Universidade de Lisboa

Faculdade de Medicina de Lisboa

PLCE1 rs2274223 A>G Polymorphism and its Functional Role in Colorectal Cancer

António Manuel Marujo Palma

Orientadores: Doutora Marta Sofia Alves Martins

Prof. Doutor Luís António Marquês da Costa

Dissertação especialmente elaborada para obtenção do grau de Mestre em Oncobiologia

2019

Universidade de Lisboa

Faculdade de Medicina de Lisboa

PLCE1 rs2274223 A>G Polymorphism and its Functional Role in Colorectal Cancer

António Manuel Marujo Palma

Orientadores: Doutora Marta Sofia Alves Martins

Prof. Doutor Luís António Marquês da Costa

Dissertação especialmente elaborada para obtenção do grau de Mestre em Oncobiologia

2019

Todas as afirmações efetuadas no presente documento são da exclusiva responsabilidade do seu autor, não cabendo qualquer responsabilidade à Faculdade de Medicina de Lisboa pelos conteúdos nele apresentados.

A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de 16 de Abril de 2019.

Agradecimentos

Quero dedicar este trabalho ao meu tio António pela força e coragem que tem tido a lutar contra esta doença e por essa razão tornou-se num exemplo para eu nunca desistir de atingir os meus objetivos.

Quero ainda dedicar este trabalho à minha avó Dolores por todos os ensinamentos, pelo carinho e pelo amor que me deu. Infelizmente não conseguiu presenciar o seu neto a finalizar mais uma importante etapa na sua vida. Seja onde for, eu sei que estás comigo e nunca me vou esquecer de ti.

Não teria sido possível chegar até aqui e terminar este trabalho sem o apoio incondicional de pessoas a quem eu quero expressar o meu mais sincero agradecimento.

Quero agradecer em primeiro lugar à minha orientadora, Dra. Marta Martins por ter confiado em mim para realizar este trabalho, por todo o apoio, ensinamentos e amizade que me deu. Quero agradecer, ao professor Luís Costa por me ter acolhido no seu laboratório como estudante de mestrado e por todos os conselhos e conhecimentos transmitidos durante este ano.

Quero agradecer a todos os meus colegas do Luís Costa Lab pela ajuda, companhia e diversão que me proporcionaram e que fizeram com que fosse muito mais fácil percorrer este caminho. Quero agradecer à Sandra pelos conselhos e ensinamentos, à Daniela, por ter sido desde o início a minha colega de bancada e todas as conversas, discussões e trabalhos que fizemos juntos não podiam ser esquecidos. Raquel, obrigado por toda a paciência que tiveste comigo. Obrigado pela ajuda, pelos conselhos e pela amizade que quero manter para sempre.

Não posso esquecer também a Inês, a Patrícia Corredeira, a Patrícia Alves e a Catarina pela ajuda e pela amizade.

Quero agradecer à Dr. Catarina Abreu pela ajuda que me deu a conseguir toda a informação clínica necessária e ao Dr. Arlindo Ferreira por todos os conselhos e pela ajuda que me deu na análise estatística.

Um agradecimento especial a todo o pessoal do Sérgio Dias Lab, com eles partilhei o espaço, mas também muitos momentos de diversão, muita amizade e muitos conhecimentos.

O meu maior agradecimento à minha família, em particular aos meus pais.

Foram eles que me incentivaram a realizar este mestrado. Nada é suficiente para agradecer todos os sacrifícios que eles passaram para tornar este momento possível, obrigado pelo apoio e palavras de motivação que me deram a todo o momento, especialmente nas horas mais difíceis. Obrigado por me ajudarem a definir o caminho.

Abstract

Colorectal cancer (CRC) is one of the cancers with highest incidence and mortality in the world. In order to decrease these numbers earlier diagnosis, better treatments and better surveillance is required.

In this context, cancer biomarkers arise as important tools useful for diagnosis, monitoring disease progression, predicting disease recurrence and therapeutic treatment efficacy. Therefore, it is necessary to find new CRC biomarkers highly sensitive and specific to help clinical decisions.

In the last decade, Phospholipase C epsilon (PLCε) has been studied as a possible biomarker for CRC, particularly its single nucleotide polymorphism (SNP) rs2274223 has been associated with the risk of CRC development.

In our study, we aimed, not only, to assess the risk of *PLCE1* SNP rs2274223 in CRC development in a Portuguese population, but also, analyse how this polymorphism affected patient's survival. Furthermore, we also investigated how this polymorphism influenced cellular processes such as proliferation, epithelial-to-mesenchymal transition (EMT), angiogenesis and inflammation.

Overall, our results show that *PLCE1* SNP rs2274223 A>G is not associated with the risk of developing CRC. Furthermore, this polymorphism was not shown to be involved in CRC survival of stages I-III and IV patients.

Moreover, we could not associate this phenotype with any abnormal cellular process.

In sum, we found by several means that *PLCE1* SNP rs2274223 A>G appears to have no role on CRC development and progression. Our findings are contrary to most of the published reports about this SNP.

Keywords: Colorectal cancer; Phospholipase Cε; SNP rs2274223; Risk; Survival.

Resumo

O cancro colorretal é o quarto cancro mais diagnosticado e o terceiro com maior mortalidade no mundo, sendo que em Portugal é o terceiro mais diagnosticado e o segundo com maior mortalidade.

De forma a diminuir a mortalidade associada a esta doença é necessário diagnosticar a mesma em estadios iniciais, desenvolver melhores terapêuticas e ainda melhorar o acompanhamento da doença. É neste sentido, que os biomarcadores podem desempenhar um papel importante.

Por essa razão, é necessário descobrir novos biomarcadores para cancro colorretal com maior sensibilidade e especificidade de forma a complementar e melhorar as decisões clínicas e terapêuticas para o doente.

As fosfolipases C (PLCs) são potenciais biomarcadores que se encontram expressos em todas as células do organismo e que participam em variadas funções celulares como proliferação, motilidade, invasão e diferenciação.

A principal função das PLCs é hidrolisar fosfatidilinositol 4,5-bisfosfato (PIP_2) existente na membrana celular dando origem a diacilglicerol (DAG) e inositol 1,4,5-trisfosfato (IP_3). O IP_3 é importante na regulação dos níveis de cálcio intracelular, enquanto que o DAG é capaz de ativar a proteína quinase C (PKC) e as suas vias a jusante.

Existem 6 famílias de PLCs (PLCβ, PLCγ, PLCδ, PLCη, PLCζ e PLCε), todas elas partilham domínios catalíticos comuns, no entanto, também apresentam domínios, estruturas e mecanismos de regulação específicos.

A PLCε tem expressão ubíqua em todos os tecidos, apesar da sua maior expressão se verificar no coração, pulmão e colon. Esta enzima apresenta domínios específicos como o domínio CDC25 na porção N-terminal e dois domínios de associação a RAS (RA1 e RA2) na porção C-terminal.

O domínio CDC25 foi demonstrado como tendo função de troca de guaninas, portanto ativador da proteína RAP1, enquanto os domínios de associação ao RAS, em particular o domínio RA2, são importantes para a translocação da enzima do citoplasma para a membrana plasmática, onde exerce a sua função.

Modelos animais transgénicos provaram que a depleção de PLCε pode levar ao desenvolvimento de hipertrofia cardíaca e ainda síndrome nefrótico.

Esta enzima também já foi associada à inflamação da pele, à neuro-inflamação e ao cancro.

No entanto, o papel desta PLC no cancro é controverso. Se por um lado foi documentado que em alguns cancros, como por exemplo do esófago, a PLCε exerce um papel oncogénico, noutros, como é o caso do cancro colorretal, a PLCε parece apresentar uma função supressora de tumor.

Nos últimos anos, têm sido vários os estudos que associam a PLCε a cancro colorretal, em particular o seu polimorfismo rs2274223 tem sido associado com um risco aumentado de desenvolver esta doença.

Este polimorfismo consiste na substituição de um nucleótido de adenina por uma guanina no gene *PLCE1*, que posteriormente se traduz na substituição de um aminoácido de histidina por um de arginina na posição 1927 da enzima, no seu domínio de ligação ao cálcio (C2).

Este polimorfismo foi sobretudo associado a um aumento do risco de desenvolver cancro esofágico e gástrico.

No entanto, em cancro colorretal, a sua função é controversa. Enquanto que foi possível associar este polimorfismo ao desenvolvimento de cancro colorretal na população chinesa e na população turca, o mesmo já não foi possível observar numa população lituana e letã.

No nosso estudo propusemo-nos a verificar se existe alguma associação entre o polimorfismo rs2274223 do gene *PLCE1* e o desenvolvimento de cancro colorretal, na população portuguesa. Para além disso, fomos ainda verificar, pela primeira vez, se este polimorfismo poderia ter algum impacto na sobrevida dos doentes com cancro colorretal em estadios I-III e em estadio IV.

Após a genotipagem de 218 pacientes com cancro colorretal e 221 respetivos controlos (equiparados para idade e sexo), os nossos resultados mostram que este polimorfismo não se associa com um maior risco de desenvolver cancro colorretal em nenhuma das diferentes associações que testámos (AA vs. AG p=1, AA vs. GG p=0,66, AA vs. AG+GG p=0,85, GG vs. AG p=0,66 e GG vs. AA+AG p=0,68). Apesar de negativo, este resultado vem corroborar os resultados dos trabalhos desenvolvidos em populações do norte da europa (lituana e letã), os quais também não encontravam correlação deste polimorfismo com o desenvolvimento desta doença.

Estes resultados podem ser explicados pelo facto de inicialmente este polimorfismo ser associado ao risco de desenvolver cancro gástrico e esofágico, podendo a PLCε apresentar um papel diferente em CRC. Além disso, a maioria dos estudos foi realizado na população chinesa, sendo que, existem estudos na população europeia que indicam que este polimorfismo não está associado com o risco de desenvolver CRC.

Uma das explicações para esta disparidade de resultados em populações diferentes pode ser a existência de um menor *Linkage Disequilibrium* (LD) na população europeia. Desta forma, este polimorfismo não seria o responsável pela carcinogénese na população chinesa, mas sim a existência de outros polimorfismos com um alto LD.

Resumindo, diferenças populacionais e de tipo de tumor podem, em parte, justificar a diferença dos nossos resultados com os anteriormente publicados.

Posteriormente, verificámos que, este polimorfismo não influencia a sobrevida dos doentes com cancro colorretal nos estadios I-III (p=0,805) em analises uni- e multivariadas controlando para as características dos pacientes e do tumor que podem afetar o prognóstico da doença como a idade, estadio, grau de diferenciação, localização do tumor, obstrução e ou perfuração, invasão vascular, linfática e ou neural e a presença de margens cirúrgicas com tumor à altura do diagnóstico.

Apesar de não existir uma correlação significativa com a sobrevida dos doentes diagnosticados em estadio IV após a análise multivariada controlando para a idade, a localização do tumor e o órgão onde se detetaram as metástases (p=0,089), os pacientes homozigóticos para o alelo G apresentam uma tendência negativa no seu tempo de vida. Uma vez que esta análise foi feita com apenas 6 pacientes homozigóticos para este alelo, seria importante aumentar este coorte de forma a obter um maior poder estatístico capaz de validar a tendência observada.

É importante referir que na análise multivariada no estadio I-III e no estadio IV o braço GG apresenta poucos doentes, pelo que um maior número de doentes neste braço poderiam conferir um maior poder estatístico à análise multivariada.

No entanto, fomos investigar *in vitro* os processos celulares relevantes para o desenvolvimento e progressão de cancro, com o objetivo de compreender melhor o papel deste polimorfismo.

Existem estudos que associam a abolição da PLCε com um aumento da proliferação, com um aumento da libertação de fatores angiogénicos como VEGF-A e com um aumento de fatores pró-inflamatórios como COX-2, CXCL-1, CXCL-2, TNF-α, IL-1β, IL-6 e STAT3.

No nosso estudo, para além de abordar a influência da sobre expressão das variantes de PLCε *wild-type* e PLCε mutante (H1927A) nestes processos, analisámos ainda de que forma este polimorfismo poderia afetar a atividade fosfolipídica da enzima e a transição epitélio-mesênquima (EMT) nas linhas de cancro colorretal HCT116 e DLD1.

Da análise *in vitro* verificamos que em termos funcionais, a sobre expressão de PLCε *wild-type* e a sobre expressão da variante PLCε mutante (H1927A) não apresentam diferenças

na atividade fosfolipídica da enzima. Contudo, a PLCε apresenta outros domínios funcionais específicos como o CDC25 e RA2, que poderão ser afetados por este SNP e que não foram testados neste trabalho.

Relativamente à expressão dos fatores angiogénicos e pró-inflamatórios anteriormente mencionados, verificámos que não existem diferenças significativas quando comparamos o efeito da sobre expressão da PLCε *wild-type* e mutante (H1927A). Contudo, a sobre expressão da PLCε *wild-type* não se traduz numa diminuição generalizada dos fatores pró-inflamatórios como já foi publicado, com exceção para TNF-α, cuja expressão diminui nas linhas DLD1, e IL-6, que diminui nas linhas HCT116 após sobre expressão de ambas as variantes de PLCε.

Fomos posteriormente verificar o efeito deste polimorfismo na proliferação celular. Assim, verificamos que a sobre expressão da PLCε (quer *wild-type* quer mutante H1927A) provoca uma diminuição na proliferação, sendo este resultado concordante com o papel supressor de tumor que é atribuído à PLCε neste tipo de tumor. No entanto, não existem diferenças na taxa de proliferação entre a sobre expressão da PLCε *wild-type* e da PLCε mutante (H1927A).

Finalmente, não conseguimos observar diferenças significativas entre o papel da PLCε *wild-type* e da sua forma mutada (H1927A) na expressão de marcadores como a E-caderina, N-caderina, Vimentina e Twist importantes no processo de EMT.

Em suma, o polimorfismo rs2274223 A>G do gene *PLCE1* não parece apresentar qualquer influência em CRC, uma vez que não verificámos qualquer associação entre o mesmo e o risco de desenvolver a doença, a sobrevivência dos pacientes e nenhuma das funções celulares por nós testadas (atividade fosfolipídica, proliferação, EMT, inflamação e angiogénese).

Palavras-chave: Cancro colorretal; Fosfolipase Cε; SNP rs2274223; Risco; Sobrevivência.

Abbreviations

^{3}H	Tritium
A	Arginine
APC	Adenomatous polyposis coli
ATCC	American Type Culture Collection
***BCL2*/BCL-2**	B-cell lymphoma 2
BMPR1A	Bone morphogenetic protein receptor, type IA
BRAF	B-RAF proto-oncogene
BSA	Bovine serum albumin
C	Cytosine
CA^{2+}	Ionized calcium
CDC25 domain	Cell division cycle 25 domain
CDH1	E-cadherin gene
CDH2	N-cadherin gene
cDNA	Complementary deoxyribonucleic acid
CEA	Carcinoembryonic antigen
CI	Confidence interval
CIMP	CpG Island methylation pathway
CIN	Chromosomal instability
CLL	Chronic lymphocytic leukaemia
CMV	Cytomegalovirus
CO_2	Carbon dioxide
COX-2	Cyclooxygenase 2
CRC	Colorectal cancer
CT	Cycle Threshold
CTC	Circulating tumour cells
CTCL	Cutaneous T cell lymphoma
ctDNA	Circulating tumour DNA
***CXCL1*/Cxcl-1**	(C-X-C motif) ligand 1
***CXCL2*/Cxcl-2**	(C-X-C motif) ligand 2
DAG	Diacylglycerol
Dil.	Dilution
DMEM	Dubelcco's modified eagle's medium
DNA	Deoxyribonucleic acid

DNase	Deoxyribonuclease
EC	European Commission
EDTA	Ethylenediaminetetra-acetic acid
EGF	Epithelial growth factor
EGFR	Epithelial growth factor receptor
EMT	Epithelial-to-mesenchymal transition
ERBB2	Receptor tyrosine-protein kinase erbB-2
ESMO	European Society for Medical Oncology
FBS	Fetal bovine serum
FOBT	Faecal occult blood test
FOLFIRI	Fluoropyrimidine plus Irinotecan
FOLFOX	Fluoropyrimidine plus oxaliplatin
G	Guanine
g	Gravitational force
G1/G2/G3	Grade 1/2/3
GC	Gastric cancer
GCO	Global Cancer Observatory
GAPDH	Human glyceraldehyde 3-phosphate dehydrogenase
gDNA	Genomic deoxyribonucleic acid
GEF	Guanine exchange factor
GFP	Green fluorescent protein
GTPase	Guanosine triphosphate hydrolase
GWAS	Genome-wide association study
H	Histidine
H1927A	Mutant PLCε
HCl	Hydrochloric acid
HR	Hazard ratio
HRP	Horseradish peroxidase
IBM	International Business Machines
***IL1B/*Il-1β**	Interleukin 1 beta
***IL6/*Il-6**	Interleukin 6
IMM	Instituto de Medicina Molecular
IP_3	Inositol 1,4,5-trisphosphate
KM	Kaplan-Meier

KRAS	Kirsten rat sarcoma viral oncogene homolog
LD	Linkage disequilibrium
LiCl	Lithium chloride
LOH	Lost-of-heterozygosity
LPA	Lysophosphatidic acid
LVNI	Lymphatic, Neural or Vascular Invasion
M	Molar
mA	Milliamperes
MAPK	Mitogen-activated protein kinase
microRNA	Micro ribonucleic acid
min	Minute
mL	Millilitre
MLH1	Human MutL homolog 1
mM	Millimolar
MMR	Mismatch repair
MRI	Magnetic resonance imaging
MSH2	Human MutS protein homolog 2
MSH6	Human MutS protein homolog 6
mRNA	Messenger ribonucleic acid
MSI	Microsatellite Instability
MUTYH	MutY homolog
N	Number of samples
NA	Not applicable
ng	Nanogram
NRAS	Neuroblastoma RAS viral oncogene homolog
°C	Celsius degrees
OR	Odds Ratio
OS	Overall Survival
P25	25^{th} percentile
P75	75^{th} percentile
pEGFP	Plasmid EGFP
Pen/Strep	Penicillin-Streptomycin
PH domain	Pleckstrin homology domain
PIK3CA	Phosphatidylinositol 4,5-bisphosphate 3-kinase

PIP_2	Phosphatidylinositol 4,5-bisphosphate
PKC	Protein kinase C
PLC	Phosphoinositide-specific phospholipase C
***PLCE1*/PLCε**	Phosphoinositide-specific phospholipase C epsilon
***PLCG*/PLCγ**	Phosphoinositide-specific phospholipase C gama
PLCβ	Phosphoinositide-specific phospholipase C beta
PLCδ	Phosphoinositide-specific phospholipase C delta
PLCζ	Phosphoinositide-specific phospholipase C zeta
PLCη	Phosphoinositide-specific phospholipase C eta
PMAS	Positive margins after surgery
PMS2	Mismatch Repair Endonuclease PMS2
p-STAT3	Phosphorylated signal transducer and activator of transcription 3
PTEN	Phosphatase and tensin homolog
PTGS2	Prostaglandin-endoperoxide synthase 2
RA1/RA2	RAS association domain 1/2
RAL	RAS related protein RAL
RHO	RAS homolog gene family
Rn	Normalized reporter value
RNA	Ribonucleic acid
Rnase	Ribonuclease
Rpm	Revolutions per minute
RTK	Tyrosine kinase receptor
RT-qPCR	Real-time reverse transcriptase polymerase chain reaction
s	Second
SD	Standard deviation
SH2/SH3	Src homology 2/3
SMAD4	SMAD family protein 4
SNP	Single nucleotide polymorphism
SPSS	Statistical Package for the Social Sciences
STAT3	Signal transducer and activator of transcription 3
STK11	Serine/threonine kinase 11
TAC	Computed tomography scan
TBS	Tris-buffered saline
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride

***TGFA*/TGF α**	Transforming growth factor α
***TNF*/TNF-α**	Tumour necrosis factor alfa
TNM	Tumour extension/Lymph Nodes/Metastasis
***TP53*/P53**	Tumour protein 53
V	Volts
***VEGFA*/ VEGF-A**	Vascular endothelial growth factor A
VIM	Vimentin
WHO	World Health Organization
µCi/mL	Micro Curie per millilitre
µg	Microgram
µL	Microliter

Index

Index of Figures

Index of Tables

1. Introduction

Cancer is a worldwide problem whose incidence and mortality is extremely high in the world, affecting every aspect of our society[1].

The efforts made in prevention, earlier diagnoses and development of better treatments has contributed to decreased cancer mortality, however the numbers of new cases and deaths related to cancer are still alarming.

In 2004, the Global Burden Disease Report showed that cancer was the 3rd disease with highest mortality, only bellow cardiovascular diseases and infectious/parasitic diseases. This report estimated that 7.4 million deaths were related to cancer internationally[2].

In 2012, cancer incidence increased to 14.1 million globally[3] and it is estimated that 8.2 million deaths occurred due this disease[3], while in 2015 cancer caused death to 8.8 million people.

The Global Cancer Observatory (GCO) estimates that in 2018, cancer incidence will rise to 18.1 million new cases and cause the death of 9.6 million[4].

Unfortunately, in 2040 cancer incidence is expected to reach 29.5 million new cases and the global number of deaths is predicted to increase up to 16.4 million[4]. These numbers can be explain by an increase in lifetime expectancy, risk factors exposure, bad lifestyle habits (e.g. smoking), among others[5].

Overall, cancer is generically defined by the uncontrolled growth and spread of malignant cells to the surrounding tissues which can, ultimately, affect almost any part of the body[6].

There are different types of cancer depending on the organ or tissue where it is formed, the type of cells and their driver mechanisms, therefore, different approaches to treat this disease are desired.

1.1. Colorectal cancer

Colorectal cancer (CRC) is the 4th most diagnosed and the 3rd with highest mortality in the world.

In Portugal, CRC is the 3rd most diagnosed and the 2nd with highest mortality for both sexes and all ages.

GCO estimates that in 2018, 1.85 million new cases will be diagnosed, and 880 792 deaths will occur. From those, 10 270 new cases and 3 050 deaths will be registered in Portugal.

It is expected that the global burden of colorectal cancer will increase to more than 3.2 million new cases and 1.6 million deaths in 2040[4].

1.1.1. Colorectal cancer development

CRC develops on the large intestine, which is part of the gastrointestinal system[7] (Figure 1).

Colon compose the major part of the large intestine and is formed by 4 portions named ascending colon (connects large intestine with the small intestine), transverse colon (connects the right

and the left part of the colon), descending colon and sigmoid colon (connects the descending colon with rectum and finally with the anus)[7].

The right side of colon (proximal colon) is composed by the ascending and two thirds of the transverse colon which developed embryologically from the midgut. While left colon (distal colon) develops from the hindgut and is composed by one third of transverse colon, descending colon and sigmoid colon[7].

In 1978, Hill et al. proposed the adenoma-carcinoma sequence to describe the transformation of normal colorectal epithelium to an adenoma (polyp), proceeding to *in situ* carcinoma, and ultimately to an invasive and metastatic tumour (Figure 2)[8].

There are different types of polyps (Figure 2) and despite they are pre-cancerous not all turn into cancer. Nevertheless, due to the risk they represent of becoming malignant (Figure 2) they must be removed[9].

The wall of colon and rectum is made of many layers, CRC start spreading from the mucosa layer outwards potentially invading blood and/or lymphatic vessels. From this point, cancer cells can spread to other parts of the body (metastasize)[9].

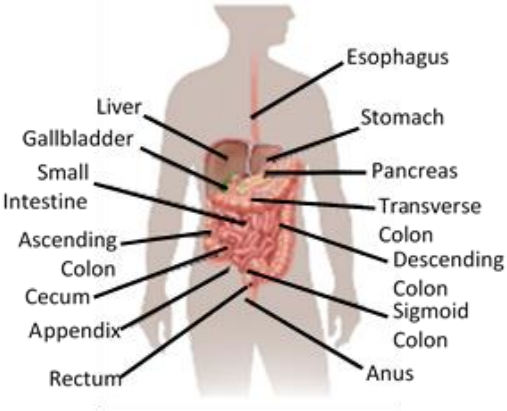


Figure 1 – Gastrointestinal System Anatomy.

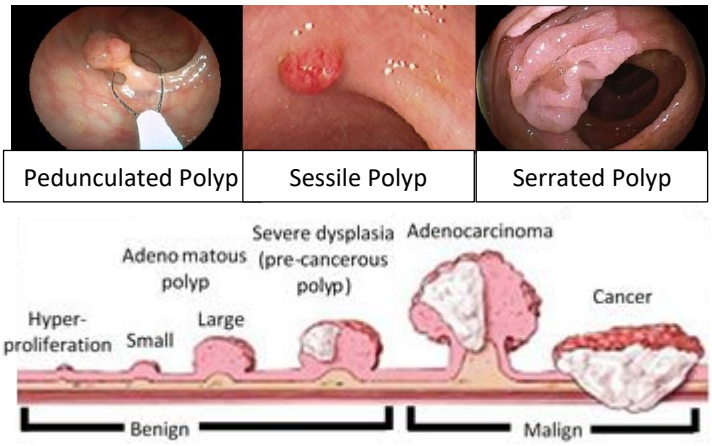


Figure 2 – Different types of polyps and how they turn into cancer.

1.1.2. Risk factors and drivers of colorectal cancer developing

Risk factors are characteristics, or substance exposure that increase the chances of a person develop a disease or injury. It is important to know which risk factors are associated to CRC development because although some are intrinsic, many can be avoided.

There are many factors pointed to increase the risk of developing this disease including overweight or obesity, excessive alcohol consumption, smoking tobacco, consumption of processed and red meat, inflammatory bowel diseases, and family history[10]. However, it is estimated that 90% of CRCs diagnosed develop sporadically and only 10% are caused by an inherited predisposition.

1.1.2.1 Sporadic colorectal cancer

Studying colorectal cancer, Fearon Vogelstein described in 1990 that accumulation of multiple mutations in epithelial cells were necessary for the acquisition of selective growth advantage[11]. Chromosomal Instability (CIN), microsatellite instability (MSI) and CpG island methylation pathway (CIMP) are now considered the three major genetic and epigenetic mechanisms responsible for sporadic CRC[12,13].

Chromosomal Instability is the most common sporadic CRC driver and it is defined as the accumulation of numerical or structural abnormalities in chromosomes which leads to lost-of-heterozygosity (LOH) in tumour suppressor loci and or chromosomal rearrangements. This allows the accumulation of mutations in critical genes like *APC, KRAS, PIK3CA, BRAF, SMAD4, TP53*, and others that activate pathways important to tumorigenesis[13].

Microsatellite Instability is caused by the abnormal number of microsatellites (short repeat sequences of DNA) in cells. The incapacity of cells to correct DNA damage caused by the silencing of DNA mismatch repair (MMR) genes leads to the accumulation of microsatellites. When these uncorrected microsatellites are in DNA coding regions, they can give rise to frameshift termination and consequently protein truncations[13].

CpG Island Methylation Pathway consist in the methylation of the genome in regions rich in cytosine (C) and guanine (G) bases that modulate DNA transcription. This modulation may silence important tumour suppressor genes and/or activate constitutively oncogenes[13].

1.1.2.2 Hereditary colorectal cancer

It is important to characterize CRC with hereditary predisposition to identify persons at risk, to provide earlier diagnosis and better therapeutic approaches[14].

Germline mutations are responsible for driving hereditary CRC with a manifestation and evolution of the disease well characterized in most of the cases. Germline mutation in *APC* gene leads to Familial Adenomatous Polyposis. Mutations in *MSH2*, *MSH6*, *MLH1* and *PMS2* genes lead to Lynch Syndrome. Biallelic mutations of *MUTYH* lead to MUTYH Associated Polyposis. *STK11* gene germline mutations lead to Peutz-Jeghers Syndrome, while *SMAD4* or *BMPR1A* lead to Juvenile Polyposis Syndrome[13,14].

1.1.3. Colorectal cancer screening and diagnosis guidelines

An earlier diagnosis of CRC is essential to achieve better outcomes for patients with this disease.

In this context, European Commission (EC) develop a set of guidelines to assure quality in CRC screening and diagnosis[15].

Despite the increasing list of methods to screen for CRC, to date only faecal occult blood test (FOBT) is recommended as screening test[15].

Some imaging techniques are being applied instead or to complement the screening of FOBT for example sigmoidoscopies and colonoscopies, due to their potential impact to prevent CRC development on-site[15].

A positive colorectal cancer must be confirmed pathologically. After a biopsy or surgery, pathologists must report every available tissue characteristic.

In order to do that, the European Society for Medical Oncology (ESMO) developed guidelines underlining the information that have a great impact on patient's prognosis.

Therefore, pathologists must report[16–18]:

a) Morphological description of the specimen;
b) Surgical procedure carried out;
c) Definition of tumour site and size;
d) Presence or absence of macroscopic tumour perforation;
e) Histological type and grade;
f) Distance of cancer from resected margins (proximal, distal and radial);
g) Presence or absence of tumour deposits;
h) Lymphovascular and/or perineural invasion;
i) Presence of tumour budding;
j) Stage TNM by reporting the extension of tumour in the bowel wall (T), the number of lymph nodes removed and how many are invaded by cancer cell (N), and finally the involvement of other organs (M).

1.1.4 Colorectal cancer treatment

To develop the best treatment, it is important to consider the extension of the disease and patient's risk to relapse.

While stage report is crucial to determine the extension of the disease, other status like involvement of resected margins, histological grade, perforation, lymphovacular and/or perineural invasion inform the probability of disease relapse[16–18].

Other parameters may also represent a high-risk condition such as mutations and altered expression of genes such as *TP53*, *KRAS*, *BCL2*, *TGFA*, *EGFR*, proliferation index and aneuploidy[18].

When developing a strategy of treatment, it is also important to consider individual patient's characteristics like physical condition and age. It is important to assess whether patients have capability to get through treatment and consider if the treatment will benefit the patient in the lifetime left.

Normally, in earlier stages of locoregional disease (Stage I) surgery is the only treatment applied[16,18]. In intermediate stages of the disease (Stage II and III) combined chemotherapy (e.g. FOLFOX – Fluoropyrimidine plus oxaliplatin) is the first line treatment after surgery (in rectal cancer chemotherapy can be combined with radiotherapy). Rectal cancers in stage III have the option to be treated with neoadjuvant therapy to diminished the size of tumour before surgery[16].

Finally, in stage IV the treatment will depend on tumour related characteristics (e.g. local of metastasis, symptoms, prognostic molecular or biochemical biomarkers), patient related factors (e.g. patient fitness condition) and treatment related factors (e.g. toxicity)[17]. Commonly, a first line treatment (e.g. FOLFIRI, FOLFOX normally in combination with cetuximab or bevacizumab) is used to shrink the tumour and metastasis and allow resection or ablation. After surgery, patients continue treatment with the first line therapy.

However, not all patients are fit to go under tumour resection, and some of them relapse after finishing the adjuvant therapy cycle. In these cases, patients may initiate a second line therapy that must be different from the first line treatment. Every time a patient relapses it is necessary to change the therapeutic strategy. There is also the possibility to include these patients in clinical trials to test new therapeutic approaches[17].

1.1.5 Colorectal cancer biomarkers

A cancer biomarker is any molecule (e.g. protein, DNA, RNA, etc.) or tumour cell that can be found in body fluids or tissue and signalize an abnormal process. Therefore, cancer biomarkers are useful to predict cancer development, prognosis or therapeutic response[19].

Table 1 shows the biomarkers recommended by ESMO and/or by American Society for Clinical Pathology and their application on the clinic[16–18,20].

Table 1 – Current biomarkers recommended by ESMO and American Society for Clinical Pathology.

Biomarker	Application	Location
FOBT	Diagnostic	All locations
CEA	Prognosis	All locations
***BRAF* mutation V600E**	Prognosis	Right colon
MSH2, MSH6, PMS2* and *MLH1	Prognosis	Right colon
***NRAS* and *KRAS* mutations**	Prognosis and therapeutic response	All locations

As previously mentioned, FOBT is the only biomarker used in CRC diagnostic[15]. Carcinoembryonic antigen (CEA) is used to monitor patient's outcome and predict CRC relapses, high levels of CEA are correlated with poor prognosis[16–18,20]. *BRAF* mutation V600E alongside with mutations on *MSH2, MSH6, PMS2* and *MLH1* are correlated with poor prognosis in right colon cancer[17,20]. Finally, *NRAS* and *KRAS* mutations were found to be correlated with poor prognosis and with resistance to anti-EGFR therapy[17,20].

Furthermore, mutations in *PIK3CA*, loss of PTEN, the levels of EGFR ligands such as amphiregulin, epiregulin and TGF-α, the levels of EGFR protein expression, amplification of EGFR and HER2 are seen by ESMO as emerging biomarkers for anti-EGFR therapeutic resistance. However, they do not have, so far, the recommendations necessary to be used by clinicians in a regular basis[20].

In CRC is also established that right and left side colon cancers have different prognosis because of their different embryological origins, as well as anatomical, histological, genetic and immunological characteristics[7,21].

Other promising biomarkers in CRC follow-up are circulating tumour cells (CTC) and circulating tumour DNA (ctDNA). In patients with CRC it is possible to find CTC and ctDNA in the blood stream. An elevated number of this cells and/or elevated levels of ctDNA is associated with the risk of developing metastasis and/or resistance to therapeutic[22].

Although several putative biomarkers have been identified, very few are effectively used in clinical practice, therefore, there is still an urgent need for highly selective and specific biomarkers able to allow the earlier detection of tumour cells.

1.2. Phosphoinositide-specific phospholipase C

Phosphoinositide-specific phospholipase C (PLC) proteins are putative cancer biomarkers that are found in every human cell and are involved in a broad range of regulatory interactions such as cell motility, growth and differentiation[23,24].

PLCs are stimulated by specific receptors for hormones, neurotransmitters, antigens, components of the extracellular matrix and growth factors[25].

Once stimulated, PLC cleaves phosphatidylinositol 4,5-bisphosphate (PIP_2) generating two second messengers: diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP_3)[25,26].

While IP_3 is important in cellular Ca^{2+} regulation, DAG activates protein kinase C (PKC) and its downstream pathway. Together, these second messengers are important regulatory factors of a variety of biological functions as cell motility, growth, survival, fertilisation and sensory transduction[25,26].

13 PLC isozymes were identified and grouped in 6 families (PLCβ, PLCγ, PLCε, PLCδ, PLCη and PLCζ) (Figure 3). All these enzymes share common core of domains but have specific structures and regulation mechanisms[26].

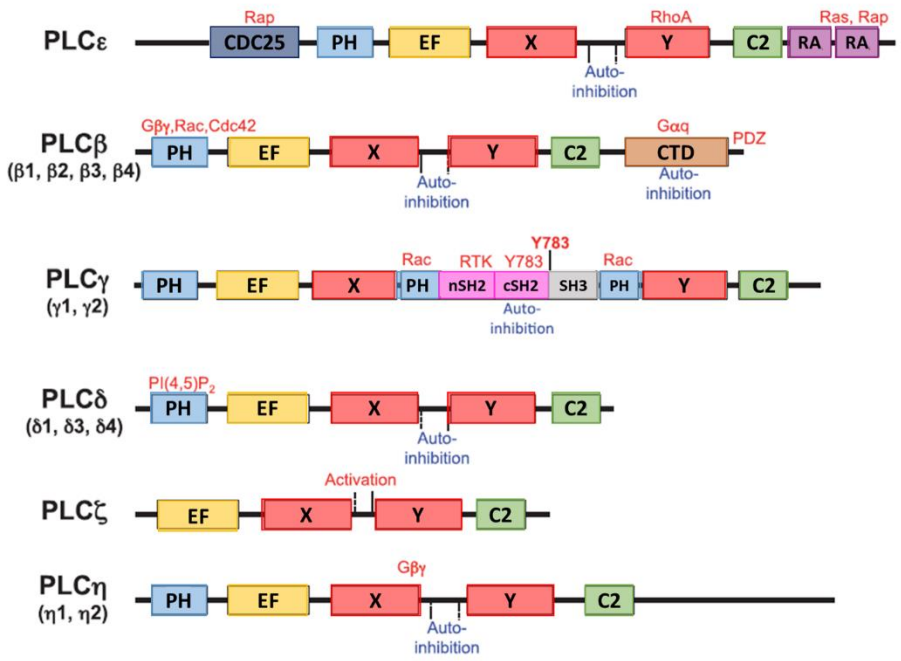


Figure 3 – Phosphoinositide-specific phospholipase C family domain organization: Domain organization of PLCε, PLCβ, PLCγ, PLCδ, PLCη and PLCζ. PLC common domains C2 (green), catalytic TIM barrel domain X-Y (red), EF-hands domain (yellow) and pleckstrin homology domain (light blue, except for PLC ζ). PLCε specific domains: CDC25 (dark blue), RAS association domain (purple). PLCβ specific domain CTD (brown). PLCγ specific domains: nSH2 and cSH2 (pink) and SH3 (gray).

PLCβ has 4 isoforms (PLCβ1, PLCβ2, PLCβ3 and PLCβ4) with different patterns of tissue distribution. While PLCβ1 and PLCβ3 are expressed in many tissues, PLCβ2 and PLCβ4 are mainly expressed in hematopoietic and neuronal cells, respectively[25–27]. Moreover, PLCβ have a GTPase activating protein function and its specific C terminal domain (CTD) play an important role in membrane binding and activation[26].

Defects in PLCβ3 has been associated with atopic dermatitis like skin inflammation and with myeloproliferative neoplasms, while PLCβ1 deficiency has been associated with myelodysplastic syndromes[26].

PLCγ has 2 isoforms (PLCγ1 and PLCγ2), while PLCγ1 is ubiquitously expressed, PLCγ2 is predominantly expressed in hematopoietic cells[25,27,28].

PLCγ has an insert between the catalytic domains X and Y constituted by two PH domains, two SH2 domains and a SH3 domain. These domains play a crucial role in PLCγ auto-inhibition and activation downstream receptors tyrosine kinases (RTK) such as epithelial growth factor receptor (EGFR)[25,27,28]. Furthermore, PLCγ is important for hematopoietic cell development, especially in immune system cells and disfunctions of PLCγ may contribute to development of auto-immune disease[27,28].

PLCγ enzymes also contribute to some oncogenic signalling pathways as signal transduction downstream of RTK. The overexpression of this enzyme has been related to breast and colon cancers

development. Driver mutations in *PLCG* genes were identified in angiosarcoma, cutaneous T cell lymphoma (CTCL) and chronic lymphocytic leukaemia (CLL)[28].

PLCδ has 4 isoforms (PLCδ1, PLCδ2, PLCδ3 and PLCδ4) and it has a broad distribution in all tissues. It is speculated that this enzyme may be the prototype of the others PLC. PLCδ only present the PLC conserved domains and do not have any unique domain. Deficiency in these enzymes may be involved in male infertility and the development of inflammatory skin diseases such as psoriasis and contact hypersensitivity[27].

PLCζ is sperm specific, its structure lacks the PH domain and disfunctions on this enzyme leads to infertility[27].

PLCη has 2 isoforms (PLCη1 and PLCη2), these isoenzymes are mainly expressed in brain and its physiological functions still unknown[27].

1.2.1. Phospholipase C epsilon (PLCε)

PLCε is broadly expressed in all tissues but has its higher expression in heart, colon and lung[29].

This enzyme has unique domains such a CDC25 domain in the N-terminal region and two RAS-Association domains (RA1 and RA2) in the C-terminal region[25–27,30–33].

While the CDC25 domain works as a guanine exchange factor (GEF) for RAP1, the association of RAS proteins with the RA2 domain is important for the translocation of this enzyme from cytoplasm to the plasma membrane where it exerts its function. Other families of small GTPases able to activate PLCε are RAL and RHO families through binding to the catalytic domain of PLCε[29].

Furthermore, PLCε activity can also be stimulated by lysophosphatidic acid (LPA) and thrombin receptors through the G-protein subunit Gα12/13[25–27,30–33].

PLCε has been shown to be crucial in the regulation of normal physiology in animal models. For example, mouse models depleted of PLCε developed heart hypertrophy in response to chronic cardiac stress. Interestingly, this enzyme has been associated with idiopathic dilated cardiomyopathy and may be associated with the development of other human cardiac diseases[34,35].

In another study, zebrafishes knockdown for PLCε developed nephrotic syndrome, since this enzyme is important to maintain the podocyte filtration barrier[30,36].

This enzyme was also associated with skin inflammation, neuroinflammation and cancer[26].

1.2.1.1. Phospholipase C epsilon and cancer

PLCε was shown to be involved in cancer development. This enzyme may participate with its phospholipase activity in many pathways associated with cancer (for example cell proliferation). On the other hand, its GEF and RAS association properties may also influence RAS/MAPK pathways[30,34,37,38].

Many studies suggest that PLCε deregulation may contribute to tumorigenesis by inducing inflammation and angiogenesis which facilitates cancer development and maintenance in skin, prostate, bladder and in intestine[39–44].

However, the role of PLCε in cancer development and progression is still controversial. For example, in skin cancer studies from Bai and co-workers and Martins et al. showed contradictory results for PLCε.

Bai and co-workers report that transgenic mice with PLCε catalytic function deleted had a delay on time of tumour development. After a two-stage chemical skin carcinogenesis protocol these mice had less tumour burden than mice with full-length PLCε. These findings suggested that PLCε has an oncogene function[45].

On the other side, Martins et al. show that PLCε behave as a tumour suppressor. Since mice knockout for PLCε had an increased tumour load, bigger tumours and an increased number of Ki67 positive cells before and after treatment in the two-stage chemical skin carcinogenesis model[46]. Thus, the controversial role of PLCε in this skin mouse model can be, at least, explained by the different nature of the transgenic animal, nevertheless more studies are needed in order to clear PLCε role in skin cancer.

However, in esophageal and gastric cancer results seem to be consensual in showing that high levels of PLCε expression were associated with tumour progression, which implicates an oncogene role of this enzyme[47–49].

In colorectal cancer, PLCε is considered a tumour suppressor protein. Danielsen et al. showed that mRNA levels of *PLCE1* are downregulated in tumour tissue when compared with normal tissue, they also found that *PLCE1* levels were associated with *KRAS* mutation[49,50]. In a separate study, Wang X. et al. also found that PLCε expression levels were downregulated when compared with colon normal tissue samples. Furthermore, they found that PLCε overexpression lead to higher apoptosis rates, slower growth and decreased migration ability in cells. PLCε overexpression also formed smaller tumours in xenograft mice[49,51].

Therefore, it is overall agreed that depending on the type of tumour, PLCε may have a tumour suppressor role or act as an oncogene[49] (Figure 4).

Finally, there are also many studies developed in digestive tract cancers that reports the presence of *PLCE1* rs2274223 single nucleotide polymorphism (SNP) as a potential risk factor for cancer development.

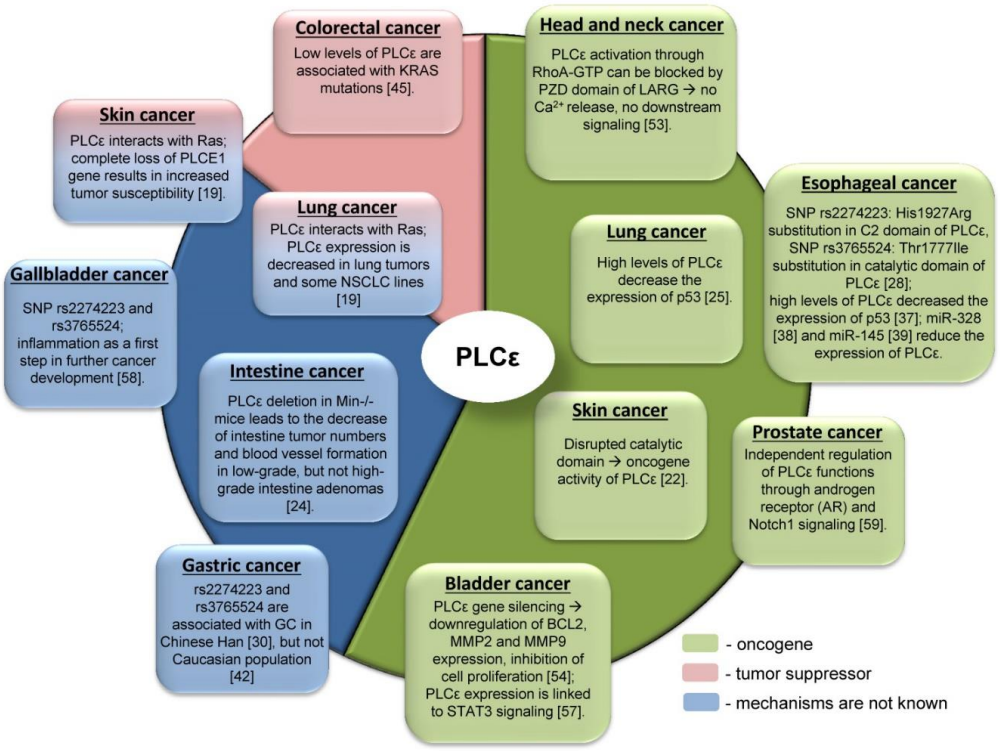


Figure 4 – PLCε controversial role in cancer. Figure from Tyutyunnykova A et al. The controversial role of phospholipase C epsilon (PLCε) in cancer development and progression, 2017[49].

1.2.1.2. Single nucleotide polymorphism rs2274223 in cancer

Single nucleotide polymorphism rs2274223 consists in the substitution of an adenine by a guanine in the position 5781 of human cDNA from *PLCE1* gene, which leads to the substitution of an histidine by an arginine in the 1927th amino-acid from PLCε, localized on the C2 domain of the enzyme[52].

This polymorphism was discovered by Abnet et al.[53] in 2010 in a genome-wide association study (GWAS), when studying esophageal cancer.

This SNP was also studied for gastric cancer (GC) by two different groups but with contrary results. While Li-Dong Wang et al.[54] report in 2010 that rs2274223 is associated with GC development in a Chinese Han population, Palmer et al.[55] could not find any association between this SNP and GC in a Polish population.

Furthermore, this polymorphism was also associated with gallbladder cancer in a North Indian population[56].

There were some retrospective studies that identified an association between rs2274223 and CRC development. In 2012, Fen-Xia Li et al.[57] reported that while the AG genotype had a malignant effect in CRC, the genotype GG seemed to have a protector role. Furthermore, in 2014, Qi Wang et al.[52] showed that in a Chinese population the phenotypes AG an GG were associate with increased risk of developing CRC, they further show that mRNA levels of the polymorphic allele *PLCE1* were

downregulated when compared to the wild-type mRNA. This could potentially lead to a reduction of PLCε enzyme and consequently reduction of its tumour suppressor role, exposing patients to increased risk of development of this disease.

In 2015, Juozas Kupcinskas et al.[58] develop the first study in a European population that evaluated the association of this SNP with CRC, the results obtained by them show no significant association between *PLCE1* SNP rs2274223 and CRC.

However, contrary to those results, in a case-control study developed by Oztas Ezgi et al.[59] in a Turkish and Caucasian population it was found an association between *PLCE1* SNP rs2274223 and CRC.

Overall, the association between SNP rs2274223 and CRC is not consensual and highly inconclusive.

2. Objectives

Many studies have explored the association between SNP rs2274223 and CRC. However, it is still unclear whether this polymorphism is related with CRC.

Interesting, most studies were developed in Chinese populations. The only two studies that focused on European population had contradictory results. Furthermore, all these studies had focused only in the risk of developing CRC leaving unanswered the question of whether the outcome of patients with CRC is affected by the presence of this polymorphism. Finally, the functional role of this polymorphism in cancer cells was never explored before.

Therefore, with this project we aimed at studying this polymorphism, not only, in a clinical setting, but also, in an *in vitro* approach.

Our specific objectives are:

i. In a clinical setting:
   a) Access the risk of *PLCE1* SNP rs2274223 in developing CRC in a Portuguese population;
   b) Determination of polymorphism rs2274223 prognostic value.

ii. In an *in vitro* approach:
   a) Determine the influence of this polymorphism in the phospholipase activity;
   b) Determine how this polymorphism affects cells properties.

3. Materials and Methods

3.1 Study Population

This was a single center case-control study which enrolled 219 patients newly diagnosed with CRC and confirmed by histopathology analysis from Hospital de Santa Maria, Centro Hospitalar Lisboa Norte, in Lisbon, Portugal. Patient's blood samples were stored at -80°C, at L. Costa Laboratory Biobank.

We retrospectively collected patient's information, namely, age at diagnosis, sex, follow-up time, tumour location, pathological tumour stage and tumour differentiation status accordingly to WHO classification of tumours of the digestive system[60], positive margins after surgery (PMAS), obstruction or perforation at diagnosis, lymphatic, neural or vascular invasion (LNVI) and *KRAS* mutations (only for Stage IV Patients). Data collection occurred between April 2006 to December 2017 and was in agreement with data protection principle.

As healthy controls, 221 samples were requested from Biobanco-IMM, Lisbon Academic Medical Centre, Lisbon, Portugal, which were matched for sex and age.

All control participants and patients agreed to make available their biological material for research purposes by filling and signing an Informed Consent.

3.2. Genomic DNA Extraction and Genotyping

Patient's blood samples were collected in ethylenediaminetetra-acetic acid (EDTA) anticoagulant tubes and stored at -80°C until genomic DNA (gDNA) extraction. 200µL of blood was used to extract gDNA following NZY Blood gDNA Isolation Kit (NZYTech) recommended protocol and stored at -20°C until analysis.

The rs2274223 SNP was genotyped using the TaqMan® SNP Genotyping Assay (Reference: C___1947254_1; Applied Biosystems). For allele A probes were marked with VIC, while for allele G probes were marked with FAM fluorescence dyes.

Patient's and control's gDNA samples were placed in MicroAmp™ Fast Optical 96-Well Reaction Plates following the Wet DNA Delivery Method of Taqman® Genotyping Master Mix Protocol in a 10µL reaction: 5µL TaqMan Genotyping Master Mix (2X), 0,5µL TaqMan genotyping assay mix (20X) and 4,5µL of sample (containing a gDNA concentration between 1 to 20ng/µL).

The PCR amplification was carried out with an initial pre-denaturation step at 60°C for 1 min and a 10 min denaturation at 95°C followed by 40 cycles of denaturation at 95°C for 15 s and annealing/extension at 60°C for 1 min, using an Applied Biosystems® 7500 Fast Real-Time PCR System according to the manufacturer's instructions. The 7500 Fast Software v2.0.6 (Applied Biosystems) was used for data analysis.

Allelic discrimination was measured by ΔRn using ROX as passive reference.

3.3. Mutagenesis

In order to test the role of this polymorphism *in vitro,* rat *PLCE1* cloned into pEGFP-C1 vector (Clontech) were obtained from Matilda Katan´s Lab at University College London. Site-directed mutagenesis was then performed in order to alter the coding of PLCε from arginine to histidine in position 1927. Mutagenesis was performed following the NZYMutagenesis kit (NZYTECH) recommended protocol using the complementary primers (forward and reverse) designed and provided by Invitrogen (Thermofisher) bellow:

Primer Forward (5'to3'):

GGAATGAGCAGTTTCTCTTCC**AC**GTTCACTTTGAAGATCTTG

Primer Reverse(5'to3'):

CAAGATCTTCAAAGTGAAC**GT**GGAAGAGAAACTGCTCATTCC

PCR was performed accordingly to manufacturer indications, in the reaction mix for mutagenesis 1µL of *PLCE1* plasmid was used and 2µL of each primer (10µM each). After PCR, template DNA was digested with DpnI restriction enzyme PCR product was transformed into NZYStar competent cells. Finally, DNA was isolated from different kanamycin resistant colonies using NZYMiniprep (NZYTech) and confirmed by sequencing.

3.4. Cell Culture

Cos-7 (Cercopithecus aethiops immortalized kidney fibroblast cell line), DLD1 and HCT116 human colon cancer cell lines were purchased from American Type Culture Collection (ATCC, Virginia, USA) and cultured in fresh complete growth medium constitute by Dubelcco's Modified Eagle's Medium (DMEM; Gibco) supplemented with 10% of Fetal Bovine Serum (FBS; Gibco) and 1% of Penicillin-Streptomycin (Pen/Strep; Gibco). Cells were grown at 37°C with 5% of CO2.

Cos-7, DLD1 and HCT116 cell lines were transfected when were at 60-70% confluence with Lipofectamine® 3000 (Thermofisher), using pEGFP-C1-*PLCE1* wild-type and mutant (H1927A) vectors, following manufacture instructions. Briefly, Lipofectamine® 3000 reagent (1µL), plasmid DNA (1µg) and P3000 reagent (2µL) were diluted in Opti-MEM® Medium (Gibco) and incubated for 20min at room temperature to create DNA-lipid complexes that were then added to 24-well plate cells in medium without serum or antibiotics. Volumes were adjusted for other format plates. For parental control, cells were incubated with Lipofectamine®3000 reagents for the same time. Roughly, 3 hours after initial transfection, the medium was replaced by fresh complete growth medium.

3.5. Cell Viability Assay

To perform the viability test 3 x 10^4 DLD1 and HCT116 cells were culture in 24-well plates with DMEM supplemented with 10% FBS and 1% Pen/Strep for 24 hours at 37°C until they were at 60-70% confluence. At that point, cells were transfected using Lipofectamine® 3000, as previously described. 24 hours after transfection, medium was changed by 1mL of fresh complete growth medium supplemented with 10% alamarBlue® from Invitrogen (Thermofisher). After 3 hours of incubation fluorescence was read with an excitation wavelength of 570nm and an emission length of 585nm in the Infinite 200 Microplate Reader (Tecan).

This last step of the procedure was repeated for 3 days.

3.6. Activity Assay

To perform the activity assay 2,5 x 10^5 Cos7 cells were seeded into a 6 well-plate with DMEM supplemented with 10% FBS and 1% Pen/Strep for 24 hours at 37°C. At that point, cells were transfected using Lipofectamine® 3000 and plasmids coding for PLCε wild-type and mutant (H1927A) as previously described.

48 hours after seeding, cells were washed twice with inositol-free DMEM (USBiological) without serum and incubated for 24h in 1,5ml of the same medium supplemented with 0,25% fatty acid free bovine serum albumin (BSA, Sigma) and 1,5µCi/ml myo-[2-3H]inositol (PerkinElmer Life Sciences). At that time, cells were incubated in 1,2ml of inositol-free DMEM without serum containing 20mM LiCl (Sigma) with or without stimulation with 100 ng/ml EGF (Calbiochem®) for 1 hour.

To lyse cells was added 1,2ml 4,5% perchloric acid (Fluka) and incubated on ice for 30min. Samples were centrifuged for 20min at 4000rpm, supernatants were separated from pellets, neutralized by addition of 3ml of 0,5M potassium hydroxide/9mM sodium tetraborate (both from Sigma) and centrifuged for 20min at 4000rpm. After that, Anion exchange AG1-X8 200–400 columns (Bio-Rad) was converted to the formate form by addition of 2M ammonium formate/0,1M formic acid (both from Sigma). These columns were equilibrated with water and supernatants were loaded. The columns were washed three times with 5ml of 60mM ammonium formate/5mM sodium tetraborate, and inositol phosphates were eluted with 5ml of 1,2M ammonium formate/0,1M formic acid. 5ml Ultima-Flo scintillation fluid (PerkinElmer Life Sciences) was added to the eluates and the radioactivity quantified by liquid scintillation counting. The values represent total inositol phosphates. The pellets from the first centrifugation were resuspended in 100µl of water and 375µl of chloroform/methanol/HCL (200:100:15) (chloroform and HCL from Sigma and Methanol from Merck), plus an additional 125µl of chloroform and 125µl of 0,1M HCL. After vortexing, the samples were centrifuged at 2500rpm for 10min. 20µl of the lower phase were placed in a scintillation vial with 2ml of Ultima-Flo scintillation fluid and the radioactivity quantified by liquid scintillation counting. The

obtained values correspond to radioactivity in total inositol lipids. PLC activity is expressed as the total inositol phosphates formed relative to the amount of [^{3}H]myo-inositol in the phospholipid pool.

3.7. Western Blot

To perform Western blotting 3 x 10^6 DLD1 and HCT116 cells were cultured in Petri dishes with DMEM supplemented with 10% FBS and 1% Pen/Strep for 24 hours at 37°C until they were at 60-70% confluence. At that point, cells were transfected using Lipofectamine® 3000.

48 hours after transfection, HCT116 and DLD1 cell lines were lysed in lysis buffer (25mM Tris 1M pH 7,5 (Sigma), 2mM EDTA 0,5M (Sigma), 10mM 10% Triton (VWR), 1mM TCEP 1M (Sigma), protease inhibitor cocktail (Roche) and phosphatase inhibitor cocktail 2 (Sigma)) for 10min on ice.

Cells with lysis buffer were incubated for 10min on ice, centrifuged by 10min at 12000rpm at 4°C, and the supernatant transferred to a new tube to quantify proteins concentration. Lysates were centrifuged for 10min at 12000rpm at 4°C, and pellets removed. Protein quantification was performed using Bradford Reagent (Bio-rad) following manufacturer's instructions. Briefly, a BSA standard curve was initially performed with 0; 0,125; 0,25; 0,5; 0,75; 1; 1,5 and 2 mg/mL concentrations. Absorbance was measured in a microplate reader at 595nm. Absorbance of samples was read at the same wavelength and concentration extrapolated by BSA standard curve.

For protein dry transfer, proteins were transferred for nitrocellulose membranes using iBlot™ 2 Gel Transfer Device from Invitrogen (Thermofisher) for 10min (1min at 20V, 4min ate 23V and 5min at 25V).

For wet transfer, proteins were transferred to a nitrocellulose membrane overnight at 40mA and at 4°C.

After the transference, the nitrocellulose membrane was blocked with 5% non-fat milk in TBS 0,1% Tween20, or in 5% BSA in TBS 0,1% Tween20 (depending on antibody dilution buffer) for 1 hour, and then incubated overnight at 4°C with specific primary antibodies (Table 2).

Table 2 – List of antibodies used in Western Blot.

Antibody	Dilution	Dilution Buffer	Source
Anti-GFP Mouse Monoclonal Antibody	1:1000	5% w/v Milk, 1X TBS, 0.1% Tween	Sigma-Aldrich Anti-GFP
Anti-β-actin Mouse Monoclonal Antibody	1:5000	5% w/v Milk, 1X TBS, 0.1% Tween	Abcam Anti-beta Actin antibody [mAbcam 8226] (ab8226)
Anti-STAT3 (124H6) Rabbit Monoclonal Antibody	1:1000	5% w/v Milk, 1X TBS, 0.1% Tween	Cell signaling® STAT3 (124H6) Mouse mAb 9139
Anti-phospho-STAT3 (Ser727) Rabbit Monoclonal Antibody	1:1000	5% w/v BSA, 1X TBS, 0.1% Tween	Abcam Anti-STAT3 (phospho S727) antibody [E121-31] (ab32143)

Next day, after wash with TBS 0,1% Tween20, the nitrocellulose membrane was incubated for 1 hour at room temperature with horseradish peroxidase-conjugated (HRP) specific secondary antibodies.

Proteins were detected using Amersham ECL Western Blotting Detection Reagent (GE Healthcare Life Sciences), according to manufacturer's instructions. Signal was detected on radiographic film (Fujifilm), using Curix60 (AGFA).

3.8. Real-time Reverse Transcriptase Polymerase Chain Reaction (RT-qPCR)

To perform RT-qPCR, 5 x 10^5 DLD1 and HCT116 cells were cultured in 6-well plates with DMEM supplemented with 10% FBS and 1% Pen/Strep for 24 hours at 37°C until they were at 60-70% confluence. At that point, cells were transfected using Lipofectamine®3000, as previously described.

For total RNA extraction, cells were lysed 48 hours after transfection following NZY Total RNA Isolation Kit protocol (NZYTech) and RNA was quantified using NanoDrop 2000™ (Thermo Fisher Scientific).

After total RNA extraction, cDNA synthesis was performed following the protocol of NZY M-MuLV First-Strand cDNA Synthesis Kit, separate oligos from NZYTech using Oligo$(dT)_{18}$.

Finally, NZY RNase H (E. coli) was added to the samples to degrade the RNA that was not converted in cDNA.

After cDNA synthesis, it was performed the qPCR. Samples were analysed in triplicates in Corbett Rotor-Gene 6000 (QIAGEN Rotor-Gene Q). To each reaction, Sybr Green (NZY qPCR Green Master Mix (2x); NZYTECH), nuclease-free water, and primers were mixed with cDNA from samples into qPCR tubes from Qiagen.

Primers were purchased from Invitrogen (Thermo Fisher), except the specific Mouse *PLCE1* primer (PPM28139A-200) which is from QIAGEN. The Table 3 is a list of the primers used in qPCR, this list also shows the primers forward and reverse sequences used. Relative mRNA expression levels were normalized to endogenous GAPDH and calculated using the $2^{-\Delta\Delta CT}$ method.

Table 3 – List of primers used in qPCR (all primers in this list are for human cDNA).

Primer Name	Forward Sequence (5'→3')	Reverse Sequence (5'→3')
GAPDH	AAC ATC ATC CCT GCC TCT ACT G	TTG ACA AAG TGG TCG TTG AGG
***CDH1* (E-cadherin)**	GGA AAC TCT CTC GGT CCA	GCC CCT CAA GTG TTA CCT CAA
***CDH2* (N-cadherin)**	GCC CCT CAA GTG TTA CCT CAA	AGC CGA GTG ATG GTC CAA TTT
***VIM* (Vimentin)**	GAA AAC ACC CTG CAA TCT T	CCT GGA TTT CCT CTT CGT G
***TWIST1* (TWIST)**	CCG GAG ACC TAG ATG TCA TTG	CCA CGC CCT GTT TCT TTG
VEGFA	GGA GGA GGG CAG AAT CAT CAC	GGT CTC GAT TGG ATG GCA GT
***IL1B* (IL-1β)**	GCC CTA AAC AGA TGA AGT GCT C	GAA CCA GCA TCT TCC TCA G
***IL6* (IL-6)**	TAC CCC CAG GAG AAG ATT CC	TTT TCT GCC AGT GCC TCT TT
***TNF* (TNF-α)**	TCA GCC TCT TCT CCT TCC TG	GCC AGA GGG CTG ATT AGA GA
***PTGS2* (COX-2)**	CCC AGG GCT CAA ACA TGA TG	GTC TAG CCA GAG TTT CAC CG

***CXCL1* (CXCL-1)**	GTC CGT GGC CAC TGA ACT	GGG GAT GCA GGA TTG AGG C
***CXCL2* (CXCL-2)**	GCA GGG AAT TCA CCT CAA GA	GGA TTT GCC ATT TTT CAG CA

3.9. Statistical Analysis

Statistical analysis was performed in IBM® SPSS® software package, version 24.0 (IBM Corporation, New York, United States). Descriptive statistics were tabulated according to *PLCE1* genotype. For continuous variables (e.g. follow-up months), median and interquartile range were calculated. On the other hand, for categorical variables, proportions were calculated and differences were tested using Chi-Square test.

To calculate the odds of developing colorectal cancer according to *PLCE1* genotype (categorical nominal variable), the odds ratio was calculated with respective confidence interval (95%), and its significance was accessed by Chi-Square test. A logistic regression was further performed to adjust the effect estimate (odds ratio) to known risk factors for CRC, as age (categorical ordinal variable) and sex (categorical nominal variable).

Given the unfavourable prognostic implications and biological differences of metastatic cancer, survival outcomes were reported in two independent groups defined as a function of the stage at time of diagnosis (I-III and IV) and patients were analysed in two arms (AA/AG vs. GG) in univariate and multivariate analysis.

Overall survival (OS) was calculated using the Kaplan-Meier method and graphical representations of survival outcomes were reported using Kaplan-Meier graphs. Survival curves of both arms were analysed by log rank test. To explore the prognostic implications of *PLCE1* genotype (categorical nominal variable), a univariate and multivariate Cox proportional hazards model was built controlling for known prognostic patients and tumour characteristics, as age (categorical ordinal variable), stage, grade, tumour location, obstruction and perforation, vascular, lymphatic and or neural invasion and the presence of unclear margins at time of diagnosis (categorical nominal variables). In stage IV patients, the multivariate analysis included tumour location, the location of organ metastasis (categorical nominal variables) and age. Median follow-up was balanced throughout the *PLCE1* genotypes.

To analyse significant differences in the PLCε Activity Assay and in the relative expression of genes *mPLCE1*, *hCDH1*, *hCDH2*, *hVIM*, *hTWIST*, *hVEGFA*, *hIL1B, hIL6*, *hTNF*, *hPTGS2*, *hCXCL1* and *hCXCL2* between controls and the polymorphisms AA and GG we performed one-way ANOVA. To determine significant differences in the Cell Viability Assay we performed two-way ANOVA. Tukey Test were performed to determine which conditions had significative differences.

4. Results and Discussion

4.1. Role of *PLCE1* rs2274223 polymorphism in the risk of developing colorectal cancer

Most of the studies about *PLCE1* rs2274223 polymorphism approach the importance of this SNP in the risk of developing colorectal cancer. However, there are contradictory results reported[52,57–59].

In order to unveil the importance of this polymorphism in the risk of developing colorectal cancer, we enrolled 218 patients with colorectal cancer and 221 healthy controls that were matched by sex and age (Table 4).

Table 4 – Demographical characteristics of patients with CRC and healthy controls enrolled in this study. Differences between groups were accessed using independent T-test for Age (Years) and Chi-square test for Age (Rank years), Sex. N represents the number of samples, ±SD represents age standard deviation. p-value were calculated using 95% confidence interval.

		Patients	Controls	p-value
Number of Samples		218	221	/
Age (Years)	**Mean**	68,78	66,26	0,011
	±SD	11,381	8,437	
Age (Rank years)	**<65**	69 (31,7%)	87 (39,4%)	0,110
N (%)	**≥65**	149 (68,3%)	134 (60,6%)	
Sex	**Male**	131 (60,1%)	129 (58,4%)	0,771
N (%)	**Female**	89 (39,9%)	92 (41,6%)	

Genomic DNA was extracted from patient's blood sample and genomic DNA from healthy controls was provided to us from Biobanco-IMM. Genotyping was assessed by real-time PCR and the allelic discrimination of patients vs. control samples were plotted considering ΔRn values for allele A and for allele G (Figure 5A and 5B).

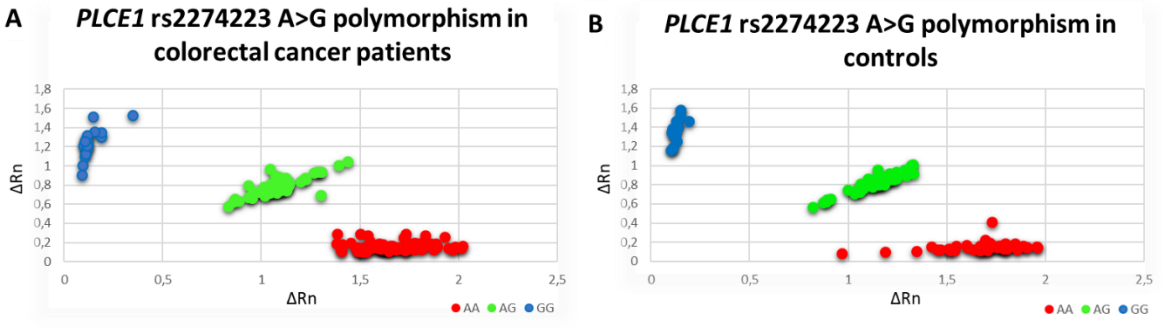


Figure 5 – Allelic discrimination of *PLCE1* rs2274223 A>G polymorphism. Plots were drawn for CRC patients (A) and controls (B) according to ΔRn of allele A (VIC) and allele G (FAM). ROX was the reference dye.

These results allowed to determine the SNP distribution in patients with CRC and healthy controls (Table 5), which was not significantly different, both in the univariate (p-value= 0,882; χ^2-test of the difference between the three genotype groups), and multivariate analysis controlling for age and sex (p-value=0,839; Logistic Regression) (Table 5). Thus, genotype homozygotic for allele A was seen in about 45% of patients and 43% of controls, whereas the heterozygote genotype was shown approximately in 42% to 43% of both groups and homozygote for allele G was seen in 13% and 15% of patients and controls, respectively.

Furthermore, the genotype frequencies of the polymorphism in the healthy control group followed the Hardy-Weinberg Equilibrium (p-value=0,32), which states that allele and genotype frequencies in a population remain constant from generation to generation in the absence of other evolutionary influences. Therefore, our results show that our control population was not affected by evolutionary processes of mutation, migration or selection.

Table 5 -Genotype distribution of *PLCE1* gene rs2274223 polymorphism in patient with CRC and healthy control.

Genotype distribution was accessed by Chi-square test and a logistic regression was performed to adjust for sex and age. N represent the number of samples. p-value were calculated using 95% confidence interval.

		Patients	Controls	p-value	Adjusted p-value
Genotype	**AA N (%)**	97 (44,5%)	96 (43,4%)	0,882	0,839
	AG N (%)	93 (42,7%)	93 (42,1%)		
	GG N (%)	28 (12,8%)	32 (14,5%)		
					Adjusted for age and sex

Nevertheless, the risk of developing CRC for each group (AA vs. AG; AA vs. GG; GG vs. AG), comparing the most common genotype with the other two genotypes together (AA vs. AG+GG) and comparing the rarest genotype with the other two (GG vs. AA+AG), do not show any significant difference between the genotypes, both in the univariate and multivariate (adjusting for age and sex) analyses (Table 6).

In any case, if a trend can be inferred from these results, although not significant, is a possible protective role of the GG genotype. Our results are in agreement with what was found by Kupcinskas, J. et al.[58], who showed no association of this SNP with CRC risk, but opposite to Wang, Q. et al.[52] and Ezgi, O. et al.[59].

Table 6 – Association between *PLCE1* genotypes and risk of CRC in patients with CRC and healthy controls. The number of samples used in the determination of risk is represented by N. Risk of developing colorectal cancer was accessed by odds ratio (OR) with a Confidence Interval of 95% (CI), p-value was calculated by two-sided Fisher's exact test with a 5% significance (α) and a logistic regression was performed to adjust for sex and age.

						Adjusted		
		N	OR	CI	p-value (α=0,05)	OR	CI	p-value (α=0,05)
Genotype	AA vs GG	253	1,16	0,65 – 2,07	0,66	0,88	0,49 – 1,59	0,68
	AA vs. AG	379	1,01	0,68 – 1,51	1,00	1,01	0,67 – 1,52	0,95
	AA vs. AG+GG	439	1,04	0,72 – 1,51	0,85	1,03	0,70 – 1,50	0,89
	GG vs. AG	246	0,88	0,49 – 1,56	0,66	0,84	0,47 – 1,51	0,56
	GG vs. AA+AG	439	0,87	0,50 – 1,50	0,68	0,85	0,49 – 1,47	0,56
						Adjusted for age and sex		

4.2. Prognostic value of *PLCE1* rs2274223 polymorphism

In order to access the prognostic value of *PLCE1* polymorphism rs2274223, we compared the overall survival (OS) of the different *PLCE1* genotypes in our patient's population. Since some studies point the genotype GG as having the worst outcome, we analysed OS in two arms, comparing genotypes AA and AG against GG[61].

Furthermore, given that the prognosis of patients in stage I-III is very distinct from patients diagnosed in stage IV, we analysed the role of *PLCE1* polymorphism rs2274223 separately in these two groups of patients.

4.2.1. *PLCE1* rs2274223 polymorphism prognostic value in stage I-III colorectal cancer patients

In patients with stage I-III CRC there are demographic, clinical and pathological characteristics such as age, sex, stage, grade, location, obstruction and/or perforation at time of diagnosis, lymphatic/neural and/or vascular invasion (LNVI) and the presence of positive margins after surgery (PMAS), that may have an impact in the prognosis of patients.

Therefore, before accessing the prognostic value of *PLCE1* rs2274223 polymorphism in stage I-III CRC patients it was necessary to analyse if these characteristics were associated with this polymorphism (Table 7).

Table 7 shows that there is no association between this SNP and any demographic, clinical or pathological characteristic investigated.

Table 7 - Clinical and demographic characteristics of stage I to III CRC patients. Clinical and demographic characteristics of CRC patients and their association with genotype were analysed by independent T-test for age, and by Chi-square test for sex, stage, grade, location, obstruction/perforation, LVNI and PMAS. N represents the number of samples, P25 and P75 represent the 25th and 75th percentiles of follow-up and age. p-value were calculated using 95% confidence interval.

		Genotype		
		AA/AG	GG	p-value
Number of Patients N (%)		149 (87,1%)	22 (12,9%)	/
Follow up (Months)	**Median**	71,475		/
	P25 - P75	47,475 - 93,279		
	Median	72,623	65,508	/
	P25 - P75	45,279 - 93,295	53,959 - 92,508	
Age (Years)	**Median**	72	66,5	0,460
	P25 - P75	63,0 - 77,5	62,00 - 72,25	
Sex N (%)	**Male**	86 (57,7%)	16 (72,7%)	0,245
	Female	63 (42,3%)	6 (27,3%)	
Stage N (%)	**I**	44 (29,5%)	3 (13,6%)	0,192
	II	50 (33,6%)	7 (31,8%)	
	III	55 (36,9%)	12 (54,5%)	
Grade N (%)	**G1**	43 (28,9%)	5 (22,7%)	0,817
	G2	85 (57,0%)	14 (63,6%)	
	G3	21 (14,1%)	3 (13,6%)	
Location N (%)	**Right Colon**	46 (30,9%)	2 (9,1%)	0,071
	Left Colon	46 (30,9%)	7 (31,8%)	
	Rectum	57 (38,2%)	13 (59,1%)	
Obstruction Perforation N (%)	**No**	93 (62,4%)	12 (54,5%)	0,631
	Yes	53 (35,6%)	9 (40,9%)	
	Unknown	3 (2,0%)	1 (4,5%)	
LVNI N (%)	**No**	119 (79,9%)	20 (90,9%)	0,207
	Yes	27 (18,1%)	1 (4,5%)	
	Unknown	3 (2,0%)	1 (4,5%)	
PMAS N (%)	**No**	139 (93,3%)	20 (90,9%)	1,000
	Yes	7 (4,7%)	1 (4,5%)	
	Unknown	3 (2,0%)	1 (4,5%)	

Following this analysis, we tested the implications of *PLCE1* genotypes on the OS of patients (by univariate and multivariate models) (Figure 6).

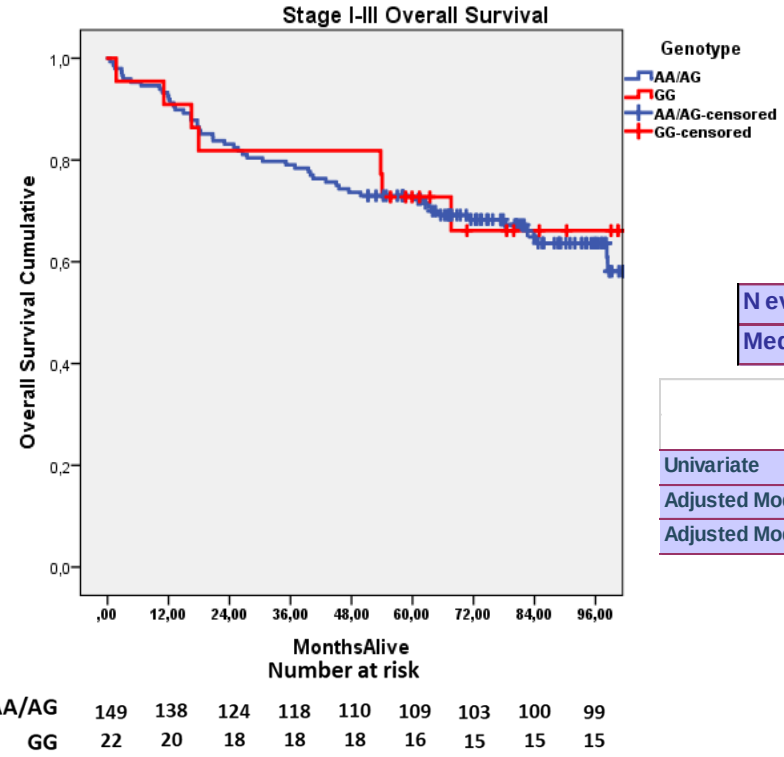
Therefore, the Kaplan-Meier (KM) representation of OS according to *PLCE1* genotype. This group of CRC patients was composed by 171 patients, of which 149 had the genotypes AA or AG (87,1%), while 22 had the genotype GG (12,9%).

Patients were followed for a maximum of 117,57 months (approximately 10 years), with a median follow-up time of 71,475 months (P25=47,475; P75=93,279). During this period 52 patients with the genotypes AA or AG and 7 patients with the genotype GG died (35,1% and 31,8% respectively). However, the median survival was not reached for any of the two groups.

Our results indicate that the genotype had no influence in the patient's outcome in a univariate analysis of OS (log-rank test p-value= 0,805; HR= 1,105; CI= 0,502 – 2,433). After adjustment for age, stage, grade, tumour location, obstruction and perforation, LVNI and PMAS at time of diagnosis (model 1, Table 8), the genotype is still seen not associated with patient's outcome (p-value= 0,992; HR= 1,005; CI= 0,409 – 2,468). Finally, a second model of analysis was performed accounting only for the variables that had shown association with the outcome on model 1 (Table 8, Age p-value= 0,001; Stage p-value= 0,006; LVNI p-value= 0,007). Confirming the previous results, model 2 shows that *PLCE1* genotype has no influence on patient's outcome (p-value= 0,793; HR= 1,124; CI= 0,469 – 2,693).

However, this multivariate analysis had a small number of patients with GG genotype and, for that reason, increasing the number of this patients could empower our analysis.

Overall, our results show for the first time that *PLCE1* rs2274223 polymorphism has no influence on patient's survival is stage I-III.



	Genotype	
	AA/AG	GG
N events (%)	52 (34,9%)	7 (31,8%)
Median	N.R.	N.R.

	HR	CI	P
Univariate	1,105	0,502 - 2,433	0,805
Adjusted Model 1	1,005	0,409 - 2,468	0,992
Adjusted Model 2	1,124	0,469 - 2,693	0,793

Figure 6 – Survival curve and univariate and multivariate analysis of OS from patients with CRC from stage I to III. Univariate analysis was performed comparing patient's genotypes AA/AG vs. GG according to their alive status across the time. N events (%) represents the number and percentage of patients that died in this analysis, N.R. represents that median survival was not reached. Hazard Ratio (HR) and p-value (P) were calculated with a 95% confidence interval (CI). Multivariate analysis was used to compare both arms (AA/AG vs. GG) in the adjusted models 1 and 2, which were calculated by cox regression with 95% confidence interval.

Table 8 – Models used for multivariate analysis of OS from patients with CRC from stage I to III. In Model 1 the variables used to calculate survival were genotype, age, stage, grade, tumour location, LVNI, obstruction/perforation and PMAS. In Model 2 the variables used to calculate survival were genotype, age, stage and LVNI. Adjustment to both models were performed using Cox regression. Hazard Ratio (HR) and p-value were calculated with a 95% confidence interval (CI).

Model 1 Variables in the Equation					Model 2 Variables in the Equation				
	HR	95% CI		p-value		HR	95% CI		p-value
		Lower	Upper				Lower	Upper	
Genotype AA+AG vs. GG	1,01	0,41	2,47	0,992	**Genotype AA+AG vs. GG**	1,12	0,47	2,69	0,793
Age	1,05	1,02	1,09	0,001	**Age**	1,06	1,03	1,09	0,000
Stage I				0,006	**Stage I**				0,003
Stage II	0,28	0,11	0,69	0,006	**Stage II**	0,27	0,12	0,63	0,002
Stage III	0,46	0,24	0,89	0,020	**Stage III**	0,50	0,28	0,91	0,024
Grade 1				0,619	**LVNI**	0,43	0,24	0,79	0,006
Grade 2	0,70	0,29	1,65	0,411					
Grade 3	0,71	0,34	1,47	0,352					
Left Colon				0,633					
Right Colon	1,41	0,68	2,94	0,352					
Rectum	1,29	0,62	2,69	0,493					
LVNI	0,41	0,21	0,78	0,007					
Obstruction Perforation	1,19	0,65	2,15	0,575					
PMAS	0,49	0,16	1,51	0,213					

4.2.2. *PLCE1* rs2274223 polymorphism prognostic value in stage IV colorectal cancer patients

It is equally necessary to account for the clinical, demographic and pathological characteristics that influence prognosis of patients with CRC stage IV such as age, sex, tumour location, presence of *KRAS* mutations and metastasis location, prior the analysis of *PLCE1* rs2274223 polymorphism association with patient's OS.

The Table 9 shows the association between clinical and demographic characteristics and *PLCE1* rs2274223 polymorphism.

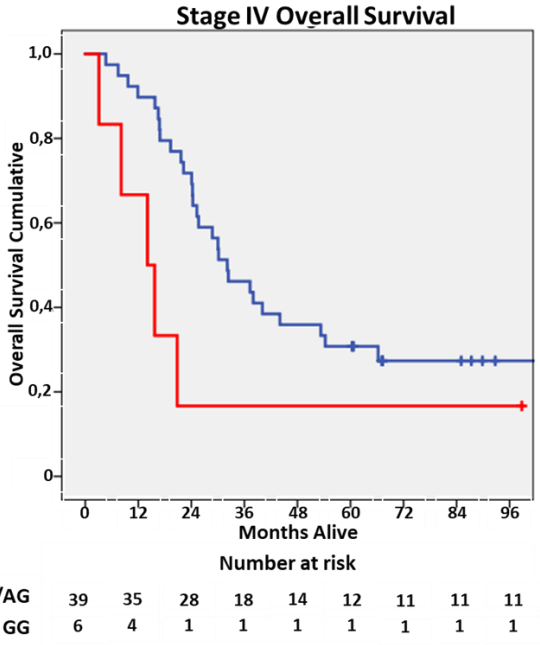
Table 9 – Clinical and demographic characteristics of CRC patients between stage IV associated with *PLCE1* rs2274223 polymorphism. Clinical and demographic characteristics of CRC patients and their association with genotype were analysed by independent T-test for age, and by Chi-square test for sex, location, *KRAS* mutations and organ of metastasis. p-value were calculated using 95% confidence interval.

		Genotype		
		AA/AG	GG	p-value
Number of Patients N (%)		39 (86,7%)	6 (13,3%)	/
OS Follow up (Months)	**Median**	30,033		/
	P25 - P75	16,836 - 60,574		
	Median	32,131	14,885	
	P25 - P75	21,639 - 60,656	6,869 - 40,279	

Age (Years)	**Median**	67	77	0,024
	P25 - P75	54 - 73	68,5 - 83,25	
Sex N (%)	**Male**	24 (61,5%)	3 (50,0%)	0,670
	Female	15 (38,5%)	3 (50,0%)	
Location N (%)	**Left Colon**	16 (41,0%)	3 (50,0%)	0,555
	Right Colon	8 (20,5%)	2 (33,3%)	
	Rectum	15 (38,5%)	1 (16,7%)	
***KRAS* mutations N (%)**	**No**	18 (46,2%)	1 (16,7%)	0,312
	Yes	13 (33,3%)	3 (50,0%)	
	Unknown	8 (20,5%)	2 (33,3%)	
Metastasis Organ N (%)	**Liver Only**	22 (56,4%)	6 (100%)	
	Lung Only	2 (5,1%)	0	
	Multiple	11 (28,2%)	0	
	Other	4 (10,3%)	0	

In stage IV CRC patients, *PLCE1* genotype was associated with age (p=0,024), being GG patients, in average, 10 years older than AA/AG CRC patients. The other clinical and demographic characteristics analysed had no association with patient's *PLCE1* genotype.

Kaplan-Meier curves of these patients are represented on Figure 7 and the adjustment model of the multivariate analysis can be consulted on Table 10.



	Genotype	
	AA/AG	GG
N events (%)	28 (71,8%)	5 (83,3%)
Median Survival	32,131	14,066
CI (95%)	21,700 - 42,562	4,975 - 23,156

	HR	CI	P-value
Univariate	0,385	0,142 - 0,975	0,044
Adjusted Model	0,389	0,131 - 1,155	0,089

Figure 7 – Survival curve, univariate and multivariate analysis of OS from patients with CRC from stage IV. Univariate analysis was performed comparing patient's genotypes AA/AG vs. GG according to their alive status across the time. N events (%) represents the number and percentage of patients that died in this analysis. Hazard Ratio (HR) and p-value (P) were calculated with a 95% confidence interval (CI). Multivariate analysis was used to compare both arms (AA/AG vs. GG) in the adjusted model and was calculated by cox regression with 95% confidence interval.

In this cohort, 45 CRC patients were analysed, of which 39 had the genotypes AA or AG (86,7%), while 6 had the genotype GG (13,3%). Patients were followed for a maximum of 106,03 months (approximately 9 years), with a median follow-up time of 30,033, months (P25= 16,836; P75= 60,574). During this period 28 patients with the genotypes AA or AG and 5 patients with the genotype GG died (71,8% and 83,3% respectively).

Both groups reached the median survival which was 32,121 months (CI= 21,700 – 42,562) for genotypes AA or AG and 14,066 months (CI= 4,975 – 23,156) for genotype GG.

In patients at stage IV, genotype had a significant influence in the univariate analysis of OS (log-rank test p-value = 0,044; HR= 0,385; CI= 0,142 – 0,975). However, given that association of *PLCE1* genotype was seen with age, a multivariate model was needed to effectively infer the value of *PLCE1* rs2274223 polymorphism in OS. After adjustment for age, tumour location and metastasis location, *PLCE1* genotype is not statistically significative for patient's survival (adjustment model p-value= 0,089; HR= 0,389; CI= 0,131 – 1,155).

KRAS mutation status was excluded from the adjustment model since there were a considerable number of patients with missing information, which could lead to incorrect results.

Altogether, these results suggest that *PLCE1* rs2274223 polymorphism may not influence patient's survival is stage IV. Although a trend for negative outcome may be seen. Given that in this analysis only 6 patients were homozygotic for this allele, increasing our cohort could effectively answer the question of whether *PLCE1* SNP rs2274223 contributes to the outcome of CRC patients in stage IV.

Table 10 – Models used for multivariate analysis of OS from patients with CRC from stage IV. In this adjustment model the variables used to calculate survival were genotype, age, tumour location and organ of metastasis. The adjustment model was performed using Cox regression. Hazard Ratio (HR) and p-value were calculated with a 95% confidence interval (CI).

Adjustment Model Variables				
		95% CI for HR		
	HR	Lower	Upper	p-value
Genotype AA+AG vs GG	0,389	0,131	1,155	0,089
Age	1,037	1,000	1,074	0,048
Left Colon				0,071
Right Colon	2,620	1,055	6,503	0,038
Rectum	3,174	1,072	9,398	0,037
Liver Metastasis				0,601
Lung Metastasis	0,381	0,091	1,587	0,185
Multiple Metastasis	0,000	0,000		0,975
Other Metastasis	0,377	0,084	1,688	0,202

4.3. PLCε activity after wild-type and mutant (H1927A) *PLCE1* overexpression

As described before, PLCε is responsible for cleaving PIP_2 into the second messengers IP_3 and DAG.

We investigated whether *PLCE1* rs2274223 polymorphism could affect PLCε lipase activity (Figure 8).

In order to do that, we performed a lipase catalytic assay in Cos-7 cells overexpressing wild-type and mutant (H1927A) PLCε (Figure 8A). This assay allows to quantify IP3 release upon PLC activation.

Therefore, as expected, our results show that overexpressing PLCε in Cos-7 cells increases significantly PLCε activity when compared to non-transfected cells (wild-type p-value=0,0014; mutant p-value=0,0008). Furthermore, after stimulation of cells with EGF, is seen an increase in PLC activity given that activation of EGF intracellular pathway activates PLCs (wild-type p-value=0,0005; mutant p-value=0,0002).

Finally, we can see that PLCε mutant has a slightly higher activity than the wild-type enzyme in cells without stimulation, and this activity is even higher after EGF stimulation (without stimulation: mutant mean=16,24; wild-type mean=12,63; with stimulation: mutant mean=39,24; wild-type mean=34,74; p=0,045). However, expression of the PLCε mutant (H1927A) was also higher than expression of wild-type protein (Figure8B). These results, therefore, suggest that H1927A mutation is likely to affect protein expression rather than its activity.

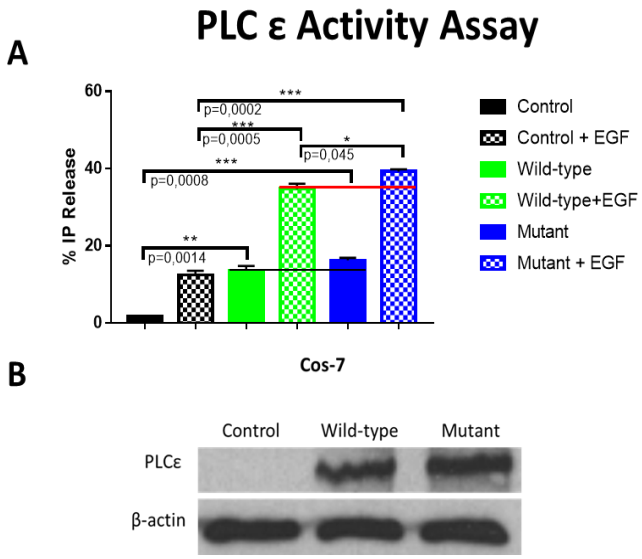


Figure 8 – PLCε Activity Assay. PLCε activity was measured by the percentage of IP release. The assay was performed in two independent experiments, each experiment was performed in duplicates. %IP Release data is present as mean ± SD and p-values (p) were calculated using one-way ANOVA and Tukey's multiple comparisons tests with a CI of 95% (A), Protein levels of PLCε were detected by mouse anti-GFP and β-actin was used as loading control in western blot (B).

4.4. *PLCE1* rs2274223 polymorphism role in EMT biomarkers

Although we didn't see differences in activity between PLCε wild-type and mutant (h1927A), we sought to analyse if PLCε enzymes could affect epithelial-to-mesenchymal transition (EMT) markers in CRC cell lines by qRT-PCR (Figure 9).

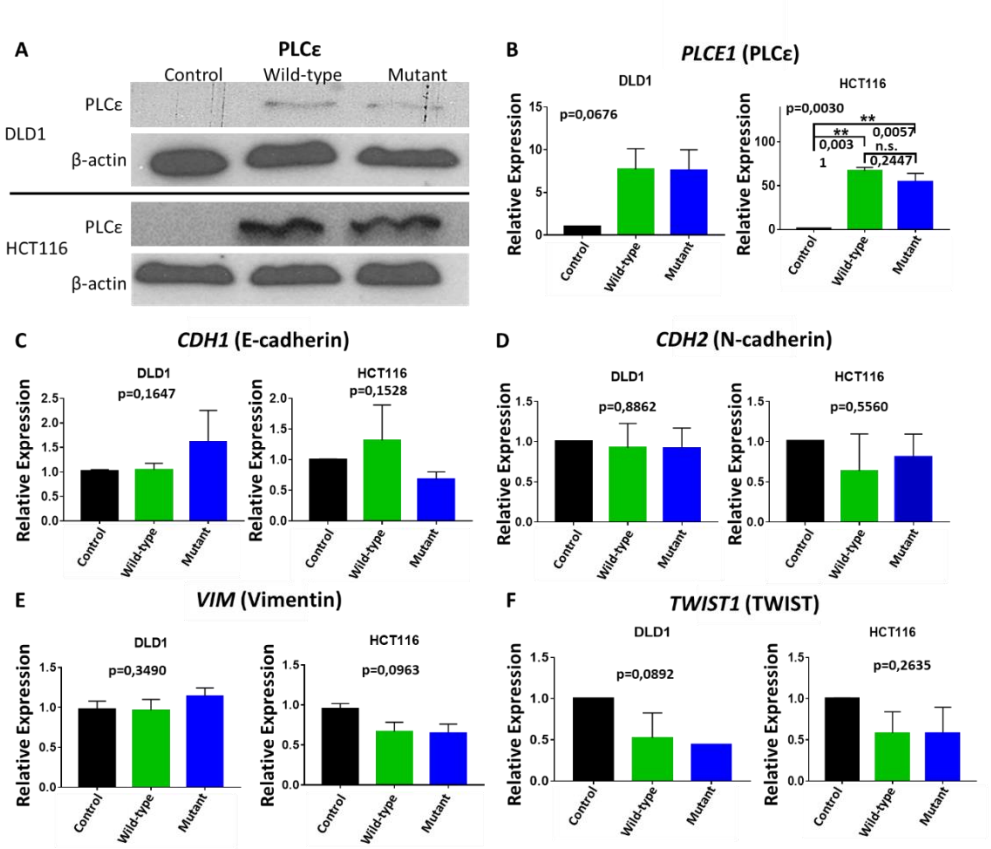


Figure 9 – Relative expression analysis of EMT Biomarkers. Western blot of PLCε overexpression with anti-GFP, β-actin was used as loading control (A), relative expression of PLCε (B), E-cadherin (C), N-cadherin (D), Vimentin (E) and TWIST (F) in DLD1 and HCT116 cell lines after overexpression of PLCε wild-type and mutant (H1927A). As controls were used non-transfected cells. The relative expression of these markers was calculated using GAPDH as standard reference. The data is present as the mean ± SD. The data shown in this figure is representative of two independent assays performed in triplicate (N=3). In the figures p-value is represented by (p) and was calculated by one-way ANOVA and Tukey's multiple comparisons tests with a CI of 95%.

Figure 9 shows that *PLCE1* overexpression with wild-type and mutant (H1927A) plasmid increased the levels of protein and mRNA of PLCε both in DLD1 and HCT116 cell lines (Figure 9A and 9B, respectively) but had no effect in the relative expression of mRNA in *CDH1* (E-cadherin), *CDH2* (N-cadherin), *VIM* (Vimentin) and *TWIST1* (TWIST) (Figure 9C, 9D, 9E and 9F). Furthermore, in CRC cell lines, expression of PLCε wild-type and mutant (h1927A) seem identical.

Overall, these results suggest that H1927A mutation of PLCε does not affect the EMT process of CRC cells.

4.5. *PLCE1* rs2274223 polymorphism role in inflammation and angiogenesis

Mingzhen Li et al. reported that PLCε promotes intestinal tumorigenesis of $Apc^{Min/+}$ mice through augmentation of inflammation and angiogenesis biomarkers, such as COX-2, CXCL-1 and CXCL-2 and VEGF-A, respectively[40].

Furthermore, STAT3 transcription factor was also associated with PLCε by Xue Yang et al. which report that knockdown of PLCε decreased the levels of STAT3 phosphorylation in bladder cancer cell lines[42]. Given that STAT3 pathway is responsible for the transcription of a variety of genes involved in the regulation of critical functions, including cell differentiation, proliferation, apoptosis, angiogenesis, metastasis, and immune responses (mainly inflammation)[62]. The same study, also reported a decrease in the levels of *IL1B*, *IL6*, *TNFA*[42].

Therefore, it is important to investigate if PLCε mutant (H1927A) protein affects the regulation and expression of these inflammatory and angiogenic markers differently than the wild type variant (Figure 10).

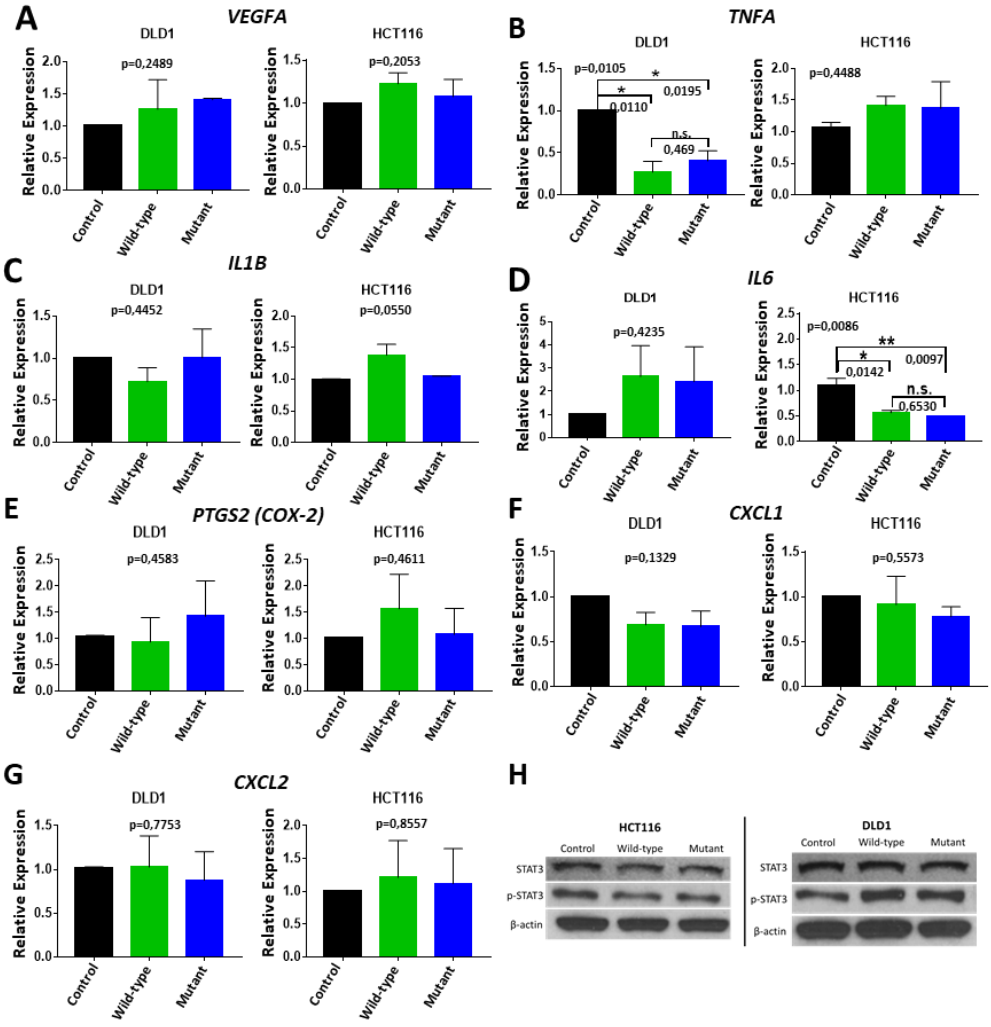


Figure 10 – Analysis of inflammatory and angiogenic biomarkers. Relative expression of *VEGFA* (A), relative expression of *TNFA* (B), *IL1B* (C), I*L6* (D), *PTGS2* (E), *CXCL1* (F) and *CXCL2* (G) in DLD1 and HCT116 cell lines after overexpression of PLCε

wild-type and mutant (H1927A). As controls were used non-transfected cells. The relative expression of these markers was calculated using GAPDH as standard reference. Relative expression data is present as the mean ± SD. The data shown in this figure is representative of two independent assays performed in triplicate (N=3). In the figures p-value is represented by (p) and was calculated by one-way ANOVA and Tukey's multiple comparisons tests with a CI of 95%. STAT3 and p-STAT3 were accessed by western blot in HCT116 and DLD1 cell lines (H). β-actin was used as loading control. The assay for western blot was performed in duplicate (N=2).

We analysed the relative expression of mRNA of *VEGFA, TNFA, IL1B, IL6, PTGS2* (COX-2), *CXCL1* and *CXCL2* (Figures 10A-G respectively), and the protein levels of STAT3 and p-STAT3 (Figure 10H).

We could not find any difference in the mRNA levels of these cytokines or phosphorylation status of STAT3 when comparing PLCε wild-type and mutant (H1927A).

We found, however, that PLCε overexpression of both variants significantly decreased *TNFA* in DLD1 cell lines and *IL6* in HCT116 cell lines.

Altogether, these results indicate that PLCε genotype does not seem to have a role in inflammatory tumour initiation and angiogenesis in colorectal cancer.

4.6. *PLCE1* rs2274223 polymorphism role in cell proliferation

Finally, we wanted to see if this polymorphism could affect cell proliferation. In order to do that, we performed a cell viability assay after overexpressing PLCε wild-type and the mutant (H1927A) in HCT116 and DLD1 cell lines (Figure 11A and 11B, respectively).

Although overexpression of both variants of PLCε seem to decrease cell proliferation in HCT116 and DLD1 cell lines, we could not see any difference between PLCε variants (p-value=0,3473 and p-value=0,8532, respectively).

Therefore, these results are in agreement with a possible tumour suppressive role for PLCε, but *PLCE1* rs2274223 polymorphism has no impact in cell proliferation.

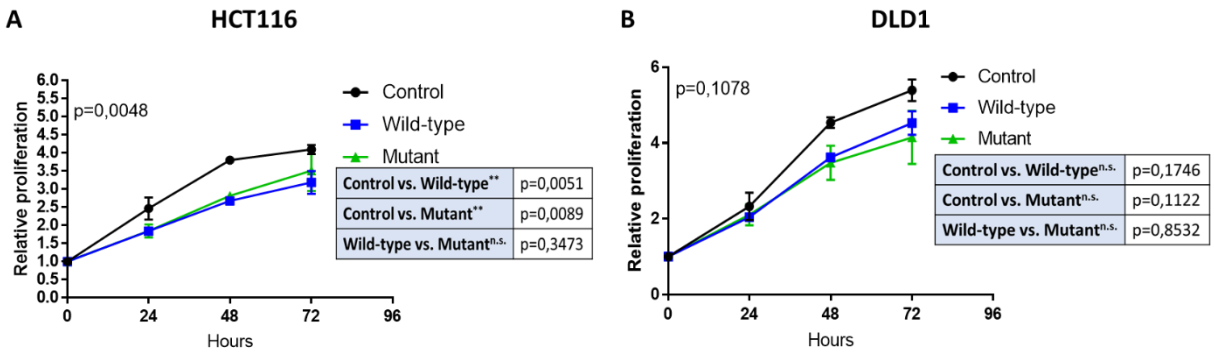


Figure 11 – Cell Viability Assay. Cell viability assay was performed with alamarBlue® in HCT116 (A) and DLD1 (B) cell lines after overexpression of PLCε wild-type and mutant (H1927A). As controls were used non-transfected cells. Relative proliferation data is present as the mean ± SD. The data shown in this figure is representative of two independent assays. In the figures p-value is represented by (p) and was calculated by two-way ANOVA and Tukey's multiple comparisons tests with a CI of 95%.

5. Conclusions and Future Perspectives

CRC is a disease which incidence and mortality has been raising along the last years. Therefore, it is important to diagnose CRC the earliest possible to improve patient's outcome.

Only after diagnosis, the best treatment is defined and patients are followed in all the course of the disease. For that reason, identifying new biomarkers will help diagnosis, treatment definition and patient's follow-up.

PLCs have been studied as possible new biomarkers[28]. One of these PLCs is PLCε. Despite, it role in cancer be controversial, this enzyme is seen as tumour suppressor enzyme in colorectal cancer[49–51].

In the last decade, *PLCE1* SNP rs2274223 association with colorectal cancer has been studied with contradictory results[49,52,58,59,61,63].

Initially this polymorphism was associated with the risk of developing gastric and esophageal cancer[47,53,54,64]. In our study we show that *PLCE1* SNP rs2274223 is not associated with an increased risk of developing CRC. Therefore, it is possible that PLCε plays a different role in CRC.

This result is opposite to most of the studies present, so far, for this polymorphism[47,52-59,61,62,64]. It is also important to refer that most of these studies approached Chinese populations[52–54,57,63,64] and there are studies in other populations with contradictory results[55,58]. This could be explained by the existence of a lower Linkage Disequilibrium[65] (LD) in other populations, which would mean that another polymorphism with a higher LD could be responsible for tumorigenesis in Chinese population, instead of *PLCE1* SNP rs2274223.

Furthermore, we analysed how this polymorphism could have an influence in patient's survival.

After analysing OS of patients at stage I-III and IV by univariate and multivariate analysis comparing arm AA/AG with GG and analysing Kaplan-Meier curves with log rank test, our findings suggest that *PLCE1* SNP rs2274223 is not associated with patient's survival. Moreover, in the OS analysis of patients at stage IV, those who were homozygotic for allele G had a negative trend on survival time.

However, in both analysis a small number of patients homozygotic for allele GG were analysed (22 in stage I-III and 6 in stage IV) and for that reason a larger cohort could empower the results of univariate and multivariate analysis and clarify if *PLCE1* SNP rs2274223 has, or not, a prognostic value in CRC patients in stage I-III and in stage IV.

Based on these results, we analysed *in vitro* how this polymorphism could influence PLCε phospholipase activity and cell functions such as cell proliferation, angiogenesis and immune responses.

First, we show that overexpressing PLCε wild-type and mutant (H1927A) enzymes did not influence PLCε phospholipase activity in Cos-7 cell line. However, this enzyme has other functional domains, such as CDC25 and RA2 domains, that were not approached in our study and might be affected by this SNP.

Next, we analysed how overexpressing PLCε with wild-type and mutant (H1927A) impact EMT. In order to do that, we analysed the mRNA expression levels of *CDH1* (E-cadherin), *CDH2* (N-cadherin), *VIM* (Vimentin) and *TWIST1* (TWIST). We could not see any difference between the overexpression of wild-type and mutant (H1927A) in any of these EMT biomarkers, which suggest that PLCε enzymes have no influence in EMT.

Since, there are studies reporting that PLCε affects inflammation and angiogenesis we decided to investigate how this polymorphism could influence the mRNA expression levels of *VEGFA, PTGS2* (COX-2), *TNFA, CXCL1, CXCL2, IL1B e IL6*[40,42]*.* Our results show that overexpressing wild-type and mutant (H1927A) PLCε did not affect angiogenic biomarker mRNA expression *VEGFA.* These results also suggest that despite overexpressing both variants of PLCε lead to a decrease in the expression of *TNFA* in DLD1 and a decrease of expression of *IL6* in HCT116 cell lines, there were no difference of expression between the wild-type and the mutant (H1927A) form of PLCε. Furthermore, we could not see any difference in the expression of *PTGS2, CXCL1, CXCL2* and *IL1B*, between control cells and after PLCε overexpression of wild-type and mutant (H1927A) forms.

We also analysed the STAT3 pathway, since some studies showed that PLCε knockdown could decrease STAT3 phosphorylation[42]. This pathway has shown to have a role in cell differentiation, proliferation, apoptosis, angiogenesis, metastasis, and immune responses[62]. STAT3 and p-STAT3 protein levels were not affected, in western blot, by PLCε overexpression of wild-type and mutant (H1927A) protein. Which indicates that PLCε has no influence on STAT3 pathway in CRC cell lines. Overall, these results are consistent with what we saw in the OS of CRC patients at stage I-III.

Finally, using alamarBlue®, we accessed how the overexpression of PLCε wild-type and mutant (H1927A) was affecting proliferation. PLCε overexpression decreased cell proliferation, which confirms that this enzyme act as a tumour suppressor. However, there were no significative differences in cell proliferation between PLCε wild-type and mutant (H1927A). This result shows that cell proliferation is not affect by PLCε rs2274223 polymorphism.

In resume, our results are consistent, since this study shows that *PLCE1* SNP rs2274223 does not increase the risk of CRC development, does not influence statistically patient's survival (although a trend was seen in patients at stage IV homozygotic for allele G), have no effect on cell proliferation, EMT, angiogenesis and inflammation.

Ultimately, this study suggests that *PLCE1* SNP rs2274223 may not have a role in CRC development and progression.

References

1. Wright EP, *et al*. Social problems in oncology. *Br. J. Cancer*, **87**, 1099-1104 (2002).

2. World Health Organization. Part 3: Disease incidence, prevalence and disability. *The Global Burden of Disease 2004: Update*. 1st Edition. Geneve, World Health Organization (2008).

3. Ferlay J, *et al.* Cancer incidence and mortality worldwide: Sources, methods and major atterns in GLOBOCAN 2012. *Int. J. Cancer*, **136**, E359-E386 (2015).

4. Bray F, *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.*, **0**, 1-31(2018)

5. Worldwide cancer statistics. Cancer research UK. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/worldwide-cancer#heading-Zero. Accessed at: November 30th, 2018

6. World Health Organization. Cancer. Available at: http://www.who.int/news-room/fact-sheets/detail/cancer. Accessed at: November 30th, 2018.

7. VanPutte CL, *et al.* Chapter 24: Digestive System. *Seeley's Anatomy & Physiology*. 11th Edition. New York, McGraw-Hill Education (2016).

8. Hill MJ, *et al*. Etiology of adenoma-carcinoma sequence in large bowel. *Lancet*, **311**, 245-247 (1978).

9. American Cancer Society. Understanding Your Pathology Report: Colon Polyps. Available at: https://www.cancer.org/treatment/understanding-your-diagnosis/tests/understanding-your-pathology-report/colon-pathology/colon-polyps-sessile-or-traditional-serrated-adenomas.html. Accessed at: November 30th, 2018.

10. Johnson CM, *et al.* Meta-analyses of colorectal cancer risk factors. *Cancer Causes Control*, **24**, 1207-1222 (2014).

11. Fearon EF, *et al*. A genetic model for colorectal tumorigenesis. *Cell*, **61**, 759-767 (1989).

12. Wood LD, *et al.* The genomic landscapes of human breast and colorectal cancers. *Science*, **318**, 1108-1113 (2007).

13. Bogaert J, *et al*. Molecular genetics of colorectal cancer. *Ann. Gastroenterol. Q. Publ. Hell. Soc. Gastroenterology*, **27**, 9-14 (2014).

14. Goel A, *et al.* Hereditary and familial colon cancer. *Gastroenterology*, **138**, 2044-2058 (2011).

15. Segnan N, *et al*. Evidence for effectiveness of FOBT screening. *European Guidelines for quality assurance in colorectal cancer screening and diagnosis*. 1st Edition. Luxembourg, Publications Office of the European Union (2010).

16. Glynne-Jones R, *et al.* Rectal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.*, **28**, iv22-iv40 (2017).

17. Van Cutsem E, *et al.* ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann. Oncol.*, **27**, 1386–1422 (2016).

18. Labianca R, *et al.* Early colon cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.*, **24**, vi64-vi72 (2013).

19. Lech G, *et al.* Colorectal cancer tumour markers and biomarkers: Recent therapeutic advances. *World J. Gastroenterol.*, **22**, 1745–1755 (2016).

20. Sepulveda AR, *et al.* Molecular biomarkers for the evaluation of colorectal cancer: Guideline from The American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and the American Society of Clinical Oncology. *J. Clin. Oncol.*, **35**, 1453–1496 (2017).

21. Lim DR, *et al*. Comparison of oncological outcomes of right-sided colon cancer versus left-sided colon cancer after curative resection. *Medicine*, **96**, 1–7 (2017)

22. Hauptman N, *et al*. Colorectal cancer blood-based biomarkers. *Gastroenterology Research and Pratice*, **2017**, 1-11 **(**2017).

23. Gresset A, *et al*. The Phospholipase C isozymes and their regulation. *Sub-Cellular Biochemistry*, **58**, 61-94 (2012).

24. Suh P, *et al.* Multiple roles of phosphoinositide-specific phospholipase C isozymes. *BMB Rep.*, **41**, 415–434 (2008).

25. Bunney TD, *et al*. PLC regulation: Emerging pictures for molecular mechanisms. *Trends Biochem. Sci.*, **36**, 88–96 (2011).

26. Nakamura Y, *et al.* Regulation and physiological functions of mammalian phospholipase C. *J. Biochem.*, **161**, 315–321 (2017).

27. Fukami K. Structure, regulation and function of phospholipase C isozymes. *J. Biochem.*, **131**, 293-299 (2002).

28. Koss H, *et al*. Dysfunction of phospholipase Cγ in immune disorders and cancer. *Trends Biochem. Sci.*, **39**, 603–611 (2014).

29. Sorli SC, *et al*. Signaling properties and expression in normal and tumor tissues of two phospholipase C epsilon splice variants. *Oncogene*, **24**, 90–100 (2005).

30. Bunney TD, *et al*. Regulatory links between PLC enzymes and Ras superfamily GTPases: Signalling via PLCε. *Adv. Enzyme Regul.*, **49**, 54–58 (2009).

31. Bunney TD, *et al.* Structural and mechanistic insights into ras association domains of phospholipase C epsilon. *Mol. Cell*, **21**, 495–507 (2006).

32. Bunney TD, *et al*. Phospholipase C epsilon: linking second messengers and small GTPases. *Trends Cell Biol.*, **16**, 640–648 (2006).

33. Song C, *et al.* Regulation of a Novel Human Phospholipase C, PLCε, through membrane targeting by Ras. *J. Biol. Chem.*, **276**, 2752–2757 (2001).

34. Smrcka AV, *et al*. Role of phospholipase Cε in physiological phosphoinositide signaling networks. *Cell. Signal.*, **24**, 1333–1343 (2012).

35. Wang H, *et al.* Phospholipase C ε modulates β-adrenergic receptor-dependent cardiac contraction and inhibits cardiac hypertrophy. *Circ. Res.*, **97**, 1305–1313 (2005).

36. Hinkes B, *et al.* Positional cloning uncovers mutations in PLCE1 responsible for a nephrotic syndrome variant that may be reversible. *Nat. Genet.*, **38**, 1397–1405 (2006).

37. Park JB, *et al.* Phospholipase signalling networks in cancer. *Nat. Rev. Cancer*, **12**, 782–792 (2012).

38. Zhang RY, *et al.* PLCε signaling in cancer. *J. Cancer Res. Clin. Oncol.*, **142**, 715–722 (2016).

39. Ikuta S, *et al*. Crucial role of phospholipase C in skin inflammation induced by tumor-promoting phorbol ester. *Cancer Res.*, **68**, 64–72 (2008).

40. Li M, *et al*. Phospholipase Cε promotes intestinal tumorigenesis of ApcMin/+ mice through augmentation of inflammation and angiogenesis. *Carcinogenesis*, **30**, 1424–1432 (2009).

41. Harada Y, *et al*. PLCε cooperates with the NF-κB pathway to augment TNFα-stimulated CCL2/MCP1 expression in human keratinocyte. *Biochem. Biophys. Res. Commun.*, **414**, 106–111 (2011).

42. Yang X, *et al.* Knockdown of PLCε inhibits inflammatory cytokine release via STAT3 phosphorylation in human bladder cancer cells. *Tumor Biol.*, **36**, 9723–9732 (2015).

43. Wang Y, *et al.* PLCε knockdown inhibits prostate cancer cell proliferation via suppression of Notch signalling and nuclear translocation of the androgen receptor. *Cancer Lett.*, **362**, 61–69 (2015).

44. Wakita M, *et al.* Phospholipase Cε activates nuclear factor-κB signaling by causing cytoplasmic localization of ribosomal S6 kinase and facilitating its phosphorylation of inhibitor κB in colon epithelial cells. *J. Biol. Chem.*, **291**, 12586–12600 (2016).

45. Bai Y, *et al.* Crucial role of phospholipase Cepsilon in chemical carcinogen-induced skin tumor development. *Cancer Res.*, **64**, 8808–8810 (2004).

46. Martins M, *et al.* Tumor suppressor role of phospholipase C in Ras-triggered cancers. *Proc. Natl. Acad. Sci.*, **111**, 4239–4244 (2014).

47. Cui X, *et al.* Prognostic value of PLCE1 expression in upper gastrointestinal cancer : a systematic review and meta-analysis. *Asian Pac J Cancer Prev.*, **15**, 9661–9666 (2014).

48. Cui X, *et al*. Comprehensive bioinformatics analysis of the mRNA profile of PLCE1 knockdown in esophageal squamous cell carcinoma. *Mol. Med. Rep.*, **16**, 5871–5880 (2017).

49. Tyutyunnykova A, *et al*. The controversial role of phospholipase C epsilon (PLCε) in cancer development and progression. *J. Cancer*, **8**, 716–729 (2017).

50. Danielsen SA, *et al.* Phospholipase C isozymes are deregulated in colorectal cancer - insights gained from gene set enrichment analysis of the transcriptome. *PLoS One*, **6**, e24419 (2011).

51. Wang X, *et al.* Phospholipase C epsilon plays a suppressive role in incidence of colorectal cancer. *Med. Oncol.*, **29**, 1051–1058 (2012).

52. Wang Q, *et al*. Association between phospholipase C epsilon gene (PLCE1) polymorphism and colorectal cancer risk in a Chinese population. *J. Int. Med. Res.*, **42**, 270–281 (2014).

53. Abnet CC, *et al.* A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nat. Genet.*, **42**, 764–768 (2010).

54. Wang LD, *et al.* Genome-wide association study of esophageal squamous cell carcinoma in Chinese subjects identifies susceptibility loci at PLCE1 and C20orf54. *Nat. Genet.*, **42**, 759–763 (2010).

55. Palmer AJ, *et al.* Genetic variation in C20orf54, PLCE1 and MUC1 and the risk of upper gastrointestinal cancers in Caucasian populations. *Eur. J. Cancer Prev.*, **21**, 541–544 (2012).

56. Sharma KL, *et al.* Association of potentially functional genetic variants of PLCE1 with gallbladder cancer susceptibility in North Indian population. *J. Gastrointest. Cancer*, **44**, 436–443 (2013).

57. Li FX, *et al.* Association of 10q23 with colorectal cancer in a Chinese population. *Mol. Biol. Rep.*, **39**, 9557–9562 (2012).

58. Kupcinskas J, *et al.* Common genetic variants of PSCA, MUC1 and PLCE1 genes are not associated with colorectal cancer. *Asian Pac. J. Cancer Prev.*, **16**, 6027–32 (2015).

59. Ezgi O, *et al*. Genetic variations in phospholipase C-epsilon 1 (PLCE1) and susceptibility to colorectal cancer risk. *Biochem. Genet.*, **54**, 826–829 (2016).

60. Bosman FT, *et al*. Tumours of the colon and rectum. *WHO Classification of Tumours of the Digestive System*. 4th Edition. Geneve, World Health Organization (2010).

61. Umar M, *et al*. PLCE1 rs2274223 A>G polymorphism and cancer risk: A meta-analysis. *Tumor Biol.*, **34**, 3537–3544 (2013).

62. Johnston PA, *et al*. STAT3-based anticancer strategies. *Molecular Interventions*, **11**, 18–26 (2011).

63. Xue W, *et al*. Association between PLCE1 rs2274223 A > G polymorphism and cancer risk: Proof from a meta-analysis. *Sci. Rep.*, **5**, 1–7 (2015).

64. Zhou R, *et al*. Phospholipase C ε-1 gene polymorphisms and prognosis of esophageal cancer patients from a high-incidence region in northern China. *Mol. Clin. Oncol.*, **8**, 170–174 (2017).

65. Slatkin, M. Linkage disequilibrium — understanding the evolutionary past and Mapping the Medical Future. *Nat. Rev. Genet.*, **9**, 477–485 (2016).