2005).

221. Husni, R. E. *et al.* DNA hypomethylation-related overexpression of SFN, GORASP2 and ZYG11A is a novel prognostic biomarker for early stage lung adenocarcinoma. *Oncotarget* **10**, 1625–1636 (2019).

222. Darwiche, N. Epigenetic mechanisms and the hallmarks of cancer: an intimate affair. *Am J Cancer Res* **10**, 1954–1978 (2020).

223. Lao, V. V. & Grady, W. M. Epigenetics and colorectal cancer. *Nat Rev Gastroenterol Hepatol* **8**, 686–700 (2011).

224. Chaffer, C. L. & Weinberg, R. A. A perspective on cancer cell metastasis. *Science* **331**, 1559–64 (2011).

225. Gutierrez, A., Demond, H., Brebi, P. & Ili, C. G. Novel Methylation Biomarkers for Colorectal Cancer Prognosis. *Biomolecules* **11**, (2021).

226. Scherer, W. F., Syverton, J. T. & Gey, G. O. Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. *J Exp Med* **97**, 695–710 (1953).

227. Ericson, K. *et al.* Genetic inactivation of AKT1, AKT2, and PDPK1 in human colorectal cancer cells clarifies their roles in tumor growth regulation. *Proc Natl Acad Sci U S A* **107**, 2598–603 (2010).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

228. Ruiz-Bañobre, J. *et al.* Noninvasive early detection of colorectal cancer by hypermethylation of the LINC00473 promoter in plasma cell-free DNA. *Clin Epigenetics* **14**, 86 (2022).

229. Mirabelli, P., Coppola, L. & Salvatore, M. Cancer Cell Lines Are Useful Model Systems for Medical Research. *Cancers (Basel)* **11**, (2019).

230. Esteller, M. Epigenetics in cancer. *N Engl J Med* **358**, 1148–59 (2008).

231. Sheaffer, K. L., Elliott, E. N. & Kaestner, K. H. DNA Hypomethylation Contributes to Genomic Instability and Intestinal Cancer Initiation. *Cancer Prev Res (Phila)* **9**, 534–46 (2016).

232. Hanssen, A. *et al.* Characterization of different CTC subpopulations in non-small cell lung cancer. *Sci Rep* **6**, 28010 (2016).

233. Brown, L. C. *et al.* Circulating Tumor Cell Chromosomal Instability and Neuroendocrine Phenotype by Immunomorphology and Poor Outcomes in Men with mCRPC Treated with Abiraterone or Enzalutamide. *Clin Cancer Res* **27**, 4077–4088 (2021).

234. Bell, R. E. *et al.* Enhancer methylation dynamics contribute to cancer plasticity and patient mortality. *Genome Res* **26**, 601–611 (2016).

235. Irizarry, R. A. *et al.* The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* **41**, 178–186 (2009).

236. de Sousa E Melo, F. & Vermeulen, L. Wnt Signaling in Cancer Stem Cell Biology. *Cancers (Basel)* **8**, (2016).

237. Tan, X. *et al.* Hypermethylation of the PTTG1IP promoter leads to low expression in early-stage non-small cell lung cancer. *Oncol Lett* **18**, 1278–1286 (2019).

238. Qi, L. *et al.* Dickkopf-1 inhibits epithelial-mesenchymal transition of colon cancer cells and contributes to colon cancer suppression. *Cancer Sci* **103**, 828–35 (2012).

239. Cowling, V. H. & Cole, M. D. Turning the tables: Myc activates Wnt in breast cancer. *Cell Cycle* **6**, 2625–7 (2007).

240. Götze, S., Wolter, M., Reifenberger, G., Müller, O. & Sievers, S. Frequent promoter hypermethylation of Wnt pathway inhibitor genes in malignant astrocytic gliomas. *Int J Cancer* **126**, 2584–93 (2010).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

241. Aguilera, O. *et al.* Epigenetic inactivation of the Wnt antagonist DICKKOPF-1 (DKK-1) gene in human colorectal cancer. *Oncogene* **25**, 4116–21 (2006).

242. Kopetz, S. *et al.* Phase II trial of infusional fluorouracil, irinotecan, and bevacizumab for metastatic colorectal cancer: efficacy and circulating angiogenic biomarkers associated with therapeutic resistance. *J Clin Oncol* **28**, 453–9 (2010).

243. Valencia, A. *et al.* Wnt signaling pathway is epigenetically regulated by methylation of Wnt antagonists in acute myeloid leukemia. *Leukemia* **23**, 1658–66 (2009).

244. Huang, Y. *et al.* Overexpression of Dickkopf-1 predicts poor prognosis for patients with hepatocellular carcinoma after orthotopic liver transplantation by promoting cancer metastasis and recurrence. *Med Oncol* **31**, 966 (2014).

245. Nusse, R. & Clevers, H. Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985–999 (2017).

246. Caspi, M., Wittenstein, A., Kazelnik, M., Shor-Nareznoy, Y. & Rosin-Arbesfeld, R. Therapeutic targeting of the oncogenic Wnt signaling pathway for treating colorectal cancer and other colonic disorders. *Adv Drug Deliv Rev* **169**, 118–136 (2021).

247. Davis, S. L. *et al.* A phase 1b dose escalation study of Wnt pathway inhibitor vantictumab in combination with nab-paclitaxel and gemcitabine in patients with previously untreated metastatic pancreatic cancer. *Invest New Drugs* **38**, 821–830 (2020).

248. Santiago, L., Daniels, G., Wang, D., Deng, F.-M. & Lee, P. Wnt signaling pathway protein LEF1 in cancer, as a biomarker for prognosis and a target for treatment. *Am J Cancer Res* **7**, 1389–1406 (2017).

249. Assaraf, Y. G. *et al.* The multi-factorial nature of clinical multidrug resistance in cancer. *Drug Resist Updat* **46**, 100645 (2019).

250. Moutinho, C. *et al.* Epigenetic inactivation of the BRCA1 interactor SRBC and resistance to oxaliplatin in colorectal cancer. *J Natl Cancer Inst* **106**, djt322 (2014).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

251. Russo, M. *et al.* Tumor Heterogeneity and Lesion-Specific Response to Targeted Therapy in Colorectal Cancer. *Cancer Discov* **6**, 147–153 (2016).

252. Klein-Scory, S. *et al.* Significance of Liquid Biopsy for Monitoring and Therapy Decision of Colorectal Cancer. *Transl Oncol* **11**, 213–220 (2018).

253. Cayrefourcq, L. *et al.* Selective treatment pressure in colon cancer drives the molecular profile of resistant circulating tumor cell clones. *Mol Cancer* **20**, 30 (2021).

254. Müller, F. *et al.* RnBeads 2.0: comprehensive analysis of DNA methylation data. *Genome Biol* **20**, 55 (2019).

255. Carmona-Saez, P., Chagoyen, M., Tirado, F., Carazo, J. M. & Pascual-Montano, A. GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. *Genome Biol* **8**, R3 (2007).

256. Luo, M., Yang, X., Chen, H.-N., Nice, E. C. & Huang, C. Drug resistance in colorectal cancer: An epigenetic overview. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1876**, 188623 (2021).

257. Mattiuzzi, C., Sanchis-Gomar, F. & Lippi, G. Concise update on colorectal cancer epidemiology. *Ann Transl Med* **7**, 609 (2019).

258. Marusyk, A., Janiszewska, M. & Polyak, K. Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. *Cancer Cell* **37**, 471–484 (2020).

259. Zhao, X.-Q., Zhang, Y.-F., Xia, Y.-F., Zhou, Z.-M. & Cao, Y.-Q. Promoter demethylation of nuclear factor-erythroid 2-related factor 2 gene in drug-resistant colon cancer cells. *Oncol Lett* **10**, 1287–1292 (2015).

260. Ju, S. *et al.* Detection of circulating tumor cells: opportunities and challenges. *Biomark Res* **10**, 58 (2022).

261. Romero-Garcia, S., Prado-Garcia, H. & Carlos-Reyes, A. Role of DNA Methylation in the Resistance to Therapy in Solid Tumors. *Front Oncol* **10**, 1152 (2020).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

262. Wen, S. *et al.* Nucleoside diphosphate kinase 2 confers acquired 5-fluorouracil resistance in colorectal cancer cells. *Artif Cells Nanomed Biotechnol* **46**, 896–905 (2018).

263. Cho, J.-W. *et al.* The importance of enhancer methylation for epigenetic regulation of tumorigenesis in squamous lung cancer. *Exp Mol Med* **54**, 12–22 (2022).

264. Smit, D. J. *et al.* High Sensitivity of Circulating Tumor Cells Derived from a Colorectal Cancer Patient for Dual Inhibition with AKT and mTOR Inhibitors. *Cells* **9**, (2020).

265. Stefani, C. *et al.* Growth Factors, PI3K/AKT/mTOR and MAPK Signaling Pathways in Colorectal Cancer Pathogenesis: Where Are We Now? *Int J Mol Sci* **22**, (2021).

266. Grossi, V., Peserico, A., Tezil, T. & Simone, C. p38α MAPK pathway: a key factor in colorectal cancer therapy and chemoresistance. *World J Gastroenterol* **20**, 9744–58 (2014).

267. Lin, J. *et al.* The POU2F1-ALDOA axis promotes the proliferation and chemoresistance of colon cancer cells by enhancing glycolysis and the pentose phosphate pathway activity. *Oncogene* **41**, 1024–1039 (2022).

268. Ha, C.-T. *et al.* ID4 predicts poor prognosis and promotes BDNF-mediated oncogenesis of colorectal cancer. *Carcinogenesis* **42**, 951–960 (2021).

269. Hua, Z. *et al.* PI3K and MAPK pathways mediate the BDNF/TrkB-increased metastasis in neuroblastoma. *Tumour Biol* **37**, 16227–36 (2016).

270. Jaboin, J., Kim, C. J., Kaplan, D. R. & Thiele, C. J. Brain-derived neurotrophic factor activation of TrkB protects neuroblastoma cells from chemotherapy-induced apoptosis via phosphatidylinositol 3'-kinase pathway. *Cancer Res* **62**, 6756–63 (2002).

271. Li, K. *et al.* Development of a Phagocytosis-Dependent Gene Signature to Predict Prognosis and Response to Checkpoint Inhibition in Clear-Cell Renal Cell Carcinoma. *Front Immunol* **13**, 853088 (2022).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

272. Qiu, L. *et al.* FKBP11 promotes cell proliferation and tumorigenesis via p53-related pathways in oral squamous cell carcinoma. *Biochem Biophys Res Commun* **559**, 183–190 (2021).

273. Wu, X. *et al.* USP3 promotes gastric cancer progression and metastasis by deubiquitination-dependent COL9A3/COL6A5 stabilisation. *Cell Death Dis* **13**, 10 (2021).

274. Peng, X., Chen, X., Zhu, X. & Chen, L. GALNT6 Knockdown Inhibits the Proliferation and Migration of Colorectal Cancer Cells and Increases the Sensitivity of Cancer Cells to 5-FU. *J Cancer* **12**, 7413–7421 (2021).

275. Liu, C. *et al.* GALNT6 promotes breast cancer metastasis by increasing mucin-type O-glycosylation of α2M. *Aging* **12**, 11794–11811 (2020).

276. Shi, Z. *et al.* Hypoxia-induced Nur77 activates PI3K/Akt signaling via suppression of Dicer/let-7i-5p to induce epithelial-to-mesenchymal transition. *Theranostics* **11**, 3376–3391 (2021).

277. Liu, X. *et al.* High expression of AP2M1 correlates with worse prognosis by regulating immune microenvironment and drug resistance to R-CHOP in diffuse large B cell lymphoma. *Eur J Haematol* (2022) doi:10.1111/ejh.13895.

278. le Duff, M. *et al.* Regulation of senescence escape by the cdk4–EZH2–AP2M1 pathway in response to chemotherapy. *Cell Death Dis* **9**, 199 (2018).

279. Guillon, J. *et al.* Chemotherapy-induced senescence, an adaptive mechanism driving resistance and tumor heterogeneity. *Cell Cycle* **18**, 2385–2397 (2019).

280. Jakhar, R. *et al.* Autophagy Governs Protumorigenic Effects of Mitotic Slippage-induced Senescence. *Mol Cancer Res* **16**, 1625–1640 (2018).

281. Loree, J. M. & Kopetz, S. Recent developments in the treatment of metastatic colorectal cancer. *Ther Adv Med Oncol* **9**, 551–564 (2017).

282. Neugut, A. I. *et al.* FOLFOX and FOLFIRI Use in Stage IV Colon Cancer: Analysis of SEER-Medicare Data. *Clin Colorectal Cancer* **18**, 133–140 (2019).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

283. Nappi, A. *et al.* Metastatic Colorectal Cancer: Role of Target Therapies and Future Perspectives. *Curr Cancer Drug Targets* **18**, 421–429 (2018).

284. Barbazán, J. *et al.* A multimarker panel for circulating tumor cells detection predicts patient outcome and therapy response in metastatic colorectal cancer. *Int J Cancer* **135**, 2633–2643 (2014).

285. Hu, M. *et al.* Circulating tumor cells in colorectal cancer in the era of precision medicine. *J Mol Med (Berl)* **100**, 197–213 (2022).

286. Hardingham, J. E. *et al.* Detection and Clinical Significance of Circulating Tumor Cells in Colorectal Cancer--20 Years of Progress. *Mol Med* **21 Suppl 1**, S25-31 (2015).

287. Krueger F. 2015. Trim Galore!: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. http://www.bioinformatics.babraham.ac.uk/projects/ trim_galore/.

288. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–2 (2011).

289. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–9 (2012).

290. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–9 (2009).

291. R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

292. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847-2849 (2016).

293. Gu, Z. et al. circlize Implements and enhances circular visualization in R. Bioinformatics 30, 2811–2812 (2014).

294. Therneau, T. M. & Grambsch, P. M. The Cox Model. in 39–77 (2000). doi:10.1007/978-1-4757-3294-8_3.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

295. Cohen, S. J. *et al.* Prognostic significance of circulating tumor cells in patients with metastatic colorectal cancer. *Ann Oncol* **20**, 1223–9 (2009).

296. Petrik, J. *et al.* Circulating Tumor Cells in Colorectal Cancer: Detection Systems and Clinical Utility. *Int J Mol Sci* **23**, (2022).

297. Wang, Z. *et al.* High-Efficiency Isolation and Rapid Identification of Heterogeneous Circulating Tumor Cells (CTCs) Using Dual-Antibody-Modified Fluorescent-Magnetic Nanoparticles. *ACS Appl Mater Interfaces* **11**, 39586–39593 (2019).

298. Ferreira, M. M., Ramani, V. C. & Jeffrey, S. S. Circulating tumor cell technologies. *Mol Oncol* **10**, 374–94 (2016).

299. Andree, K. C., van Dalum, G. & Terstappen, L. W. M. M. Challenges in circulating tumor cell detection by the CellSearch system. *Mol Oncol* **10**, 395–407 (2016).

300. Yu, M. *et al.* Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* **339**, 580–4 (2013).

301. Bhagat, A. A. S. & Lim, C. T. Microfluidic technologies. *Recent Results Cancer Res* **195**, 59–67 (2012).

302. Mendelaar, P. A. J. *et al.* Defining the dimensions of circulating tumor cells in a large series of breast, prostate, colon, and bladder cancer patients. *Mol Oncol* **15**, 116–125 (2021).

303. Wang, L. *et al.* Circulating tumor cells as an independent prognostic factor in advanced colorectal cancer: a retrospective study in 121 patients. *Int J Colorectal Dis* **34**, 589–597 (2019).

304. Ocaña, A. *et al.* Circulating DNA and Survival in Solid Tumors. *Cancer Epidemiol Biomarkers Prev* **25**, 399–406 (2016).

305. Das, A. *et al.* Clinico-pathological correlation of serial measurement of circulating tumor cells in 24 metastatic colorectal cancer patients receiving chemotherapy reveals interpatient heterogeneity correlated with CEA levels but independent of KRAS and BRAF mutation. *Cancer Biol Ther* **16**, 709–13 (2015).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

306. Yalcin, S. *et al.* Determination of circulating tumor cells for detection of colorectal cancer progression or recurrence. *Hepatogastroenterology* **57**, 1395–8 (2010).

307. Lu, C.-Y. *et al.* Molecular detection of persistent postoperative circulating tumour cells in stages II and III colon cancer patients via multiple blood sampling: prognostic significance of detection for early relapse. *Br J Cancer* **104**, 1178–84 (2011).

308. Uen, Y.-H. *et al.* Persistent presence of postoperative circulating tumor cells is a poor prognostic factor for patients with stage I-III colorectal cancer after curative resection. *Ann Surg Oncol* **15**, 2120–8 (2008).

309. Friedlander, T. W., Premasekharan, G. & Paris, P. L. Looking back, to the future of circulating tumor cells. *Pharmacol Ther* **142**, 271–80 (2014).

310. Esteller, M. *et al.* Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med* **343**, 1350–4 (2000).

311. Zhang Z.M. et al. TFAP2E hypermethylation was associated with survival advantage in patients with colorectal cancer. J Cancer Res Clin Oncol 140, 2119-2127 (2014).

312. Adhikari, S. *et al.* The paradigm of drug resistance in cancer: an epigenetic perspective. *Biosci Rep* **42**, (2022).

313. Parikh, A. R. & Corcoran, R. B. Monitoring resistance through liquid biopsy. *Ann Oncol* **29**, 8–11 (2018).

314. Esteller, M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* **21**, 5427–5440 (2002).

315. Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* **13**, 484–92 (2012).

316. Wallace, D. C. Mitochondria and cancer. *Nat Rev Cancer* **12**, 685–98 (2012).

317. Lin, Y.-H. *et al.* Functional Role of Mitochondrial DNA in Cancer Progression. *Int J Mol Sci* **23**, (2022).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

318. Kubo, Y. *et al.* Low mitochondrial DNA copy number induces chemotherapy resistance via epithelial-mesenchymal transition by DNA methylation in esophageal squamous cancer cells. *J Transl Med* **20**, (2022).

319. Liu, Z., Tian, J., Peng, F. & Wang, J. Hypermethylation of mitochondrial DNA facilitates bone metastasis of renal cell carcinoma. *J Cancer* **13**, 304–312 (2022).

320. Liu, Y. *et al.* An Epigenetic Role of Mitochondria in Cancer. *Cells* **11**, (2022).

321. Mu, Z. *et al.* Detection and Characterization of Circulating Tumor Associated Cells in Metastatic Breast Cancer. *Int J Mol Sci* **17**, (2016).

322. Hodgkinson, C. L. *et al.* Tumorigenicity and genetic profiling of circulating tumor cells in small-cell lung cancer. *Nat Med* **20**, 897–903 (2014).

323. Paoletti, C. *et al.* Comprehensive Mutation and Copy Number Profiling in Archived Circulating Breast Cancer Tumor Cells Documents Heterogeneous Resistance Mechanisms. *Cancer Res* **78**, 1110–1122 (2018).

324. Vishnoi, M. *et al.* The isolation and characterization of CTC subsets related to breast cancer dormancy. *Sci Rep* **5**, 17533 (2015).

325. Palmirotta, R. *et al.* Next-generation Sequencing (NGS) Analysis on Single Circulating Tumor Cells (CTCs) with No Need of Whole-genome Amplification (WGA). *Cancer Genomics Proteomics* **14**, 173–179 (2017).

326. Heymann, D. & Téllez-Gabriel, M. Circulating Tumor Cells: The Importance of Single Cell Analysis. *Adv Exp Med Biol* **1068**, 45–58 (2018).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

AGRADECIMIENTOS

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

AGRADECIMIENTOS

Tengo que empezar agradeciendo a mis directores de tesis Ángel y Rafa, por haber apostado por mí y haberme dado la oportunidad de iniciarme con Oncomet en esta increíble experiencia en el mundo de la investigación y por haberme guiado durante estos años, en los que he aprendido tanto.

Gracias, a los pacientes y sus familiares, por vuestra generosidad y altruismo.

Gracias a todo Oncomet, al Servicio de Oncología, especialmente a Yolanda, Francis, Sonia y, sobretodo, a Elena, por resolvernos tantas dudas y el aprendizaje que me llevo de todas nuestras reuniones. Gracias a la UM y al laboratorio de Biopsia líquida, gracias a todos por vuestra ayuda y por el aprendizaje de todos los momentos que hemos compartido. En especial, gracias a Laura, por tu implicación y tu disposición siempre a ayudar, por tus felicitaciones cuando sale algo bien y por interesarte cuando algo va mal.

Gracias también al grupo que me abrió las puertas en Francia. Merci beaucoup au Laboratoire Cellules Circulantes Rares Humaines de Montpellier. Thank you, Catherine, for being so welcoming, so kind and so inspiring and encouraging, it is a pleasure working with you. Thank you to all the lab team, Laure, Françoise, Thomas, Doryan and Luis for your help, for all the chats and the crepes under the sun; but very especially to Zahra (and Mostafa) for being so nice and kind and for all the help, and of course to Keerthi, for taking care of me, for being incredibly nice and generous and for being always willing to help. Thank you for the long conversations while fighting with the DEPArray and for not hating me for the infinite mails asking you to print something for me. Merci du fond du cœur pour tout, le temps que nous avons partagé me manque.

Gracias al laboratorio 13, a Manuel, Sabela, Alba, Francisco, Pilar y especialmente a Pablo, por tu paciencia y por enseñarme tantísimo. Mil

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

gracias por vuestra ayuda y por todos los momentos geniales que compartimos.

Gracias a los compañeros del IDIS con los que he coincidido en esta etapa, a Tais, a Majo por tus ánimos y los ratos de desahogo; y al laboratorio 2, a Jesús, Paula y en especial a Andrea, por tu ayuda y tus consejos.

Gracias a Ali, por tu ayuda, has sido indispensable; y a Ramón, el MacGyver de la ddPCR, por echar siempre una mano para lo que fuera, por escuhar siempre a los de Epi y por tantos momentos de partirnos de risa que hemos compartido en el café, en la 5ª, con unas cañas o donde fuera.

Gracias inifinitas, a los oncolineros. A mis compis de Unidad, los chicos Epi, Nico y Aitor, este trabajo también es vuestro. A Nico, por estar siempre dispuesto a echar una mano y por tus conocimientos de bioinformática que tanto me han salvado la vida ¡Mil gracias! A Aitor, no sé ni por dónde empezar… Cuántas veces hemos dicho eso de "Que bien hemos dado" y la verdad es que sí. Gracias por estos años, por tu apoyo, por tu disposición a ayudar siempre, por implicarte en mis proyectos como si fueran propios e intentar buscar siempre soluciones a todo. Gracias también por las numerosas obras de arte pegadas en mi escritorio, los bolis fantasía y el alijo de Schoko-Bons de tu mesa del que tanto me he surtido estas últimas semanas. No me imagino otra persona con la que procesar muestras a las tantas al ritmo de Sonia y Selena, ni con quien vivir todos los momentos surrealistas (que no han sido pocos) que tanto hemos comentado hasta llorar de la risa. Muchas gracias, trabajar codo con codo con un amigo lo ha hecho todo mucho más llevadero. Gracias también a Manu, eres casi uno más de Epi.

Gracias a Inés y Nuria, por escuchar mis anécdotas y por tener siempre palabras de ánimo, a Pablo, por tu optimismo y estar siempre para echar una mano cuando gritaba ¡Ayuda! desde la sala del microscopio, si no fuera por ti, aún estaría allí sentada micromanipulando PBS; a las nanos, pero

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

en especial a las que fueron labo 13, a Sai y Sandra A. por echar siempre una mano y a Sandra, mi San, por estar siempre dispuesta a ayudar y a escucharme, siempre, sin juzgar, da igual la locura que te contara, has sido imprescindible, hemos llevado caminos paralelos durante toda esta etapa y en este tiempo de escritura compartir nuestras quejas y nuestro agobio, me han hecho sentir mucho más ligera y acompañada, no te imaginas cuánto ¡Mil gracias chuli!; a Ana y a Rachel, por vuestro buen rollo y positivismo; a Carol y a Óscar, por vuestros ánimos y buen humor y contagiar siempre una sonrisa, por estar para lo que hiciera falta y por ser mis compañeros infalibles de zumito de piña; a Carlos, mi Carlis, gracias por tu generosidad, por estar echando una mano antes de saber que la necesito y por hacerme saber siempre que puedo contar contigo para lo que sea, por haber sido un apoyo en los momentos más bajos, gracias por tu sofá y por ser el mayor "generador de planes" del universo, has sido fundamental para mi; a Patri, mi Pachuliña, por apuntarte la primera ya sea a echar una mano o a un tiroteo, por tu positivismo, por buscar el lado bueno de todo y por escuchar mis dramas y darme siempre consejo, has sido imprescindible para mi. Gracias chicos, por tantos y tantos momentos memorables que hemos vivido juntos, por estar siempre para celebrar los triunfos y para ahogar las penas, por todas las anécdotas, los incontables momentos de morirnos de risa y por haber hecho esta etapa tan especial. Habéis sido imprescindibles para mí y sin vosotros no habría llegado hasta aquí, esta frase es un clásico, pero en mi caso, no podría ser más real. Sois lo mejor que me llevo de esta etapa, con diferencia, y ya solo por eso, estaré siempre agradecida de haber iniciado este camino. Gracias de corazón a todos, por llenarme tanto la mochila y haberme hecho crecer tantísimo en lo profesional y en lo personal.

A mis padres, por vuestro apoyo infinito, por creer en mi más que yo misma y por animarme siempre a hacer lo que me gusta. A mi hermano, por salir al rescate en los momentos bajos. A Mario, por estar siempre, por tu cariño y tu apoyo incondicional.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

ANNEXES

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

ANNEX I

ETHICS COMMITTEE APPROVAL

XUNTA DE GALICIA
CONSELLERÍA DE SANIDADE
Secretaría Xeral Técnica

Edificio Administrativo San Lázaro
15703 SANTIAGO DE COMPOSTELA
Teléfono: 881546425
ceic@sergas.es

galicia

DICTAMEN DEL Comité de ética de la investigación con medicamentos de Galicia (CEIm-G)

Paula M. López Vázquez, Secretaria del Comité de ética de la investigación con medicamentos de Galicia (CEIm-G)

CERTIFICA:

Que este Comité evaluó en su reunión del día 28/02/19 :

Título: Desarrollo de una firma epigenética en biopsia líquida para predecir la respuesta terapéutica en cáncer colorrectal metastásico
Promotor*:* Ángel Díaz Lagares
Tipo de estudio: EPA OD
Version: v02 de 09/12/2019 y HIP/CI de la misma fecha
Código del Promotor:ADL-FOL-2019-01
Código de Registro: 2019/017

Y, tomando en consideración las siguientes cuestiones:

- La pertinencia del estudio, teniendo en cuenta el conocimiento disponible, así como los requisitos legales aplicables.
- La idoneidad del protocolo en relación con los objetivos del estudio, justificación de los riesgos y molestias previsibles para el sujeto, así como los beneficios esperados.
- Los principios éticos de la Declaración de Helsinki vigente.
- Los Procedimientos Normalizados de Trabajo del CEIm-G de Galicia

Emite un **INFORME FAVORABLE** para la realización del estudio por el/la investigador/a del centro:

Centros	Investigadores Principales
C.H. Universitario de Santiago	Ángel Díaz Lagares

Xacobeo 2021

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ANNEX I

XUNTA DE GALICIA
CONSELLERÍA DE SANIDADE
Secretaría Xeral Técnica

Edificio Administrativo San Lázaro
15703 SANTIAGO DE COMPOSTELA
Teléfono: 881546425
ceic@sergas.es

galicia

Y HACE CONSTAR QUE:

1. El comité cumple los requisitos legales vigentes aplicables a los Comités de ética de investigación.
2. El comité tanto en su composición como en sus PNTs cumple las Normas de Buena Práctica Clínica (CPMP/ICH/135/95).
3. La composición actual del comité es:

Susana María Romero Yuste (Presidenta). Médico Especialista en Reumatología.
Irene Zarra Ferro. (Vicepresidenta). Farmacéutica de Atención Especializada.
Paula Mª López Vázquez, (Secretaria). Médico Especialista en Farmacología Clínica.
Belen Pérez Marcos (Vicesecretaria). Farmacéutica de Atención Especializada
Jesús Alberdi Sudupe. Médico especialista en Psiquiatría.
Rosendo Bugarín González. Médico Especialista en Medicina Familiar y Comunitaria.
Juan Casariego Rosón. Médico Especialista en Cardiología.
Nuria Carballeda Feijóo. Miembro lego. Representante de los intereses de los pacientes.
Juana Mª Cruz del Río. Trabajadora Social.
Ana Belen Cruz Valiño. Licenciada en Derecho.
Juan Fernando Cueva Bañuelos. Médico Especialista en Oncología Médica.
José Álvaro Fernández Rial. Médico Especialista en Medicina Interna.
José Luis Fernández Trisac. Médico Especialista en Pediatría.
Mª José Ferreira Díaz. Diplomada Universitaria de Enfermería.
Agustín Pía Morandeira. Farmacéutico de Atención Primaria
Carmen Rodríguez-Tenreiro Sánchez. Licenciada en Farmacia.
Diego Santos García. Médico Especialista en Neurología.
Juan Vázquez Lago. Médico Especialista en Medicina Preventiva y Salud Pública.
Mª Asunción Verdejo González. Médico Especialista en Farmacología Clínica.

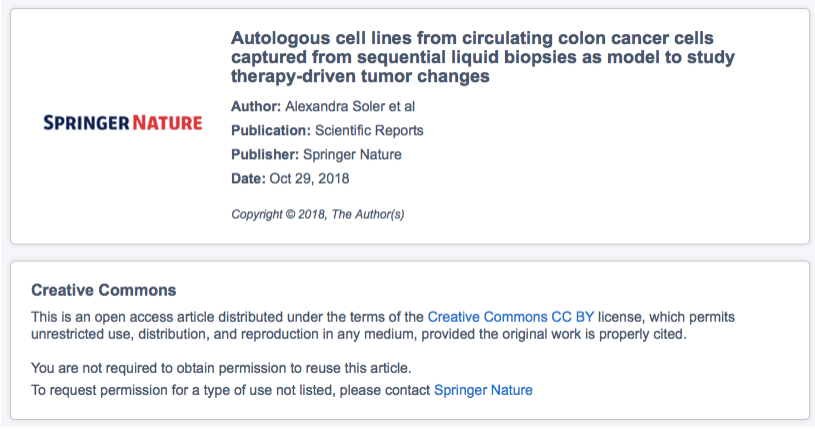
En Santiago de Compostela,

Xacobeo 2021

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

IMAGES USE

Images from Figure 1 to Figure 9 have been created by the author using Biorender.com. In the case of Figure 5, this was adapted from other the published article “Autologous cell lines from circulating colon cancer cells captured from sequential liquid biopsies as model to study therapy-driven tumor changes” by Alexandra Soler, Laure Cayrefourcq, Thibault Mazard, Anna Babayan, Pierre-Jean Lamy, Said Assou, Eric Assenat, Klaus Pantel, Catherine Alix-Panabières, with permission of the publishers.



SPRINGER NATURE

Autologous cell lines from circulating colon cancer cells captured from sequential liquid biopsies as model to study therapy-driven tumor changes

Author: Alexandra Soler et al
Publication: Scientific Reports
Publisher: Springer Nature
Date: Oct 29, 2018

Copyright © 2018, The Author(s)

Creative Commons

This is an open access article distributed under the terms of the Creative Commons CC BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

You are not required to obtain permission to reuse this article.
To request permission for a type of use not listed, please contact Springer Nature

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

ANNEX III

PUBLICATIONS

1. Bao-Caamano A, Costa-Fraga N, Cayrefourcq L, Jácome MA, Rodriguez-Casanova A, Muinelo-Romay L, López-López R, Alix-Panabières C, Díaz-Lagares A. Epigenomic analysis reveals a specific DNA methylation program of metastasis-competent circulating tumor cells in colorectal cancer. (*Submitted*)

2. Bao-Caamano A, Costa-Fraga N, Cayrefourcq L, Rodriguez-Casanova A, Muinelo-Romay L, López-López R, Alix-Panabières C, Díaz-Lagares A. Epigenomic reprogramming of circulating tumor cells is associated with the development of therapy resistance in colon cancer. (*In preparation*)

3. Bao-Caamano A, Costa-Fraga N, Cayrefourcq L, Abalo A, Kurma K, Earl J, Castillo M, Brozos-Vazquez E, Longo F, Vidal-Insua Y, Vazquez-Rivera F, Candamio-Folgar S, Ferreiro R, Reguera P, Villalón-López J, Lago-Lestón RM, Rodriguez-Casanova A, López-López R, Alix-Panabières C, Muinelo-Romay L, Díaz-Lagares A. Whole-genome bisulfite sequencing of circulating tumor cells identifies epigenetic signatures of primary and secondary resistance to therapy in colorectal cancer patients. (*In preparation*)

4. Rodriguez-Casanova A*, Costa-Fraga N*, **Bao-Caamano A***, López-López R, Muinelo-Romay L, Diaz-Lagares A. Epigenetic Landscape of Liquid Biopsy in Colorectal Cancer**.** Front Cell Dev Biol 2021. 9:622459. DOI: 10.3389/fcell.2021.622459. *Co-first authors. Review

5. Bao-Caamano A, Rodriguez-Casanova A, Diaz-Lagares A. Epigenetics of Circulating Tumor Cells in Breast Cancer. Adv Exp Med Biol 2020. 1220:117-134. DOI: 10.1007/978-3-030-35805-1_8. Review

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

OTHER PUBLICATIONS

1. Rapado-González Ó, Brea-Iglesias J, Rodríguez-Casanova A, **Bao-Caamano A**, López-Cedrún JL, Triana-Martínez G, Díaz-Peña R, Santos MA, López-López R, Muinelo-Romay L, Martínez-Fernández M, Díaz-Lagares Á, Suárez-Cunqueiro MM. Somatic mutations in tumor and plasma of locoregional recurrent and/or metastatic head and neck cancer using a next-generation sequencing panel: A preliminary study. Cancer Med 2022. DOI: 10.1002/cam4.5436.

2. Rapado-González Ó, Costa-Fraga N, **Bao-Caamano A**, López-Cedrún JL, Álvarez-Rodríguez R, Crujeiras AB, Muinelo-Romay L, López-López R, Díaz-Lagares Á, Suárez-Cunqueiro MM. Genome-wide DNA methylation profiling in tongue squamous cell carcinoma**.** Oral Dis 2022. DOI: 10.1111/odi.14444.

3. Sonzini G, Granados-Aparici S, Sanegre S, Diaz-Lagares A, Diaz-Martin J, de Andrea C, Eritja N, **Bao-Caamano A**, Costa-Fraga N, García-Ros D, Salguero-Aranda C, Davidson B, López-López R, Melero I, Navarro S, Ramon Y Cajal S, de Alava E, Matias-Guiu X, Noguera R. Integrating digital pathology with transcriptomic and epigenomic tools for predicting metastatic uterine tumor aggressiveness. Front Cell Dev Biol 2022. 10:1052098. DOI: 10.3389/fcell.2022.1052098.

4. Rodriguez-Casanova A, Costa-Fraga N, Castro-Carballeira C, González-Conde M, Abuin C, **Bao-Caamano A**, García-Caballero T, Brozos-Vazquez E, Rodriguez-López C, Cebey V, Palacios P, Cueva JF, López-López R, Costa C, Díaz-Lagares A. A genome-wide cell-free DNA methylation analysis identifies an episignature associated with metastatic luminal B breast cancer. Front Cell Dev Biol 2022. 10:1016955. DOI: 10.3389/fcell.2022.1016955.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

5. Ruiz-Bañobre J, Rodriguez-Casanova A, Costa-Fraga N, **Bao-Caamano A**, Alvarez-Castro A, Carreras-Presas M, Brozos-Vazquez E, Vidal-Insua Y, Vazquez-Rivera F, Candamio-Folgar S, Mosquera-Presedo M, Lago-Lestón RM, Muinelo-Romay L, Vázquez-Bueno JÁ, Sanz-Pamplona R, Moreno V, Goel A, Castillo L, Martin AC, Arroyo R, Esteller M, Crujeiras AB, López-López R, Díaz-Lagares A. Noninvasive early detection of colorectal cancer by hypermethylation of the LINC00473 promoter in plasma cell-free DNA. Clin Epigenetics 2022. 14(1):86. DOI: 10.1186/s13148-022-01302-x.

6. Rodríguez-Casanova A*, **Bao-Caamano A***, Lago-Lestón RM, Brozos-Vázquez E, Costa-Fraga N, Ferreirós-Vidal I, Abdulkader I, Vidal-Insua Y, Rivera FV, Candamio Folgar S, López-López R, Muinelo-Romay L, Diaz-Lagares A. Evaluation of a Targeted Next-Generation Sequencing Panel for the Non-Invasive Detection of Variants in Circulating DNA of Colorectal Cancer. J Clin Med 2021. 10(19):4487. DOI: 10.3390/jcm10194487. *Co-first authors

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA





Metastatic colorectal cancer continues to be a major health burden despite advances in detection and therapy. Molecular characterization of circulating tumour cells (CTCs), main drivers of the metastatic process, and the study of epigenetic mechanisms hold great promise to identify new targets for drug development and biomarkers to assess therapy response. In this frame, this thesis characterized the DNA methylome of CTCs in advanced colorectal cancer, providing new information about the biology of these cells and the metastatic process. This study also identified new potential therapeutic targets and epigenetic biomarkers that could help guide therapeutic decisions in patients with advanced disease, contributing to achieve a more personalized medicine in colorectal cancer.