UNIVERSIDADE DE LISBOA

Faculdade de Medicina

Gene Expression and Phenotypical Analysis of Winner and Loser Cells Engaged in Cell Competition

Alexandre Miguel Guedes da Silva Afonso

Orientadores:

Doutora Ana Cristina Alves Queirós

Professora Doutora Sandra Cristina Cara de Anjo Casimiro

Dissertação especialmente elaborada para obtenção do grau de Mestre em Oncobiologia

Especialização em Investigação em Oncobiologia

2021

UNIVERSIDADE DE LISBOA

Faculdade de Medicina

Gene Expression and Phenotypical Analysis of Winner and Loser Cells Engaged in Cell Competition

Alexandre Miguel Guedes da Silva Afonso

Orientadores:

Doutora Ana Cristina Alves Queirós

Professora Doutora Sandra Cristina Cara de Anjo Casimiro

Dissertação especialmente elaborada para obtenção do grau de Mestre em Oncobiologia

Especialização em Investigação em Oncobiologia

2021

A impressão desta dissertação foi aprovada pelo Concelho Científico da Faculdade de Medicina de Lisboa em reunião de 23 de Março de 2021.

Agradecimentos

Fico eternamente grato a todos aqueles que de alguma forma contribuíram para a realização deste projeto de final de mestrado.

Agradeço veementemente, em primeiro lugar, à Doutora Ana Queirós, por ter aceitado ser a minha orientadora e pelo seu constante sentido crítico e imensurável apoio demonstrados logo desde o início da realização do projeto.

Quero também agradecer a todos os docentes da Faculdade de Medicina da Universidade de Lisboa, pela disponibilidade e empenho oferecidos na nossa formação académica. Deixo também os meus parabéns aos Professores Doutores João Ferreira, Luís Costa, e Sandra Casimiro pelo fantástico trabalho na criação e promulgação do Mestrado em Oncobiologia, trabalho esse que nos deu a nós alunos uma exímia educação.

Agradeço ao Moreno Lab na Fundação Champalimaud e a todos os seus membros por me terem recebido e por toda a ajuda prestada no desenvolvimento desta dissertação.

Deixo também os meus agradecimentos aos meus tios, que gentilmente me acolheram em sua casa como seu próprio filho, dando-me assim a oportunidade de frequentar este mestrado.

Agradeço também à minha família, que tanto me apoiou na realização deste projeto, dando especial destaque aos meus avós por terem sempre acreditado em mim.

A todos estes e muitos outros, muitíssimo obrigado!

Resumo

O cancro continua a ser um dos maiores focos da comunidade científica, o que resulta em milhares de estudos publicados anualmente com o intuito de mitigar o aparecimento e avanço desta doença, já que está associada a uma crescente taxa de mortalidade. Diversos destes estudos focam-se no processo de oncogénese e nos mecanismos que promovem ou previnem a mesma. Recentemente, o estudo da Competição Celular (CC) revelou que este processo possui um papel importante durante a progressão tumoral. A CC é um mecanismo através do qual células mais aptas ("*winners*") são selecionadas, por meio de eliminação de células menos aptas ("*losers*"), e é descrito como estando envolvido em vários processos fisiológicos, como a homeostasia e regeneração dos tecidos e também em algumas patologias como o cancro e doenças neurodegenerativas.

A função da CC durante a oncogénese é, contudo, dependente do contexto, podendo atuar como um supressor ou promotor tumoral. No primeiro caso, a CC identifica células que possuam alterações potencialmente carcinogénicas, como mutações no *HRAS* e a sobre expressão de *ERBB2*, e promove a sua eliminação, impedindo assim eventuais transformações neoplásicas. Por outro lado, a CC pode ser manipulada pelas células tumorais para estas adquirirem mais espaço para proliferar, através da eliminação das células não tumorais circundantes. Coincidentemente, vários marcadores de CC têm sido associados a um pior prognóstico dos pacientes. Assim sendo, torna-se imperativo compreender os mecanismos de CC presentes no desenvolvimento tumoral. No entanto, ainda são desconhecidos os fatores que condicionam as funções da CC no cancro, assim como os principais reguladores da mesma.

Com base nas evidências referidas, o objetivo deste projeto foi a identificação de genes cuja expressão desencadeasse fenómenos de CC em células de cancro colorretal, e determinar os impactos fenotípicos da mesma nas células. Desta forma, foi usado um modelo de co-cultura entre duas linhas celulares de cancro colorretal, LoVo e LS 174T. Foi observado um aumento significativo na apoptose das células LoVo após 48h de co-cultura, enquanto que as LS 174T apresentaram um aumento de viabilidade e proliferação, em comparação com as respetivas condições de monocultura. De acordo com a definição canónica de CC, os resultados sugerem que as células LoVo, quando

colocadas em co-cultura com as células LS 174T, são classificadas como "*losers*" e eliminadas, ao passo que as últimas assumem o papel de "*winners*". Esta eliminação é dependente das proporções de células "*winner*" em co-cultura, sendo que a taxa de apoptose é amplificada quando estas se encontram em maioria. No entanto, um aumento no número total de células em cultura, mantendo as proporções, resulta numa mitigação da eliminação das *"loser".* Estes resultados não só comprovam que esta eliminação é dependente da área da superfície de contacto entre "*winner"* e "*loser"*, e não de fatores humorais libertados pelas *"winner"*, como também indicam que a mesma depende do espaço disponível para a proliferação celular. Curiosamente, as "*loser"* apresentaram uma eliminação menos pronunciada às 72h, e as "*winner*" mantiveram a sua viabilidade acrescida, sugerindo que a duração do efeito da CC é diferente nos dois tipos de células.

De forma a averiguar se a apoptose observada é ativada por caspases, o mesmo ensaio foi repetido na presença de emricasam, um inibidor de caspases. A eliminação das células "*loser*" foi completamente inibida, apesar da viabilidade das células "*winner*" ter permanecido inalterada. Surpreendentemente, as "*loser"* demonstraram também uma maior viabilidade. Estes dados propõem que as células "*loser*" são eliminadas num processo mediado por caspases, e indicam que podem estar presentes fatores secretados que exercem o seu efeito nas células "*winner*", mas também nas células "*loser"* quando a apoptose é inibida.

Coincidentemente, já foi previamente descrita uma proliferação compensatória dependente de apoptose, na qual a morte das células "*loser*" resulta numa maior proliferação das células "*winner*". No entanto, a inibição da apoptose não alterou a taxa de proliferação das células "*winner*", o que sugere que a proliferação não é dependente da apoptose das "*loser"*.

Devido à aparente correlação entre vários marcadores de CC e a presença de metástases, a capacidade de migração das duas linhas celulares em co-cultura foi avaliada. Os diversos ensaios revelaram um aumento da migração das células "*winner*", acoplado a uma marcada sobre-expressão de vimentina, o que sugere que a CC intervém na transição epitélio-mesênquima.

A CC pode atuar durante todas as etapas da doença oncológica e afetar a resposta à terapêutica, pelo que foi testada a resposta à quimioterapia no modelo de

co-cultura. A presença de oxaliplatina causou um aumento da viabilidade das células LoVo e, em contrapartida, um incremento na apoptose das células LS 174T. Este efeito poderá dever-se à maior taxa de proliferação das células "*winner*" em co-cultura, o que as torna mais vulneráveis à citotoxicidade da quimioterapia usada. Porém, é possível que fatores externos alterem o contexto da CC, pelo que o "*fitness"* das células é relativo e dependente do ambiente no qual estas se encontram. Assim sendo, nestas condições, as LoVo agem como "*winner*", eliminando as LS 174T que agora se comportam como *"loser".* A ideia de que algo extrínseco às células influência a CC é inovadora, e evidencia que esta pode contribuir à eliminação de células mais invasivas aquando do tratamento com agentes quimioterapêuticos.

Com o intuito de identificar quais os genes, e respetivas vias de sinalização, potencialmente implicados em CC, foi realizada uma análise do transcriptoma, através de RNA-Seq. Para tal, a expressão génica das células "*winner*" e "*loser*" em co-cultura foi comparada com a respetiva monocultura. Após análise bioinformática e validação independente por RT-qPCR dos genes diferencialmente expressos, 8 genes foram identificados como sendo sobre-expressos nas células "*winner*", e 3 genes foram identificados como sub-expressos nas células "*loser*", quando as mesmas se encontram em co-cultura. O *knock-down* (KD) dos 8 genes nas células "*winner*" permitiu avaliar os seus efeitos fenotípicos na co-cultura. O KD dos genes *AQP3*, *MYT1* e *NRIP1* em particular bloqueou por completo a eliminação das células "*loser*", sugerindo que o aumento da expressão destes é crucial à eliminação das "*loser"* em CC. Adicionalmente, genes inseridos em vias de sinalização previamente implicadas na oncogénese foram igualmente validados por RT-qPCR. Os resultados expõem um aumento da sinalização mTOR e JAK/STAT nas células "*winner*", e uma diminuição destas nas células "*loser*", o que poderá explicar as diferenças observadas na viabilidade e proliferação celular.

Este projeto permitiu associar, pela primeira vez, a CC à migração celular e à modulação da resposta à quimioterapia em cancro do cólon. Adicionalmente, foi possível a identificação de novos genes cuja sobre-expressão é essencial para a eliminação das células "*loser*". Este projeto constitui não só uma valiosa adição ao conhecimento atual sobre os mecanismos que regem a CC, como também poderá contribuir para o desenvolvimento de potenciais terapêuticas que a bloqueiem, atrasando a progressão tumoral.

Palavras-Chave: Competição Celular; Sequenciação de RNA; Migração Celular; Progressão Tumoral; Resposta à Quimioterapia.

Abstract

Cancer remains a major focus of the scientific community, despite the growing efforts to thwart this pathology. Research continues to progress in order to counter the rising mortality rate, striving to better understand the natural process of oncogenesis. A very exciting topic is the implication of cell competition (CC) in cancer, where it can act as a tumour suppressor or promoter. Cell competition, a homeostasis mechanism, is responsible for comparing cells with their counterparts and promoting the elimination of the lesser fit ones. Although CC aims to maintain tissue health, several of its markers have been linked to a worse prognosis and metastasis development. Since CC may play a role in both cancer prevention and progression, understanding the molecular basis of this process would be invaluable. Thus, this project intended to unveil new genes which expression triggers CC in colorectal cancer cells, and to uncover the phenotypic impact that these cells suffer when engaging in CC.

With this aim, a co-culture assay using two colorectal cancer cell lines, LoVo and LS 174T, was used. These cell lines were found to engage in CC, with LoVo being the "loser" cells, thus suffering elimination, and LS 174T the "winner" cells, with enhanced survival. The loser cells' elimination was found to be caspase dependent, as the use of a pan-caspase inhibitor was sufficient to block the effect.

To investigate other ways that CC influenced cells, a proliferation and a migration assay were carried out. During CC, winner cells exhibited a higher proliferation rate, which was not dependent on the apoptosis of the loser cells, while loser cells exhibited a lower proliferation. It was also found that winner cells migrated more in co-culture, whilst loser cells migrated less. Both these results suggest that CC transforms the winner cells into a more aggressive phenotype.

Next, to study the influence of CC in response to cancer treatment, the co-culture assay was also repeated in the presence of oxaliplatin. A complete ablation of loser cell elimination was observed, with an increase in winner cell apoptosis, which can indicate that CC improves chemotherapy response in more invasive cells.

To discover the genes implicated in the signal transduction of CC, an RNA-Seq was performed on the competing cells. This gene expression analysis revealed several genes which were upregulated on the winner cells during CC. These genes were knocked down (KD) in winner cells using shRNAs, followed by co-culture with non-altered loser

cells, to test if their upregulation was necessary for CC occurrence. The knockdown of *AQP3*, *MYT1* and *NRIP1* abrogated CC, suggesting that the upregulation of each of them on the winner cells is necessary for the event.

Overall, this project revealed new ways in which CC affects cell behaviour and unveiled novel genes necessary for cell competition, thus suggesting new targets for the future development of therapies with the potential to improve cancer prevention and hinder its progression.

Keywords: Cell Competition; RNA-Sequencing; Cell Migration; Tumour Progression; Chemotherapy Response.

Abbreviations, acronyms and symbols

% - Percentage

°C – Degree Celsius

AKT - Protein kinase B

APC - Adenomatous polyposis coli

AQP3 – Aquaporin 3

Aβ – Amyloid Beta

BMP – Bone morphogenetic protein

BRCA - Breast invasive carcinoma

CACFD1 - Calcium channel flower homolog

CC – Cell Competition

CDC42 - Cell division control protein 42 homolog

CDH1 - Cadherin-1

cDNA – Complementary DNA

cm – Centimetre

COAD – Colon adenocarcinoma

COL6A3 - Collagen alpha-3(VI) chain

Ct – Cycle threshold

CYP1A1 - Cytochrome P450, family 1, subfamily A, polypeptide 1

DCAF1 - DDB1- and CUL4-associated factor 1

DCN – Decorin

DFS – Disease-Free survival

DMEM - Dulbecco's Modified Eagle's Medium

DMSO – Dimethyl sulfoxide

DNA - Deoxyribonucleic acid

Dpp – Decapentaplegic

ECM – Extracellular matrix

EDAC – Epithelial defence against cancer

EdU - 5-ethynyl-2´-deoxyuridine

EGFR – Epithelial growth factor receptor

EMT – Epithelial mesenchymal transition

ERBB2 - Receptor tyrosine-protein kinase erbB-2

Et al. – Et alia

FACS – Fluorescent activated cell sorting

FBS – Fetal bovine serum

FPKM-UQ - Fragments per kilobase of transcript per million mapped reads upper quartile

GC – Guanine-cytosine

GDC – Genomic Data Commons

GEPIA – Gene Expression Profiling Interactive Analysis

GFP – Green fluorescent protein

GPX2 - Glutathione peroxidase 2

h – Hours

H_0 – Null hypothesis

HDI – Human development index

HEK 293 - Human embryonic kidney 293

HRAS - HRas proto-oncogene

Il-7 - Interleukin 7

IQR – Interquartile region

ITH – Intra-tumour heterogeneity

JAK – Janus kinase

JNK - c-Jun N-terminal kinases

KIRC - Kidney renal clear cell carcinoma

KRAS - Kirsten rat sarcoma virus

LAMC2 - Laminin subunit gamma-2

LB – Luria Bertani culture medium

LEDs - Light-emitting diode

LIHC - Liver hepatocellular carcinoma

LLGL1 - Lethal(2) giant larvae

LUAD - Lung adenocarcinoma

MAPK1 - Mitogen-activated protein kinase 1

MDCK – Madin-Darby canine kidney

mg – Miligrams

MGP - Matrix Gla protein

mL – Millilitre

MSMO1 - Methylsterol monooxygenase 1

MTOR – Molecular target of rapamycin

MUC4 - Mucin 4

MYC - Myc proto-oncogene protein

MYT1 - Myelin transcription factor 1

NCBI - National Center for Biotechnology Information

NRIP1 - Nuclear receptor-interacting protein 1

NS – Non significant

OS – Overall survival

PAAD - Pancreatic adenocarcinoma

PBS - Phosphate buffered saline

PDK4 - Pyruvate dehydrogenase kinase isozyme 4

PI – Propidium iodide

PRAD - Prostate adenocarcinoma

PSA – Prostate specific antigen

PTGS2 - Prostaglandin-endoperoxide synthase 2

Px – Pixels

RNA – Ribonucleic acid

RNA-Seq – Ribonucleic acid sequencing

RPL24 - 60S ribosomal protein L24

RPL32 - 60S ribosomal protein L32

rpm – Rotations per minute

RT-qPCR – Real-Time quantitative polymerase chain reaction

SCRIB – Scribble

SEM - Standard error of the means

shRNA – Short hairpin RNA

SKIL - SKI Like proto-oncogene

SLC34A2 - Sodium-dependent phosphate transport protein 2B

SLC6A14 - Sodium- and chloride-dependent neutral and basic amino acid transporter B(0+)

SPARC - Secreted protein acidic and rich in cysteine

SPINK4 - Serine protease inhibitor Kazal-type 4

SRC - Proto-oncogene tyrosine-protein kinase Src

STAT - Signal transducer and activator of transcription

STAT3 - Signal transducer and activator of transcription 3

SULF1 - Sulfatase 1

SYNPR – Synaptoporin

T-All - T-cell acute lymphoblastic leukaemia

TCGA – The Cancer Genome Atlas

TdTomato - Tandem dimeric derivative of DsRed

THBS2 - Thrombospondin-2

tkv – thickveins

TP53 - Tumor protein P53

TPM - Transcripts per kilobase million

TRPM8 - Transient receptor potential cation channel subfamily M (melastatin) member 8

UCSC - University of California Santa Cruz

UGT1A1 - UDP-glucuronosyltransferase 1A1

UGT1A10 - UDP-glucuronosyltransferase 1A10

UV – Ultraviolet

VIM – Vimentin

VPS25 - Vacuolar protein-sorting-associated protein 25

vs - Versus

wg - wingless

WNT1 - Proto-oncogene protein Wnt-1

WT – Wild Type

YAP - Yes-associated protein 1

µL – Microliter

µm – Micrometer

µM – Micromolar

List of Figures

List of Tables

Table of contents

1. Introduction

1.1. The Burden of Cancer

Albeit the ongoing efforts, cancer still remains one of the world's most impactful diseases, with a growing incidence worldwide, due to a complex interaction of several factors such as longer life expectancy and *westernization* of middle or third world countries[1-3] (Figure 1).

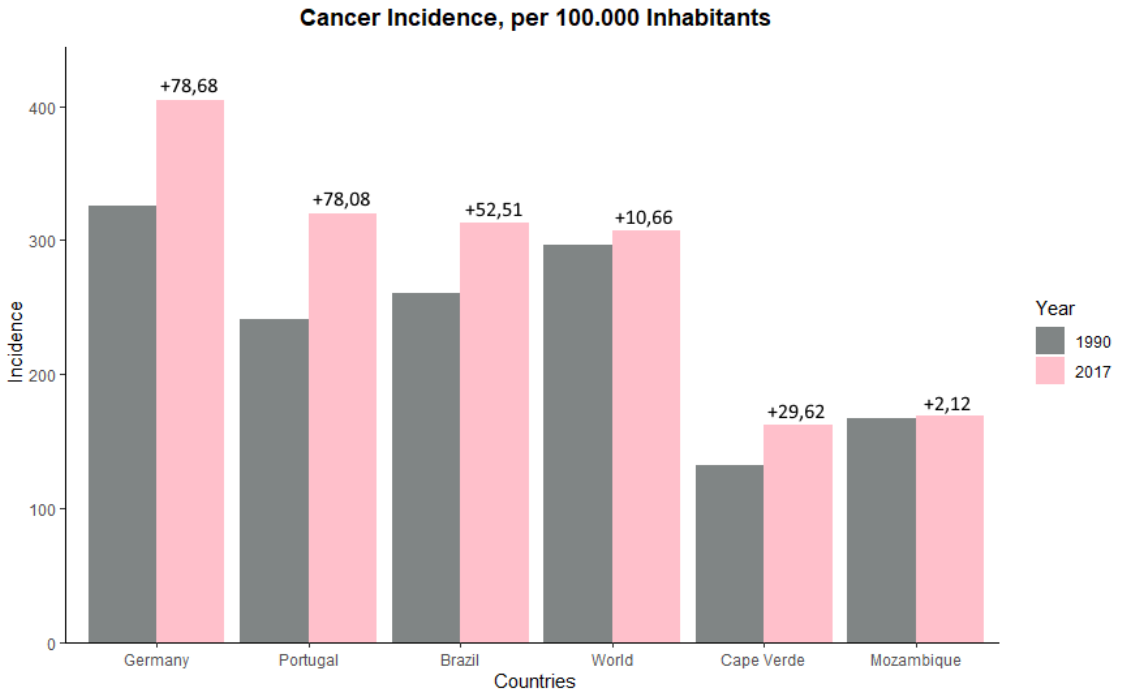


Figure 1. Rising cancer cases per 100.000 people. The countries displayed on the image are ordered by, and were chosen as representative of different levels of Human Development Index (HDI) as described by the United Nations[4] (Germany-Very High HDI; Portugal-Very High HDI; Brazil-High HDI; World-Representing the Global Average; Cape Verde-Medium HDI; Mozambique-Low HDI). The numbers on the bars represent absolute change per 100.000 people. Adapted from Our World in Data[2,3].

The same tendency is reflected in Portugal, with the age-standardized number of new cases per 100.000 people increasing from 241,4 in 1990, to 319,45 in 2017, representing a 32,3% increase[2,3] (Figure 1). This rising incidence is accompanied by the growing significance of cancer as a cause of death, in spite of better survival rates[5], with both factors partly explained by the declining mortality due to cardiovascular diseases, including stroke or coronary heart disease, which have long had the highest share of mortality in developed nations[1,6,7] (Figure 2, data pertaining to Portugal only), and lower mortality rates of infectious diseases in third-world countries[1,6,7].

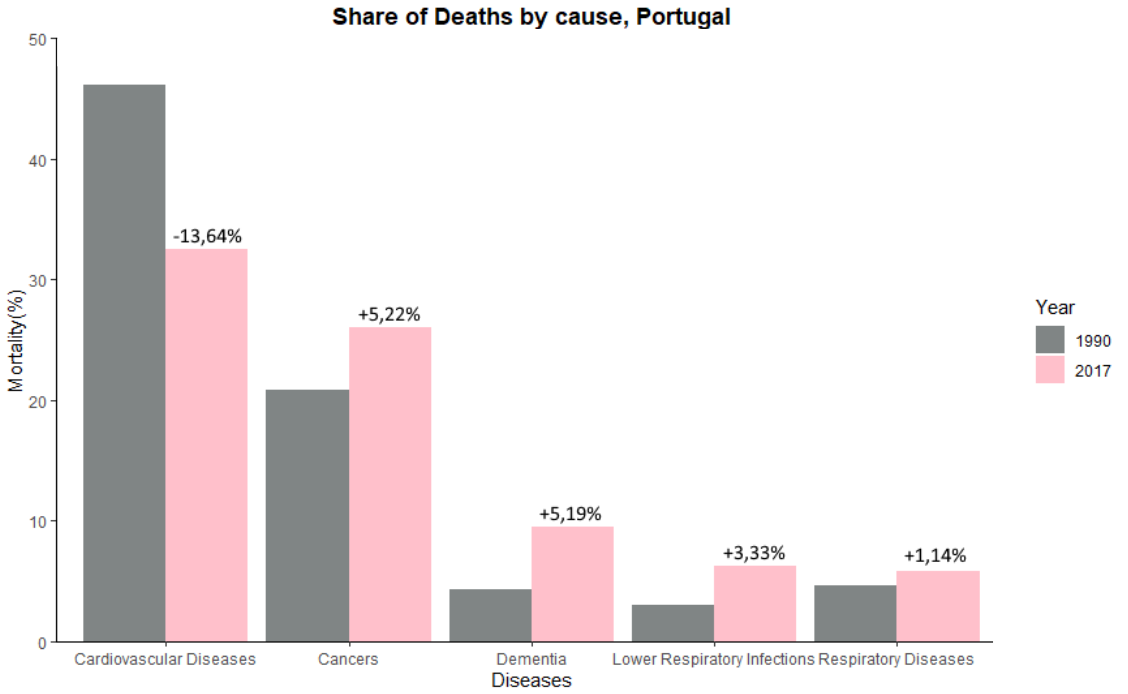


Figure 2. The rising burden of cancer mortality in Portugal. The causes displayed on the image are the leading causes of death in the country. The numbers on the bars represent the absolute change in the share of mortality. Adapted from Our World in Data[3,7].

Another phenomenon that can explain this rise is the onset of new methods which allowed for earlier and more sensitive diagnosis[1,5]. However, this only explains a transient effect, as cases are now detected in early stages or even sub-clinically, instead of at a later stage, where they would eventually be detected using older methodology[1,5]. A proper example of this is the approval of the prostate-specific antigen (PSA) test, which resulted in a higher detection rate of early stage cancers and a decline in late stage ones[8]. Given that there is a clear correlation between detection stage and likelihood of survival, these novel methods decrease cancer related mortality at the expense of incidence[1,5]. Nevertheless, this can lead to overdiagnosis or to the report of subclinical disease, that wouldn't clinically manifest in the patient's expected lifetime[9]. Contrarily, in developing nations, there is evidence of a vast underreport of cancer incidence and mortality[10,11], suggesting that the observed upsurge is probably higher than previously stated.

The epidemiological distribution of different cancer types presents itself heterogeneously, depending on geography (UV radiation exposition), economic development (healthcare), and religious/cultural factors (diet and condom usage)[5].

Generally, countries with higher Human Development Index (HDI) display higher incidence rates of lung cancer (includes tracheal and bronchus cancers) and colorectal cancer (includes colon and rectal cancers)[12], consistent with higher consumption of tobacco[13,14] and processed meat[15], known risk factors for these cancers, respectively[16,17] (Figure 3). Conversely, countries with lower HDI show high rates of cervical and liver cancers, resulting from a deficient vaccination coverage against the oncogenic viruses that represent the main causes for these neoplasms[18-20] (Figure 3).

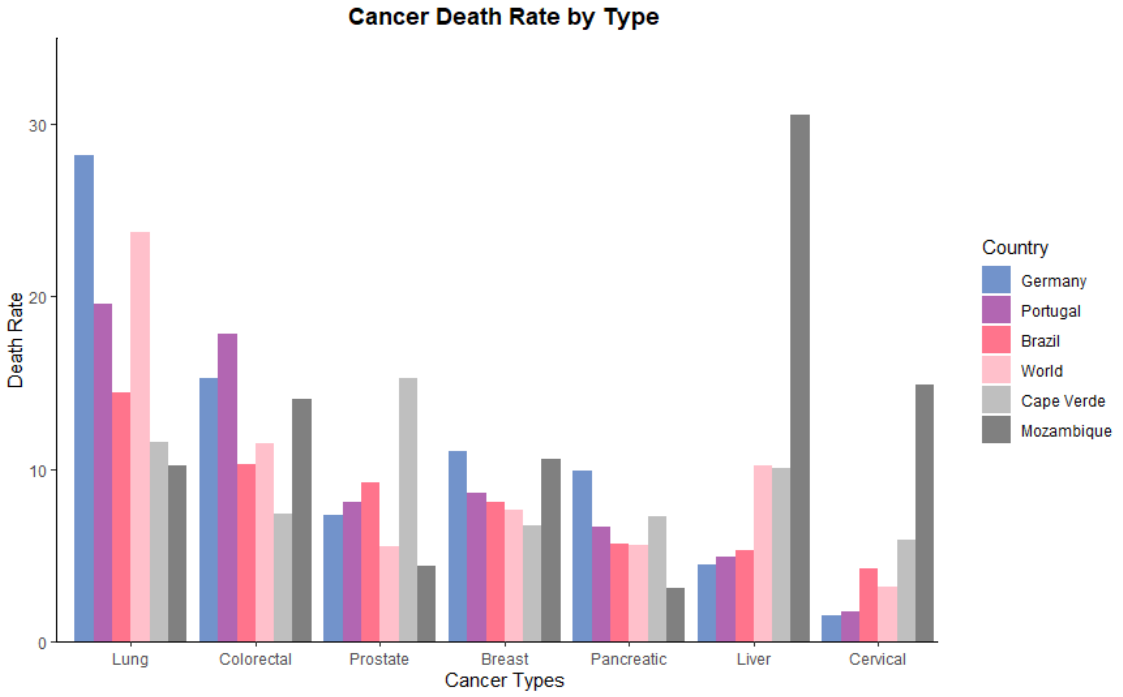


Figure 3. Differential death rates of cancers by primary location. The countries displayed on the image are ordered by, and were chosen as representative of different levels of Human Development Index (HDI) as described by the United Nations[4] (Germany-Very High HDI; Portugal-Very High HDI; Brazil-High HDI; World-Representing the Global Average; Cape Verde-Medium HDI; Mozambique-Low HDI). Adapted from Our World in Data[3,12].

Regardless of this disparity between nations at opposite ends of the spectrum, the majority of cancers are attributed to preventable causes[12,21], and therefore, can be acted upon, either by changing lifestyles, or by improving vaccination programs[1,5]. Furthermore, although cancer mortality rates have been dwindling in the past few years[5], the resulting morbidity from the disease, and sometimes the treatment[22,23], is still a main cause of concern, negatively impacting cancer patients' and survivors' quality of life[1,24]. Accordingly, the most preeminent approach to decrease cancer derived

morbidity and death might be one that enforces on this preventable incidence, instead of directly acting on the mortality *per se*[25].

1.2. Cancer Mortality and Emergence of Resistances

Conforming to the World Health Organization, cancer can be defined as "(…) a large group of diseases that can start in almost any organ or tissue of the body when abnormal cells grow uncontrollably, go beyond their usual boundaries to invade adjoining parts of the body and/or spread to other organs."[25]. It is a disease with a varied prognosis, depending on several factors, and on the local of primary origin[1,5] (Figure 4, data pertaining to Portugal only). Some have a 5-year survival rate above 75%, usually those that are frequently detected early and/or allow for surgical resection (ex: prostate)[1,5,26], while others are marked by a low 5-year survival rate, those which are only diagnosed when the disease is already advanced, generally with several metastasis (ex: pancreas)[5,27]. Notwithstanding, the stage of disease is not the only hindrance in cancer survival, with resistance to therapy being another hurdle[28], wherein intra-tumour heterogeneity (ITH) plays a major role[29-32].

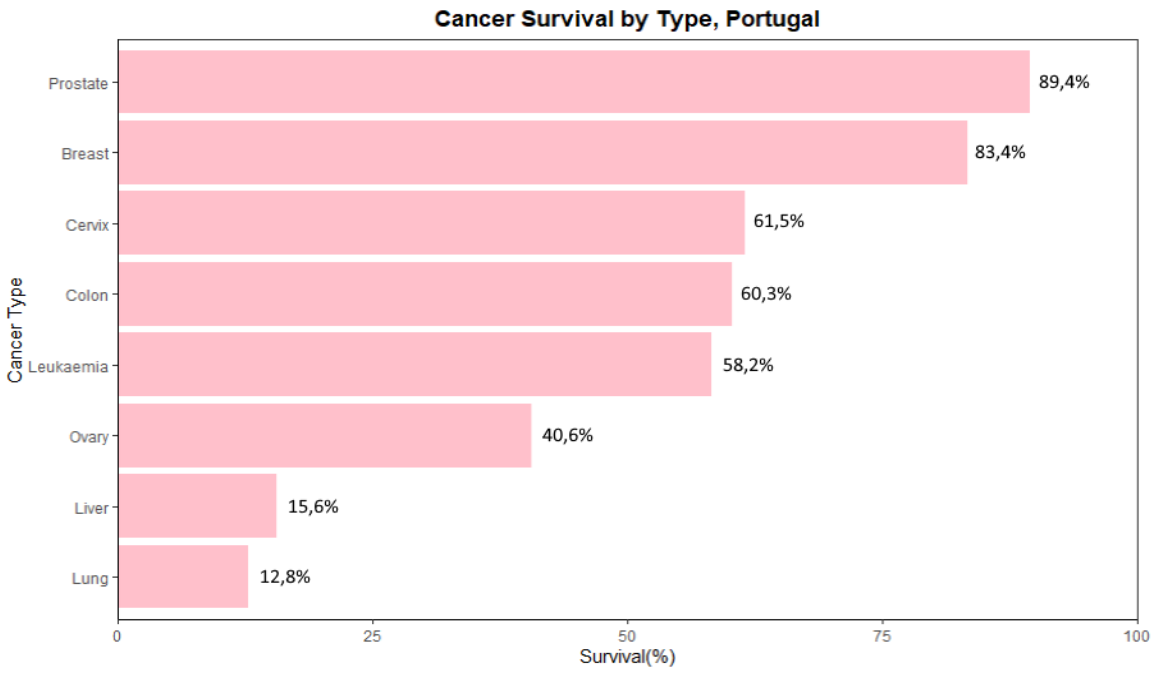


Figure 4. Disparate 5-year survival of several cancers in Portugal. Graph shows percentage of patients who were alive 5 years after diagnosis. The numbers on the bars represent Survival (%). Data from 2009. Adapted from Our World in Data[33,34].

As the oncogenic process progresses, each individual cancer cell can accumulate stochastic genetic and epigenetic alterations, culminating on a tumour that is not comprised of a single genetic identity, but of a mosaic of different subclones, each with their own mutations, some of which conferring therapy resistance[30,31]. These resistant subclones, due to the heavy pressure from the therapeutic, are then "selected" and survive, preventing the totality of the tumour from being eliminated[31,35,36]. These subclones might even impart their advantage to others, with some resistant cells having the ability to convey chemoresistance to their neighbours, in a paracrine manner[37]. This tumour mosaicism accentuates the importance of using combination strategies to circumvent this issue, as there is a lower change of existing cells being resistant to both drugs, especially if used simultaneously as to prevent acquisition of novel resistance mutations[29-32,38].

1.3. Intra-tumour Heterogeneity

It has long been postulated that the Darwinian evolution principles[39] could be applied in the context of carcinogenesis, with the pre-tumoural cells having a higher proliferative ability, and thus being selected in detriment of their wild type neighbours[32,40,41]. In consonance with this clonal evolution theory, as cells garner mutations and as cancer progresses, the resulting phenotype would be gradually more aggressive[40] (Figure 5A).

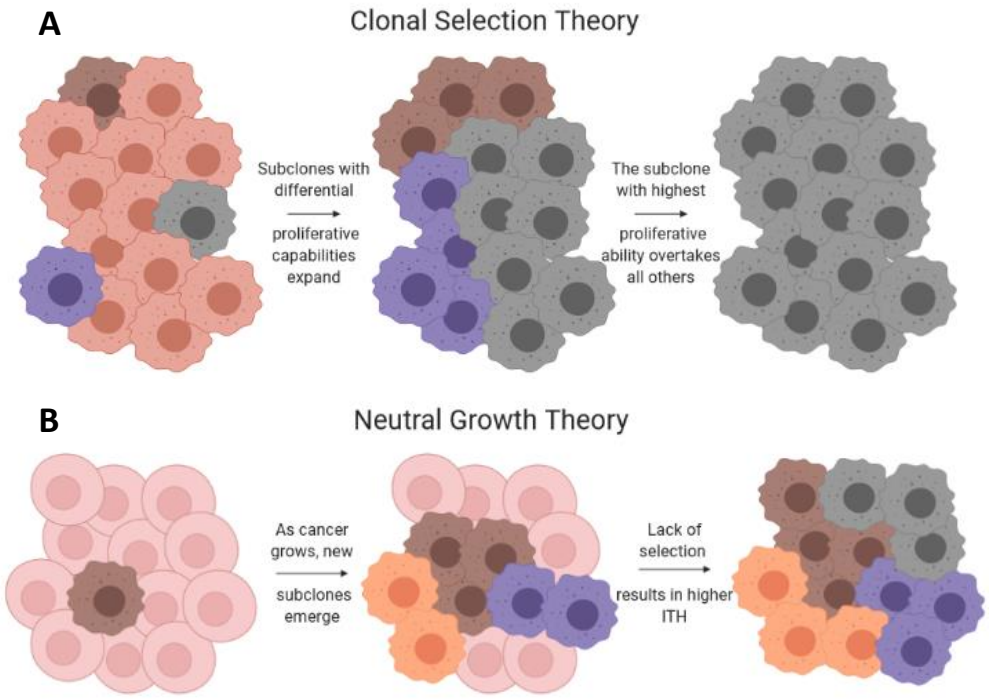


Figure 5. Different theories of cancer progression. (A) Clonal Selection Theory of Cancer Growth. According to the Darwinian principles, as more proliferative subclones arise, these overtake and outcompete lesser ones, eventually becoming the bulk to the tumour. (B) Neutral Theory of Cancer Growth. As cancer cells divide, they acquire new mutations that, despite translating into differential fitness, are not subjected to selection. As such, those subclones become a constituent part of the tumour, leading to a higher intra-tumour heterogeneity (ITH). The colours represent different subclones with distinct fitness.

Yet, this rationale fails to explain the high degree of intra-tumour heterogeneity usually seen, far surpassing what would be expected[42]. Instead, it is proposed that most cancers grow under the neutral law where the expected genetic diversity would be much higher, consistent with the observed phenotype[35,36,42] (Figure 5B). This hypothesis proposes that after the initial driver mutations, any descendant subclones grow at similar rates, without one dominating over others[36], save for high selective pressures, such as treatments[43]. Due to spatial constraints and naturally high division rates, most tumour cells obey the neutral growth law, explaining how adjacent populations with distinct fitness can seemingly coexist[35], with only the cells at the boundary between them, having the possibility to engage in selection events[42]. As such, the major situations where traditional Darwinian selection[39] is predominant are in the initial steps of oncogenesis and after cancer drugs administration[35,36,42], and taking into consideration the importance of these two stages, unfolding the various ways in which this selection occurs would prove invaluable as a research focus. In that regard, a noteworthy mechanism, known as Cell Competition (CC), might act as a main driver of this selection[44].

1.4. Cell Competition

CC is a phenomenon in which more capable cells, dubbed the “winners”, outcompete the less fit ones, the “losers”, resulting in their elimination[45,46] (Figure 6).

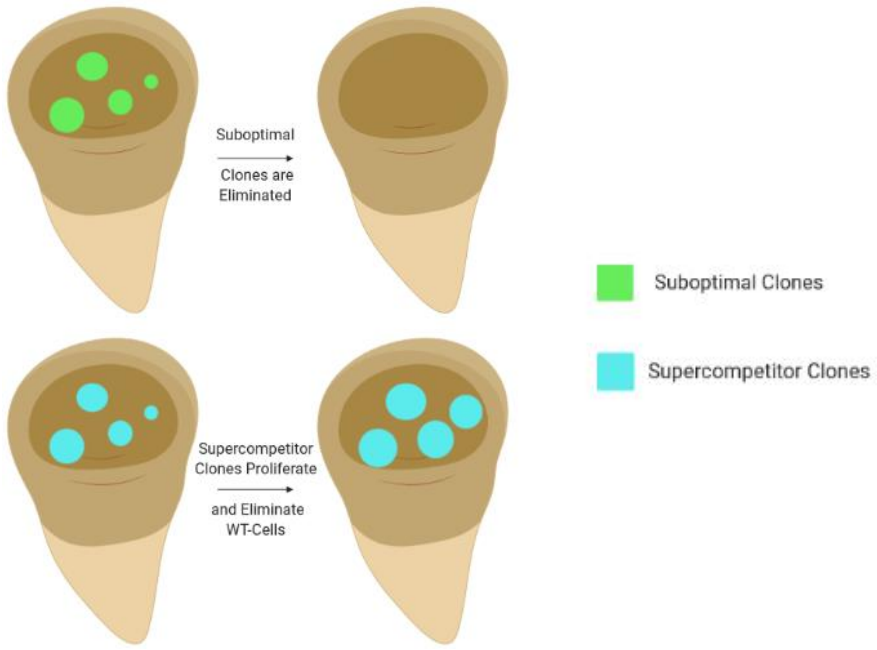


Figure 6. Example of cell competition in *Drosophila* wing disc. In a mosaic Wing Disc, when clones with differential fitness are present, they are either eliminated, if they are suboptimal comparing to the WT tissue (above), or proliferate, if they have higher levels of fitness relative to the WT tissue (below).

First discovered in the 70's when studying a *Minute* mutation[47], CC is a biological process found to be implicated in various situations, from natural tissue homeostasis[48], wound healing[49] and aging[46], to diseases like Alzheimer's[50] or cancer[51-54]. When in heterogeneity (M+/M-), these *Minute* mutants presented with a deficient ribosomal protein, which led to inferior proliferation rates[47,55,56], but were otherwise viable (individuals reached adulthood)[47]. Still, when these cells where surrounded by their wild-type (WT) counterparts (M+/M+), they were promptly eliminated[47], which sparked the belief that these heterozygous cells underwent apoptosis when near WT cells, later confirmed to be the case[45]. This differs from other elimination scenarios because cells aren't intrinsically winners or losers, they require some degree of interaction with one another for the less fit to be eliminated, both being viable within their own homotypic populations[45-47]. Although it was originally described in *Drosophila melanogaster* imaginal discs[47], it has since been verified to be conserved across several tissues and species, including model organisms like *Mus musculus*[57,58], or human cell lines[59,60], with many other mutations besides *Minute* known to originate CC (Table 1). All in all, several variations of this process have been characterized, each with their own particularities.

Table 1. List of CC related genes. Non-exhaustive list of genes found to trigger cell competition and their respective phenotype and brief summary of functions. The corresponding human homologue of the genes was found by using HomoloGene[79] and Fly Base[80].

Gene ID	*Phenotype*	*Observations*
rpL32[47]	When mutated, cells are eliminated[47].	The *Minute* mutation. Encodes a ribosomal protein in *Drosophila melanogaster*.
Dpp[45]	When in insufficient quantity, cells are eliminated[45].	Survival factor. The *Drosophila melanogaster* homologue of the human BMP family.
Scrib/SCRIB[61-63]	When knocked-down, cells are eliminated from the epithelium[61-63].	Polarity gene. Implicated in mechanical cell competition[62,63].
L(2)gl[59]	When knocked-out, cells are eliminated[59].	Polarity gene. *Drosophila* homologue of human *LLGL1*.
Vps25[64]	When mutated, cells are eliminated[64].	Vesicular protein. Leads to tumour growth when CC is blocked[64].
Fwe[65]	Its isoforms communicate the cell's relative fitness to its neighbours[65].	Calcium channel. *Drosophila* homologue of human *CACFD1*.
Mahj[59]	When knocked-down, cells are eliminated[59].	Cell cycle protein. *Drosophila* homologue of human *DCAF1*.
Myc[60,66,67]	Cells with overexpression eliminate WT tissue[60,66,67].	Transcription factor that leads to proliferation.
Yki/YAP1[68,69]	In Drosophila, cells with overexpression eliminate WT tissue[68]. In MDCK cells, overexpressing cells are extruded[69].	Transcriptional regulator that leads to proliferation.
Stat92E[70]	Cell with lower expression are eliminated by WT cells[70].	Transcription Factor. *Drosophila* homologue of the human STAT protein family.

	Cells with overexpression eliminated WT cells[70].	
wg[71]	When in insufficient quantity, cells are eliminated[71].	Survival factor. *Drosophila* homologue of the human *WNT1*.
Apc[72]	When mutated, cells eliminate WT cells[72].	Negative regulator of the WNT pathway.
Trp53/TP53[73,74]	Lower levels lead to outcompetition of stem cell progenitors[73]. Mutant-p53 expressing cells suffer necroptosis by WT MDCK cells[74].	Major tumour suppressor gene. Also plays a role in mechanical CC[63].
HRAS[75]	Cells with this mutation are extruded from the epithelium[75].	Mutation commonly found in cancer.
v-Src[76]	Cells with this mutation are extruded from the epithelium[76].	Retroviral oncogene. Encodes a tyrosine kinase receptor.
CDC42[77]	Cells with this mutation are extruded from the epithelium[77].	Regulation of cell cycle.
ERBB2[78]	Cells with this mutation are extruded from the epithelium[78]	Tyrosine kinase receptor. Oncogene.

1.4.1. Competition for Survival Factors

In this type of CC, cells compete for a survival factor, and those that possess a deficiency in competing for, or in transducing these survival signals, suffer apoptosis[45,46]. It was the first type of CC described, with the aforementioned *Minute* mutants being at a disadvantage to capture the survival factor Dpp[81], eventually resulting in their elimination[45,46]. This reduction of Dpp signaling in the prospective loser cells was sufficient to trigger apoptosis in a JNK-dependent manner[45]. In turn, the constitutive activation of the Dpp receptor tkv led to apoptosis of the surrounding WT tissue[82],

reiterating the importance of context in CC, with the same cells being able to phenotypically behave as losers or winners. CC also occurs with mutations in *Rpl24*, a *Mus musculus* riboprotein gene, with heterozygous cells for this mutation being outcompeted by non-mutants on a mosaicism scenario[58]. Although this event is akin to what is observed in *Drosophila*, there is no description of survival factors involved in this competition, contrasting with the role of Dpp in fly. Moreover, wg, another signaling factor[83,84], was found to trigger similar results, with wg insensitive clones being eliminated when bordering WT cells[71]. Withal, when mutant clones over-activating wg were induced, they instead induced apoptosis on the WT cells, becoming "Supercompetitors"[71]. These findings were all reported in mutants, but identical results were found in a murine study, which found that in the WT thymus, progenitor cells compete for IL-7, with younger progenitor cells outcompeting the older ones, leading to their elimination, constituting a natural homeostatic process[53,85].

1.4.2. Mechanically Induced Cell Competition

Mechanical forces have long been known to act as transducers of information[86], in a process known as mechanotransduction, where physical stimuli are translated into biological responses[86]. It is present on various circumstances, just as blood flow dynamics influencing endothelial homeostasis[87], the adaptive response to changes in cardiac function[88], and even the necessity of gravity to maintain bone integrity[89]. Remarkably, mechanotransduction has also been linked to CC[90]. When cells proliferate at distinct rates on a limited space, they generate mechanical stress[91-93], stretching and compressing the slower growing cells, and if these forces prove sufficient, and these cells are more sensitive than their fast growing neighbours, cell death/extrusion might ensue [91,94-97]. This is most evident on the *Drosophila* pupal notum where, as cells divide and crowding occurs, delamination follows, allowing proper maintenance of natural tissue growth[95,96]. This manner of cell competition seems to be conserved in mammals, with Madin-Darby canine kidney cells with knockdown of *SCRIB* being extruded from the epithelium when surrounded by WT cells[62,63], via p53 induced lethality[*63*]. In this scenario, the *SCRIB* deficient cells are more sensitive to physical compression, and through crowding, they end up being eliminated from the heterozygous tissue, resulting in an overall fitter population[62,63,90]. Since mechanical forces propagate through the

tissue, this could act as a long distance cell competition[91], diverging from the other types which are mostly cell-to-cell[46].

1.4.3. Fitness Fingerprints Comparison

A cell has the inherent ability to present their capacity to survive and proliferate, through presentation of a "*tag*". Each cell's tag is then compared to their neighbours' and those who exhibit a tag encoding for a lesser fitness status undergo apoptosis[65]. The most striking example of this event is the *Fwe* gene, which codes a transmembrane calcium channel[65]. Cells express different isoforms of the protein at different proportions according to their own fitness[65]. When a cell is unfit, it expresses more of Fwe^{Lose} isoforms, whereas winner cells express more Fwe^{Win} isoforms[65]. Cells that are adjacent can then compare the relative levels of their isoforms and those who have more loser tags are eliminated[65]. Also worthy of note is the upregulation of Sparc/Osteonectin in the loser cells, an extracellular glycoprotein that was found to transiently protect from apoptosis[98]. These different isoforms are proposed to be downstream of most competition scenarios, like the above mentioned *Minute* or *SCRIB* mutants[46]. When cells are deficient for these genes, they express Fwe^{Lose} isoforms on their membrane, being posteriorly eliminated when they come into contact with cells with a Fwe^{Win} tag[46]. A point of contention in this model, however, is how the signal transduction between the competing cells occurs. Whereas some studies state that soluble factors are involved, and are sufficient to result in CC[99], the majority of recent studies suggest that cell-cell contact is strictly necessary[65,100,101]. This machinery was also found to be involved in Alzheimer's disease, with neurons exhibiting Aβ damage being culled, which allowed for attenuation of the disease symptoms, protecting against cognitive decay[50,102,103]. CC's part in this disorder highlights its importance in the regulation and maintenance of normal organ function[104], paving the way for further research in other conditions.

1.5. Cell Competition in Cancer

Taking into account its functions in the elimination of unfit cells, and in the perpetuation of cells with "desirable" characteristics, CC has been proposed to act both as a tumour suppressor or tumour promoter mechanism[44,105]. In loose terms, cell

competition acts as two faces of the same coin, hindering proliferation of pre-tumoral clones or bolstering tumoral clones' mitotic and invasive capabilities, for they act as supercompetitors, being more viable than the WT surrounding tissue[44]. A thought-provoking part of this dichotomous behaviour is identifying the moment at which pre-tumoural cells start evading CC, and exploit it to fuel their own carcinogenesis[44] (Figure 7).

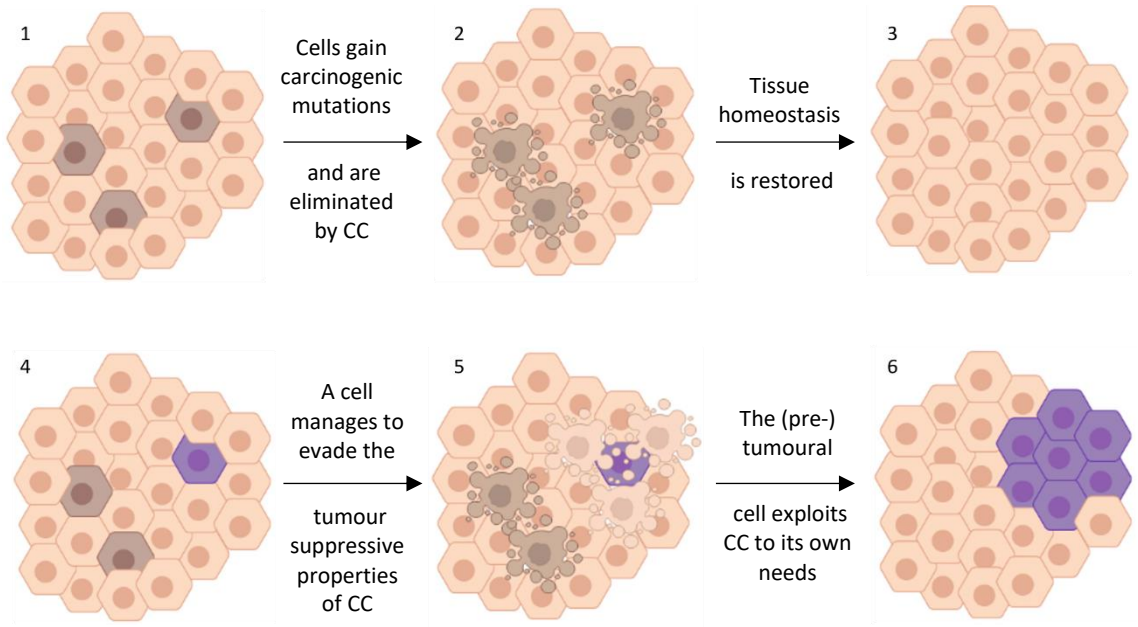


Figure 7. The role of cell competition in cancer. As cells gain potentially carcinogenic mutations, they are promptly eliminated by their WT neighbours, through CC, re-establishing homeostasis. When a cell develops the ability to evade the tumour suppressive properties of CC, it survives, being able to proliferate. When the cell harbours sufficient carcinogenic mutations, it may exploit CC to induce apoptosis on the surrounding tissue, fuelling its growth and progression into cancer. The numbers signify sequential order.

1.5.1. Cell Competition as a Tumour Suppressor

CC's function as a tumour suppressor sits on its elimination of aberrant cells, which could be pre-cancer cells, thus preventing oncogenesis[44,105]. *In vitro* conducted studies show that cells with the *RAS-V12* mutation, one of the most prevalent in cancers[106], are extruded from the WT epithelium by CC, preventing them from proliferating, both *in vitro*[75] and *in vivo*[107]. Similar evidence is found in Src activated cells[76] in ERBB2-overexpressing cells[78], in YAP constituently activated cells[69] and also in CDC42 activated cells[77], all well described to be oncoproteins or related to cancer[108-111].

Strikingly, obesity, which is a well described risk factor for several cancers[112-114], was found to hamper this extrusion[115]. Hyperinsulinemia, a hallmark of type 2 diabetes mellitus, which is a disorder that is linked with higher cancer risk[116], was also found to abrogate CC, leading to epithelial tumourogenesis[117]. Additionally, when *TP53*-mutant cells, an common ocurrence in cancer[118], were present on a WT epithelium, they suffered necroptosis via cell competition[74]. These mechanisms constitute what is known as Epithelial Defence Against Cancer (EDAC), where cells that harbour potentially carcinogenic mutations are eliminated from the epithelium[52]. Worthy of note are the limitations of this defence, for cells which are both mutant for *TP53* and encode a *HRAS-V12* mutation are not culled from the WT tissue, remaining in it and proliferating[74]. *In vivo*, flies with *lgl* knockout developed neoplastic masses in their imaginal discs, but when they were intermingled with WT cells, they were eliminated, preventing the tumours from surfacing[119]. *SCRIB* mutants are also eliminated, and if this elimination is blocked, aggressive tumour growth occurs[120]. Identical findings were seen in mice. On a WT thymus, progenitor cells from the bone marrow arrive and replace the resident progenitor cells, which have impaired IL-7 signaling[53]. If this replacement is blocked, however, tumoural transformation occurs, with the development of a disorder resembling human T-ALL[53].

1.5.2. Cell Competition as a Tumour Promoter

On the other side of the coin, lies CC's role in cancer progression. In *Drosophila* mosaics overexpressing *MYC*, one of the most famous protooncogenes[121], the cells trigger apoptosis on the surrounding WT environment, thus behaving as supercompetitors[65,66,122]. Cell mixing of the *MYC* mutant cells with their normal counterparts enhances cell elimination, showing that the rate of apoptosis is dependent on surface area contact between winners and losers[100,123]. Moreover, cells with deficient expression of *APC*, which is a common occurrence in colon cancer[124], also outcompete the surrounding WT epithelium, allowing for tumoural growth[72]. Remarkably, if CC is blocked through apoptosis suppression, tumour growth is impeded, suggesting that CC drives its growth[72]. Cells which co-express *EGFR*, also implicated in cancer[125], and the *miR-8* microRNA also behave as supercompetitors, engulfing nearby cells, a process that is also prevented by apoptosis inhibition[51]. Furthermore, mice knockout for the

abovementioned *Fwe* have resistance against skin papilloma formation, hinting that CC plays a role in cancer initiation[126]. Human cancer cell lines can also engage in CC events, particularly in a MYC-Mediated manner[60,67]. In this scenario, the cells with higher levels of *MYC* expression are the winners, whereas the ones with lower levels, are the losers[60,67]. The CC event was changeable by exogenous *MYC* modulation through the use of shRNAs, inverting which cells were the prospective winners and losers[60], and was blocked by the use of apoptosis inhibitors[60,67]. The apoptosis of the loser cells was also found to induce proliferation of the winner cells[60], and this apoptosis-induced proliferation has also been linked to cancer progression[127]. Also noteworthy is the effect of the extracellular protein SPARC, the previously mentioned protective signal for loser cells[98], in cancer. Higher levels of *SPARC* expression in cancer are commonly linked with a worse prognosis[128], and since this protein is secreted by loser cells on the onset of CC, it could signify that when cancers actively use CC to grow, they do so more aggressively, leading to worse overall survival (OS) and disease-free survival (DFS)[128]. The expression of this gene is even upregulated at the tumour-normal tissue boundary, where theoretically CC is taking place[54].

The multitude of roles of CC in cancer, in particular its context dependency, offer a new paradigm with which to tackle the various stages of cancer development. Further understanding of the underlying molecular pathways and their effects, could therefore potentially constitute an important step to better combat the mortality and morbidity which stem from this disease.

2. Objectives

Although some genes are known to trigger a CC phenomenon, the exact effectors of the pathway in humans are not yet recognized. Furthermore, CC has also been implicated in cancer, with some CC related genes linking to worse survival. However, the way CC leads to this worse prognosis, is also not yet fully understood. To tackle these two issues, a gene expression analysis and several phenotypical characteristics associated with worse prognosis were evaluated.

As such, the first specific aim for this project was to:

- Unveil novel genes implicated in cell competition.
    - Optimization of a co-culture assay;
    - RNA extraction, sequencing and analysis;
    - Gene hits validation;
    - shRNA validation;
    - Co-Cultures with shRNA treated cells.

The second specific aim for this project was to:

- Discover new phenotypic changes that occur during cell competition.
    - Evaluate proliferation of cells engaged in CC;
    - Evaluate migration of cells engaged in CC;
    - Study the influence of chemotherapy in CC.

3. Materials and Methods

3.1. Materials Used

A complete list of all the material and reagents used in this project is provided as Supplementary Data.

3.2. Cell Lines

The following cell lines were used in this study: LoVo (ATCC® CCL-229™), LS 174T (ATCC® CL-188™), HCT 116 (ATCC® CCL-247™) and HEK-293 (ATCC® CRL-1573™). LoVo and LS 174T colon cancer cell lines were gently provided by the Grifoni Lab from the University of l'Aquila. HCT 116 and HEK-293 were kindly ceded by the Rita-Fior Lab from Champalimaud Center for the Unknown. All cell lines were tested and were mycoplasma free.

3.3. Cell Culture

Cell lines were maintained frozen in DMEM with 10% DMSO (Sigma-Aldrich). After thawing, they were grown on either 25cm^2 or 75cm^2 flasks, in accordance with cell quantity need for downstream experiences. Cell lines were cultivated in DMEM high glucose (Biowest), supplemented with 10% FBS (Gibco) and 1% Penicillin/Streptomycin (ABM). Cells were maintained in an incubator at 37°C with 5% CO_2. When confluence was reached, culture medium was removed from the flask, followed by a wash with PBS 1X (Biowest), to remove medium which would otherwise prevent TrypLE (Gibco) action. TrypLE 1X was then added, and cells were incubated for 3 minutes at 37°C to achieve detachment. Culture medium was added to block TrypLE action, and cells were resuspended, followed by centrifugation at 1200 rpm for 4 minutes. The supernatant was removed, cells were resuspended, and 1:10 of them transferred to a new flask, where new complete medium was added.

3.4. Cell Membrane Labelling

Cell lines were grown and, following centrifugation (see above), were resuspended, and counted using an automatic counter (TC20, Bio-Rad). 4x10^6 cells of each cell line were then subjected to cell membrane labelling, either with green fluorescent PKH67 dye (Sigma-Aldrich) or with CellVue Claret Far Red fluorescent dye

(Sigma-Aldrich), in accordance to the manufacturer protocol, scaled down to account for the lower number of cells used.

3.5. Generation of Modified Cell Lines through Lentiviral Transduction

Lentivirus were produced to create cell lines stably expressing GFP or TdTomato, and cell lines knocked-down for several genes of interest.

All lentiviral production was performed using the Lipofectamin 3000 Transfection Reagent (Invitrogen), according to the "Improve lentiviral production using Lipofectamine 3000 reagent" manufacturer protocol. GFP and TdTomato constructs were respectively inserted in the FUGW (Addgene, Plasmid #14883) and FUtdTW (Addgene, Plasmid #22478) plasmids. shRNA constructs were inserted in either a pLKO.1-puro or a TRC2-pLKO-puro plasmid backbone. HEK-293 cell line was used to produce virus particles and was cultivated with Opti-MEM Reduced Serum Medium (Gibco), supplemented with 5% FBS and kept in an incubator at 37°C with 5% CO2. The packaging vectors used for virus production were pLP1, pLP2, VSVG (ViraPower mix).

For transduction, LoVo, LS 174T and HCT 116 cells were cultivated in Opti-MEM Reduced Serum Medium, to which 8 μg/mL of polybrene were added with 2 mL of the lentivirus supernatant from the above protocol. Cells transducted with GFP and TdTomato lentivirus were later subjected to a new transduction, using the lentivirus containing shRNAs. Selection of cells expressing GFP or TdTomato was performed by sorting through FACS (FACSAria Fusion), whereas cells expressing the shRNAs were selected through Puromycin (Sigma-Aldrich), changed every 2 days for 8 days.

3.6. Co-Culture Assay

After either the membrane labelling or lentiviral transduction protocol, cells were recounted and diluted to the concentration of $1x10^6$ cells per mL. Afterwards, cells were cultured, as usual, on monoculture (seeding of $2x10^5$ cells per mL) or in co-culture against other cell lines with a different colour labelling (seeding of $1x10^5$ green cells and $1x10^5$ red cells per mL), in triplicates. Other proportions were also tested in the co-culture experiments: cells were seeded at a 1:4 ratio ($1,5x10^5$ cells of red/green + $5x10^4$ cells of green/red per mL) or at a double initial seeding ($2x10^5$ green cells plus $2x10^5$ red cells per mL).

3.7. Apoptosis Assay

To measure cell apoptosis, the Annexin V (BioLegend) staining protocol was used, based on the manufacturer's protocol. Following 48 or 72 hours after seeding in co-culture, cells were detached from the 24-well plates, resuspended in fresh medium, and transferred to a 1,5 mL microtube. Cells were then counted and disposed of to a maximum quantity of 4 x 10^5 cells per microtube, followed by centrifugation for 4 minutes at 1200 rpm. The supernatant was removed, and the pellet was dislodged in 100 µL of Annexin V Binding Buffer (BioLegend). Afterwards, an optimized quantity 1 µL of Pacific Blue Annexin V (BioLegend) and 1 µL of Propidium Iodide (PI) (1mg/ml) (Sigma-Aldrich) were added to each microtube, which were incubated for 15 minutes in the dark. After incubation time, 250 µL of Annexin V Binding Buffer was added to each microtube, and cells were analysed for apoptotic signals on a flow cytometer. When cells were labelled by lentiviral transduction, only the Pacific Blue Annexin V was used, as the PI signal couldn't be discernible against the TdTomato signal on the flow cytometer. PI staining information is not included on the graphs, to allow comparison between both protocols.

3.8. Cell Proliferation Assay

To measure cell proliferation, 48 hours after the co-culture assay, cells were incubated for 1 hour with 10 µM EdU and then detached, pelleted, and subjected to Click-iT EdU proliferation assay Alexa Fluor 555 Azide (Thermo-Fisher), with some minor alterations to the manufacturer's protocol to scale down the assay for the 24-well plates which were used in this study. Posterior to the kit protocol, cells were analysed on a flow cytometer for fluorescent signal.

3.9. Flow Cytometry

Apoptosis and proliferation assay experiments were tested for their fluorescence through flow cytometry. All flow cytometry data was analysed through FlowJo version 10.7.1, where cells were subjected to several gates, sequentially:

For the apoptosis assay cells were first subjected to the forward versus side scatter gating, allowing for the removal of cell debris, while maintaining alive and

apoptotic cells. Afterwards, forward scatter height versus forward scatter area was used, to exclude doublets. Cells were then stratified in accordance to their labelling, using different flow cytometry filters. Those two resulting groups were then evaluated for their Annexin V and PI fluorescence, resulting in 4 final populations. Owing to flow cytometry limitations, when the lentiviral protocol is used, only Annexin V signal is evaluated, as TdTomato and PI are detected using the same flow cytometry filter (Phycoerythrin (PE)).

For the proliferation assay cells were first subjected to the forward versus side scatter gating, allowing for the removal of cell debris and apoptotic cells, while maintaining alive cells. Afterwards, forward scatter height versus forward scatter area was used, to exclude doublets. Cells were then stratified in accordance to their labelling, using different flow cytometry filters. Those two resulting groups were then evaluated for their EdU flourescence.

3.10. Emricasam, Oxaliplatin and Puromycin Concentration Optimization

For caspase inhibition optimization, a total of 2 x 10^5 LoVo and LS 174T cells per mL were co-cultured at a 1:1 proportion and incubated for 48 hours with different emricasam (Selleckchem) concentrations. Cells were then measured for apoptosis using the aforementioned protocol, and the lowest concentration of emricasam which inhibited LoVo cell elimination was chosen. In the following experiments, the optimized concentration of 10 µg/mL emricasam was added to the cell culture medium, followed by normal apoptosis protocol.

For cytotoxicity experiments, a previously optimized[37] concentration of 100 µM oxaliplatin (Sigma-Aldrich) was added to the cell culture medium 24 hours after seeding, followed by normal apoptosis protocol after another 24 hours.

Considering the shRNA plasmid contains a puromycin resistance gene, selection of shRNA expressing cells was made using the antibiotic (Sigma-Aldrich). The optimal concentration of puromycin for each cell line selection was tested by plating 1 x 10^5 cells per mL in 24 well-plates, and culturing for 7 days in the presence of different concentrations of puromycin (Sigma-Aldrich), ranging from 1ug/mL to 10ug/mL. The optimal concentration was chosen as the lowest which was sufficient to kill all cells in

the well. For LS 174T cells, the concentration of puromycin chosen was 3 µg/mL, for LoVo cells it was 5 µg/mL, and 2 µg/mL for HCT 116 cells.

3.11. Scratch-wound Assay

48 hours after co-culture seeding, cell culture medium was removed, and a 200 µL pipet tip was used to make a vertical scratch on the well surface area. The well was then washed with PBS to remove excessive resulting cell debris, followed by adding of new cell culture medium. Using a microscope (Leica DFC3000 G) coupled to a LED (CoolLed), cells were stimulated with both Green or Blue laser, and a photo was taken immediately after scratch, and from 2 to 2 hours, until a total of 8 hours. Since the plate could not be fixed in spot, photo location might have changed slightly with each subsequent picture. To dilute this issue, on each photo, 5 different points were chosen, and the "wound" width was measured using FiJi software version 1.53c. Rate of cell migration was defined as

$$\frac{Wound\ Width_{t=0h} - Wound\ Width_{t=8h}}{Hours\ Elapsed}$$

3.12. Fluorescence Activated Cell Sorting (FACS)

Following 48 or 72 hours after seeding in co-culture, cells were detached from the 24-well plates and pelleted by centrifugation at 1200 rpm for 4 minutes. Cells were resuspended in FACS Buffer and filtered through a 70 µm cell strainer (Fisher Scientific) into a flow cytometry tube. All cells were subjected to FACS, including monocultures, to assure proper controls. The separated cells were then either cultured or subjected immediately to RNA extraction.

3.13. RNA Extraction

Cells were pelleted and the RNA was extracted using RNeasy Mini Kit (Qiagen), in accordance to the manufacturer protocol. RNA was eluted in 20 µL of Nuclease Free Water (Invitrogen) and then evaluated quantitatively and qualitatively with a spectrophotometer (Nanodrop 2000). RNA was then stored at -80°C.

3.14. RNA-Sequencing

LoVo and LS 174T cells were seeded in monoculture and co-cultured with each other. After 48 hours, all cultures were sorted through FACS, monocultures included, and cells were subjected immediately to the RNA extraction protocol. Following extraction, resulting RNA was then evaluated quantitatively and qualitatively with a spectrophotometer. RNA was then sent for sequencing at CeGaT, Germany. Resulting reads were assessed for quality and then aligned to the human genome assembly of Dec. 2013 (hg38) using the R package STAR (Bioconductor). Afterwards, a count of the number of reads was assigned to each gene, and differentially expressed genes were identified using the R DESeq2 package (Bioconductor). Further filters were then applied to narrow down the number the hits, in order to facilitate subsequent analysis.

3.15. Primer Design

Primer sequences used in this study were either previously reported or newly designed. When previously reported the intended target was confirmed with the UCSC genome browser[129] and their melting temperature and non-intended targets were evaluated with NCBI Primer BLAST[130].

Alternatively, sequence of the gene of interest was procured from UCSC genome browser[129], and NCBI Primer Blast[130] was used to design the primers. Primer pairs with exon-exon junction were chosen or, if not available, those with closer GC content, melting temperature and with fewer off targets were chosen. For these purposes, the human assembly of Dec. 2013 was used (hg38). All primers were ordered from Eurofins Scientific. A complete list of the primer sequences used in this study is presented as Supplementary Data (Table S1).

3.16. RT-qPCR

The previously extracted RNA was converted to cDNA QuantiTect Reverse Transcription Kit (Qiagen) in accordance to the manufacturer's protocol. A quantity of 200 ng was used for all reverse transcription protocols. All reverse transcription protocols were made on a thermal cycler (Bio-Rad). Afterwards, qPCR was performed

with the cDNA using PowerUp SYBR Green Master Mix (Applied Biosystems), in accordance to the manufacturer's protocol. All experiments were performed on a StepOnePlus (Applied Biosystems), on 96-well reaction plates with a final reaction quantity of 20 µL per well. The RT-qPCR data was analysed using StepOne Software version 2.3. All relative fold change was acquired through the $2^{-\Delta\Delta Ct}$ Method.

3.17. Plasmid DNA extraction

shRNAs ordered against genes of interest were all MISSION shRNA Bacterial Glycerol Stock (Sigma-Aldrich). While near an open flame, a 20 µL pipet tip was used to scratch the bacterial glycerol stock and the tip was added to 100 mL of LB broth in an Erlenmeyer. The bacteria were left to grow overnight at 37°C at 200 rpm. A Midiprep was then carried out (Zymo Research) following the manufacturer's alternative centrifugation protocol. Resulting Plasmid DNA was then evaluated quantitatively and qualitatively with a spectrophotometer. The Plasmid DNA was then stored at -20°C. A complete list of the shRNAs used in this project is presented as Supplementary Data (Table S2).

3.18. Bioinformatics Analysis

GDC TCGA datasets were accessed, analysed, and expression levels of the gene hits were compared between normal and tumour tissues. The values were in FPKM-UQ to facilitate cross-sample comparison and differential expression analysis[131]. The datasets used for this analysis were BRCA (Breast invasive carcinoma), COAD (Colon adenocarcinoma), KIRC (Kidney renal clear cell carcinoma), LIHC (Liver hepatocellular carcinoma), LUAD (Lung Adenocarcinoma), PAAD (Pancreatic Adenocarcinoma) and PRAD (Prostate Adenocarcinoma). Pathway analysis was performed using Reactome[132]. Survival analysis was performed using GePia[133], using median as a group cut-off and a 95% Confidence Interval.

3.19. Data Analysis

All RT-qPCR results were analysed using the ΔΔCt Method. For the co-culture experiments, all data was considered as having a normal distribution, and the statistical significance was given by independent two-sample t-test. All cells in co-culture were

compared to their monoculture. The null hypothesis was presented as H_0 = There is no difference between the apoptosis/migration/proliferation of the two groups. For the GDC TCGA expression analysis, all data was considered to have a normal distribution and statistical significance was given by independent two-sample t-test. The tumour expression values were compared to the normal tissue expression values. The null hypothesis was presented as H_0 = There is no difference between the expression of the gene between the two groups. All values with "0" were removed as they were considered technical mistakes and diluted the expression values (previous analysis demonstrated that their removal did not alter the statistical significance). "Whiskers" on the box-plots are given by 1,5 x IQR. The data analysis software used for statistical analysis was RStudio, R version 3.6.2.

3.20. Image Design

All graphs on this study were made using RStudio, R version 3.6.2, with the ggplot2 package. All other images were designed using the free version of Biorender.

4. Results

4.1. Optimization of a Co-Culture Assay

To analyse the behaviour of cells in and out of CC, a co-culture assay was established between two cancer cell lines. LoVo and LS 174T were chosen, as those colon cancer cell lines were previously describe to engage in CC with one another[60]. As such, cells were differentially colour-labelled, either through lentiviral transduction with GFP and TdTomato, or through use of either red or green fluorescent dyes that incorporate into the cell membrane, which allowed to separately analyse the contrasting coloured cell lines through flow cytometry (Figure 8 and Figure 9).

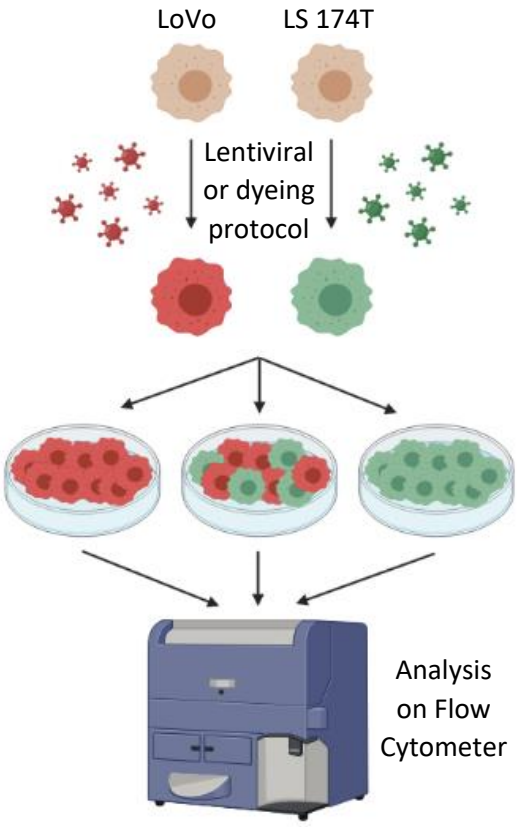


Figure 8. Co-culture assay workflow. Cell lines are subjected to differential labelling through lentiviral transduction or lyophilic dyeing protocol and then cultured in monoculture and in co-culture with each other. After 48 or 72 hours, cells are detached from the plates and analysed on the flow cytometer for apoptosis levels.

Afterwards, cells were co-cultured and mixed together on a 1:1 proportion. 48 hours later, cells were gauged for apoptosis by using Annexin V, simultaneously with propidium iodide (PI). Annexin V is commonly used as an apoptotic marker as it binds to

specific membrane lipids that are only externalized during apoptosis[134]. PI is an intercalating agent, which binds to DNA when the cell is ruptured[134]. As both Pacific Blue Annexin V and PI emit fluorescence on different wavelengths upon binding, this permits discrimination between intact cells (Annexin V-|PI-) (Figure 9D, Q4), early apoptotic cells (Annexin V+|PI-) (Figure 9D, Q3) and late apoptotic cells (Annexin V+|PI+) (Figure 9D, Q2)[134]. Due to limitations of the sensitivity of the flow cytometer, when TdTomato was used, the apoptosis analysis was limited to Annexin V signal (Figure 9E). To allow comparison between results obtained from the two protocols, henceforth, cell apoptosis is always given as total cell positive for Annexin V.

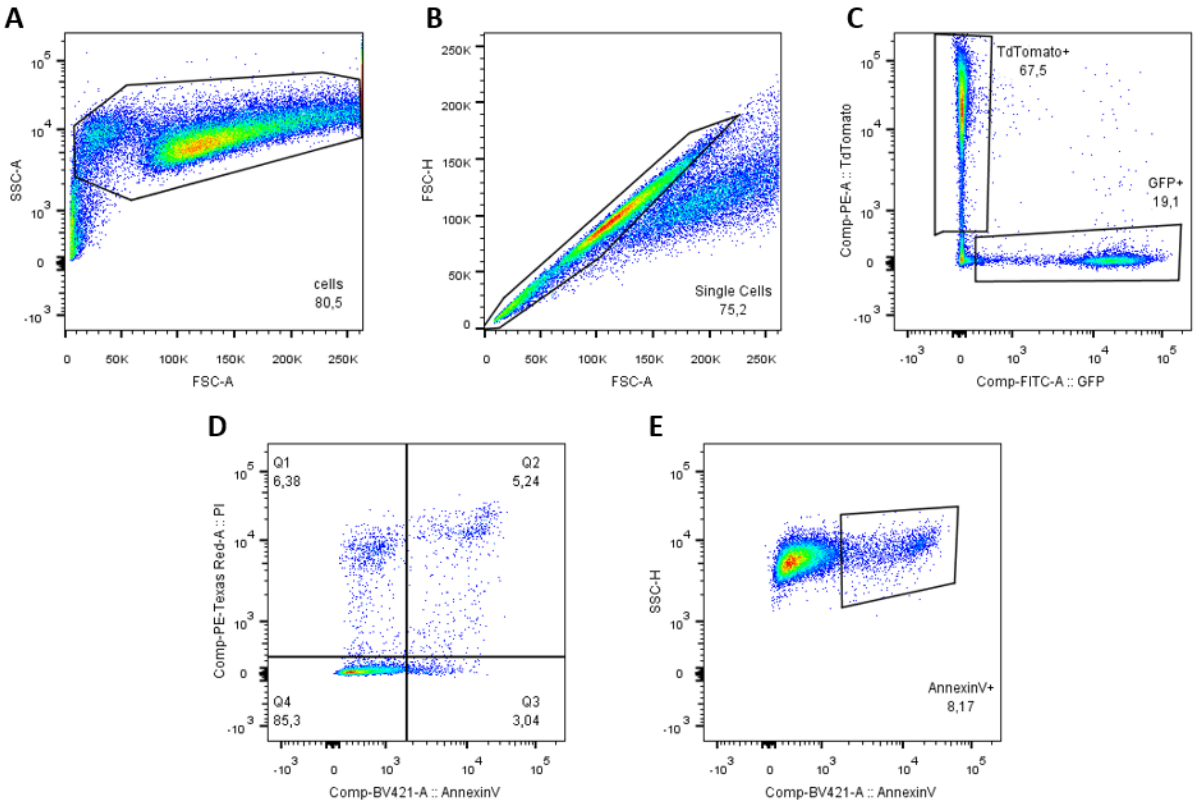


Figure 9. Flow cytometry apoptosis analysis. (A) Example of the first gating of the analysis. The area selected corresponds to the alive cells and apoptotic cells, allowing the removal of cell debris. (B) Second gating of the analysis. The area selected corresponds to the single cells, with the remainder being doublets. (C) Selection of the two differentially labelled cell lines. This gating allows to evaluate each cell line individually. Data from this chart is from the lentiviral protocol of co-culture. (D) After selecting either differentially labelled population, the cells are then gated in accordance to their Annexin V+ and PI+ signal. The Q4 corresponds to intact cells (Annexin -| PI-), Q3 to the early apoptotic cells (Annexin + | PI -) and Q2 to the late apoptotic

cells (Annexin + | PI+). In this specific chart, the total Annexin + signal would be given by adding the Q2 and Q3 quadrants, which would add up to 8,28 % of apoptotic cells in the selected cell line. Data from this chart is from the dyeing protocol. (E) Owing to flow cytometry limitations, when the lentiviral protocol is used, only Annexin V+ signal is evaluated. On this specific chart, 8,17 % of the cells of the selected population are considered to be apoptotic. Data from this chart is from the lentiviral transduction protocol.

4.1.1. LoVo and LS 174T Colon Cancer Cell Lines Engage in CC

To confirm that CC takes place when LoVo and LS 174T are co-cultured together, a co-culture assay was carried out, with both cell lines growing alone, or together with each other at a 1:1 proportion. When in co-culture, LoVo cells suffer more apoptosis than their monoculture counterparts, using both the lentivirus and the dye protocols (Figure 10A). Conversely, the opposite is observed with the LS 174T cells, which have lower levels of apoptosis in co-culture than in monoculture. Data from PI staining reveals that most cells are in the late apoptotic stage. Following the traditional CC definition, LoVo cells therefore behave as losers, since they suffer more apoptosis in co-culture, and LS 174T behave as winners, as they present an improved survival in co-culture.

Cells were also seeded at a higher number per well, to determine the influence of total cell number on CC. When cells are seeded at this higher confluency, the same tendency is observed, with LoVo cells dying more and LS 174T cells dying less, although the phenomenon seems to be attenuated (Figure 10B), suggesting that high confluency inhibits the phenomenon of CC.

To check if cell mixing had an effect in this elimination, different proportions of prospective losers and winners were tested (Figure 10C). When the LoVo cells were seeded at a proportion of 1:4, they suffered an even more drastic elimination, having a percentage of apoptotic cells three times as higher as their basal levels in co-culture (Figure 10C). In this same proportion, LS 174T cells, experience no difference in apoptotic rate when compared to their monoculture and on the inverse seeding of 4:1, LS 174T cells suffer lower apoptosis (Figure 10C). Even when LoVo cells are in the majority, on the 4:1 seeding, they are still subjected to more elimination than their monoculture analogues (Figure 10C). The same conditions were repeated and evaluated at 72 hours instead of the normal 48 hours, and for the LoVo cells, similar results were observed, with the exception that the seeding at 4:1 proportion no longer has more apoptosis (Figure 10D). On the LS 174T cells, all effects appear to be exacerbated at the

72 hours' time point, with all conditions featuring lower apoptosis (Figure 10D). The increase of loser cell elimination on the 1:4 proportion suggests that cell mixing does influence CC, as in that condition there is more contact between loser and winner cells.

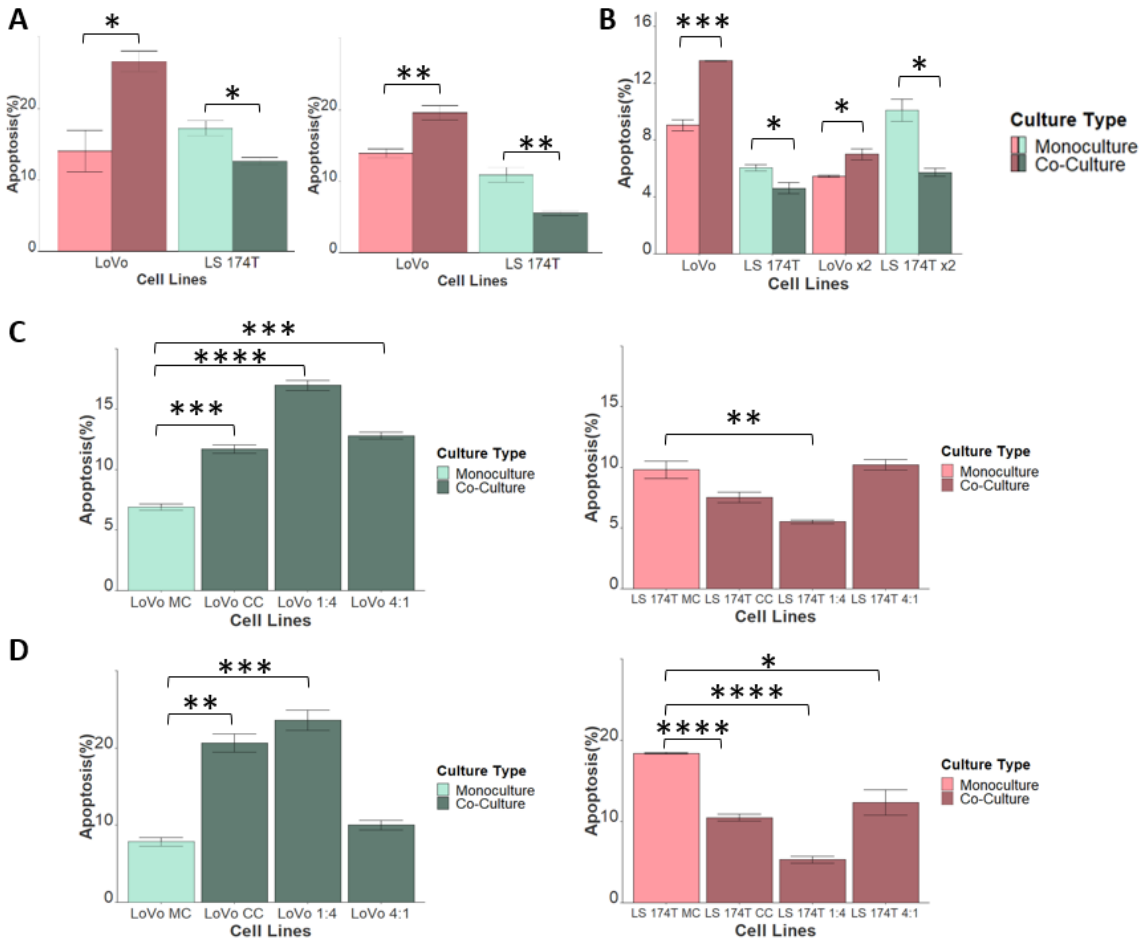


Figure 10. LoVo and LS 174T cells engage in cell competition. (A) LoVo and LS 174T cells in co-culture at normal seeding confluency after 48 hours. Left - Co-culture assay with the lentiviral protocol. Right - Co-culture assay with the dye protocol. All left graph bars have n=3, and all right graph bars have n=4. (B) LoVo and LS 174T cells in co-culture at normal and double seeding confluency after 48 hours. The 4 rightmost bars represent seeding at double confluency. Data from lentiviral protocol. (C) LoVo and LS 174T cells in co-culture at different seeding proportions after 48 hours. Left - LoVo cells apoptosis levels at the different seeding proportions. Right – LS 174T cells apoptosis levels at the different seeding proportions. (D) LoVo and LS 174T cells in co-culture at different seeding proportions after 72 hours. Left - LoVo cells apoptosis levels at the different seeding proportions. Right – LS 174T cells apoptosis levels at the different seeding proportions. Data from lentiviral protocol. All bars have n=3. Error bars correspond to ± SEM (Standard Error of the Means). Non-significance is not represented on the graphs. *P<0,05, **P<0,01, ***P<0,001, ****P<0,0001 when compared to corresponding cells in monoculture.

4.1.2. HCT 116 Cells Engage in CC with LS 174T Cells but Not With LoVo Cells

To check if this apoptosis was conserved between other cell lines, the co-culture assay was repeated using a third cell line, HCT 116, which was co-cultured with LoVo and LS 174T (Figure 11A). When cultured with LoVo cells, HCT 116 cells had no difference in apoptosis, while the LoVo themselves survived better than in their monoculture (Figure 11A). Since none of the cell lines has higher levels of apoptosis in co-culture, this implies that LoVo and HCT 116 cells do not engage in CC with each other. By contrast, when cultured with LS 174T, HCT 116 cells had higher elimination, while the LS 174T cells showed no difference (Figure 11A). Since the HCT 116 cells had more apoptosis, it could be that a CC phenomenon is occurring, albeit less noticeable than what happens with LoVo and LS174T cells cultured together.

To refine these results and confirm that cell mixing also has an impact in the CC between these two cell lines, the experiment with the different proportions was repeated with this co-culture (Figure 11B). In regards to the HCT 116 cells, the results were akin to the ones obtained in the LoVo vs LS 174T experiment, with all seeding proportions having a higher percentage of apoptosis when in co-culture with LS 174T cells, particularly when the HCT 116 are at a minority at seeding time (Figure 11B). The LS 174T cells, however, have an equal apoptosis percentage, and when they are in the minority, they even have a higher percentage of apoptosis (Figure 11B), contrasting to the LoVo vs LS 174T experiment, where they always have an equal or lower percentage. The results suggest that HCT 116 and LS 174T engage in a phenomenon of CC, with HCT 116 behaving as losers and LS 174T behaving as winners. As per the results obtained, this CC is influenced by cell mixing, and it’s not as distinct as the one between LoVo and LS 174T, as the apoptosis of the loser cells is lower, and there are no lower levels of apoptosis on the winners.

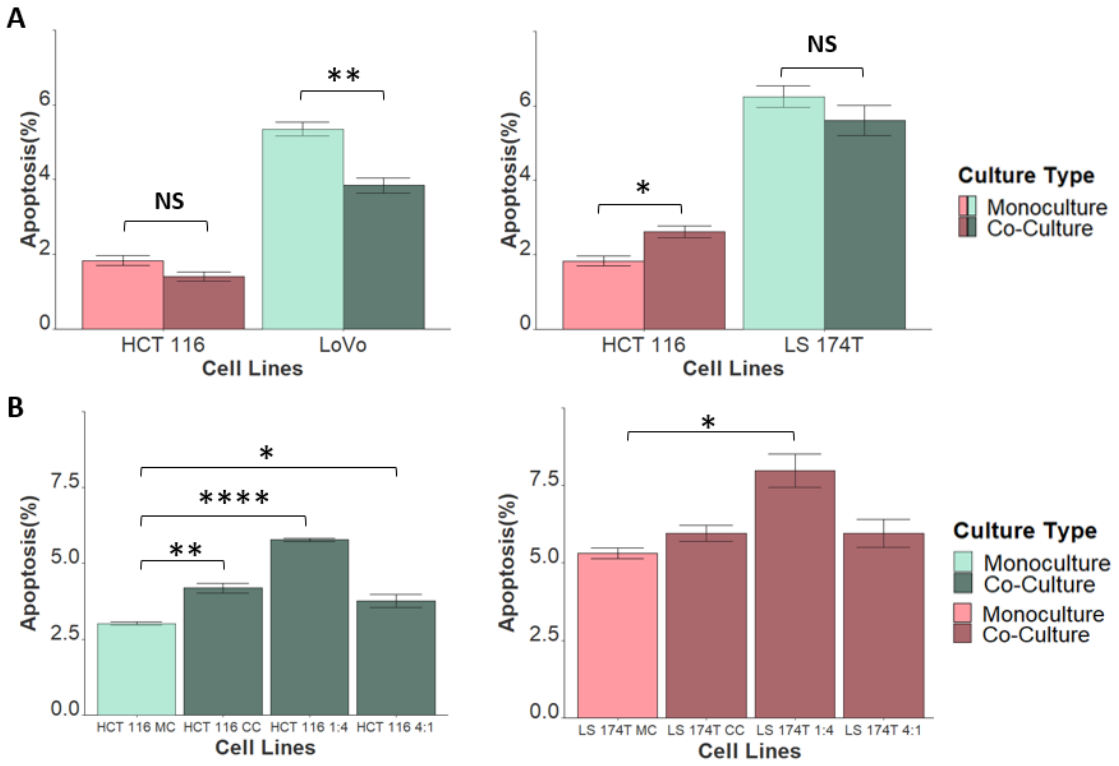


Figure 11. HCT 116 and LS 174T cells engage in cell competition. (A) Left - HCT 116 and LoVo cells in co-culture at normal seeding confluency after 48 hours. Right - HCT 116 and LS 174T cells in co-culture at normal seeding confluency after 48 hours. The same HCT 116 monoculture is used as control on both graphs. (B) HCT 116 and LS 174T cells in co-culture at different seeding proportions after 48 hours. Left - HCT 116 cells apoptosis at the different seeding proportions. Right - LS 174T cells apoptosis at the different seeding proportions. All data are from lentiviral protocol. All bars have n=3. Error bars correspond to ± SEM (Standard Error of the Means). Non-significance is not represented on the "B" labelled graphs. NS Non-significant, *P<0,05, **P< 0,01, ****P<0,0001 when compared to corresponding cells in monoculture.

4.1.3. Apoptosis of Loser Cells in Co-culture is Caspase-dependent

Previous results had shown that caspase inhibition was sufficient to block loser cell apoptosis[96,123]. To determine if the elimination of the loser cells in these scenarios is caspase-dependent, this competition scenario was assessed for its dependence on these enzymes. With this aim, cells were seeded with Emricasam, a powerful pan-caspase inhibitor (Figure 12). When caspases are inhibited, the apoptotic behaviour of the loser cells is rescued, with LoVo cells surviving better in co-culture than in their monoculture (Figure 12). LS 174T maintain the same phenotype of having lower apoptosis levels in co-culture (Figure 12). These observations point towards the fact that LoVo cells' apoptosis in co-culture is caspase-dependent, and that the protective effect observed on the LS 174T is not loser apoptosis dependent.

In addition, the incubation with emricasam was performed with the LS 174T vs HCT 116 co-culture as well, but it had no effect on the levels of apoptosis (Figure 12), indicating that in this case, CC is not caspase-dependent.

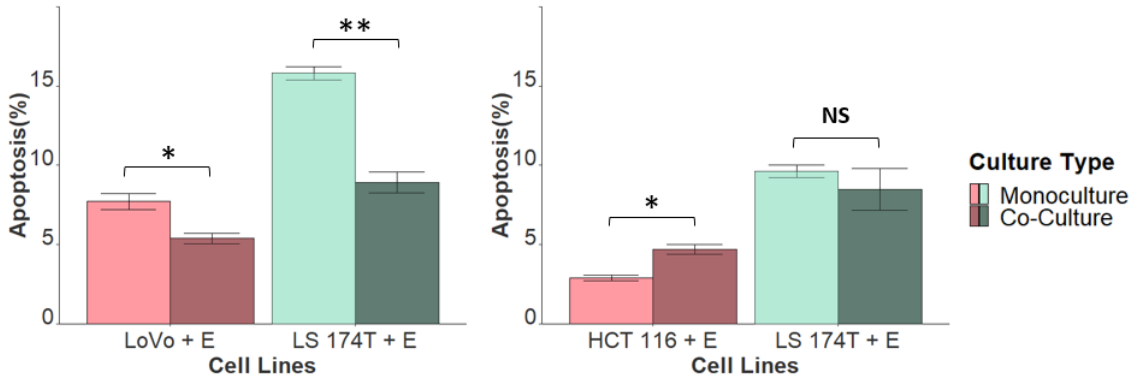


Figure 12. Emricasam inhibits LoVo vs LS 174T CC but not HCT 116 vs LS 174T CC. Different co-culture assays with Emricasam. Left – LoVo and LS 174T co-culture in the presence of pan-caspase inhibitor Emricasam. Right – HCT 116 and LS 174T co-culture in the presence of pan-caspase inhibitor Emricasam. All data are from lentiviral protocol. All bars have n=3. Error bars correspond to ± SEM (Standard Error of the Means). NS Non-significant, *P<0,05, **P< 0,01 when compared to corresponding cells in monoculture.

4.2. Winner Cells Have Higher Proliferation during CC

Although LS 174T cells were found to survive better in co-culture with LoVo, it was not clear if both cells had any other alteration linked to worst cancer prognosis, like higher proliferation. With this goal, co-cultures were incubated with EdU for 2 hours (Figure 13). This thymidine analogue is incorporated into novel DNA during its synthesis, and using flow cytometry methods, allows for determination of the percentage of cells which are in the S-phase, which are cells that are replicating their DNA and are preparing for mitosis. When in co-culture after 48 hours, LS 174T cells have a higher percentage of S-phase cells, while LoVo cells have a lower percentage, when comparing to their respective monocultures (Figure 13). Since the higher apoptosis effect on the LoVo cells was less noticeable after 72 hours, the EdU incubation was repeated at that time point. Whilst LS 174T cells maintain a much higher proliferation on the co-culture, LoVo cells bear no difference between mono and co-culture, with the decrease seen at 48 hours no longer detected (Figure 13).

A well described event is apoptosis-induced compensatory proliferation, where the death of some cells lead to the proliferation of the surviving cells[135,136], which has

also been described to happen in CC[60]. To test if this was the case, cells were incubated with the pan-caspase inhibitor, to evaluate whether the elimination of LoVo cells had an impact on LS 174T cells proliferation. While LoVo cells proliferation decrease was hindered, LS 174T increased proliferation was maintained when apoptosis was blocked (Figure 13). The results indicate that CC has a direct effect on the winner cells, as they proliferate more, and on the loser cells, as they proliferate less when in co-culture. This higher proliferation on the winner cells does not seem to be apoptosis-induced compensatory proliferation, since the blockade of the pan-caspase inhibitor did not rescue the effect.

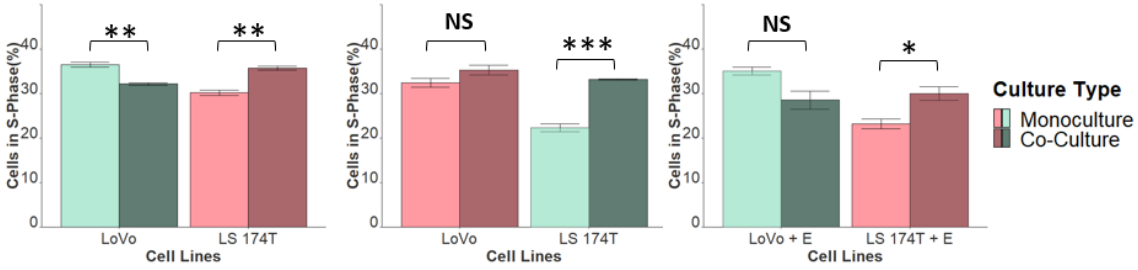


Figure 13. CC increases winner cell proliferation. Right - After 48 hours, LoVo cells have a lower percentage of S-phase cells and LS 174T have a higher percentage. Middle - After 72 hours, LS 174T have a much higher proportion of S-phase cells while LoVo cells exhibit no difference. Right - After 48 hours, in the presence of the pan-caspase inhibitor Emricasam, LoVo show no difference, and LS 174T cells have a higher percentage of S-phase cells. All data are from lentiviral protocol. All bars have n=3. Error bars correspond to ± SEM (Standard Error of the Means). NS Non-significant, *P<0,05, **P< 0,01, ***P<0,001 when compared to corresponding cells in monoculture.

4.3. Winner Cells and Loser Cells Have Different Migration Rates in CC

Considering that CC has been linked to metastasis, it was postulated that CC could also have an impact on cell migration. To test this, cells were co-cultured and a scratch-wound assay was done on the seeding surface. A photo was taken immediately after scratching, followed by subsequent photos every two hours until a total of 8 hours was reached. Cells were stimulated with coloured LEDs before photos to allow for selective visualization of only GFP+ or TdTomato+ cells (Figure 14A). LoVo cells had a lower migration rate when in co-culture, and LS 174T cells had a higher rate (Figure 14B). The experiment was repeated using the reverse colours and both cell lines exhibited the previously mentioned differential migration (Figure 14C). This allowed to infer that CC

has a direct effect on both loser and winner cells, transforming their invasive capabilities.

To investigate if this was due to social migration[137], where cells anchor to each other and migrate at similar rates, scratch-wound assay was conducted using two cell lines which do not engage in CC, LoVo and HCT 116 (Figure 15A). With this co-culture, HCT 116 presents with less migration on co-culture and LoVo presents no difference (Figure 15B). The fact that only HCT 116 cells have their migration changed indicates that this effect is different from the co-cultures where CC is in play, because on those experiments, both cells exhibited alterations in their migration.

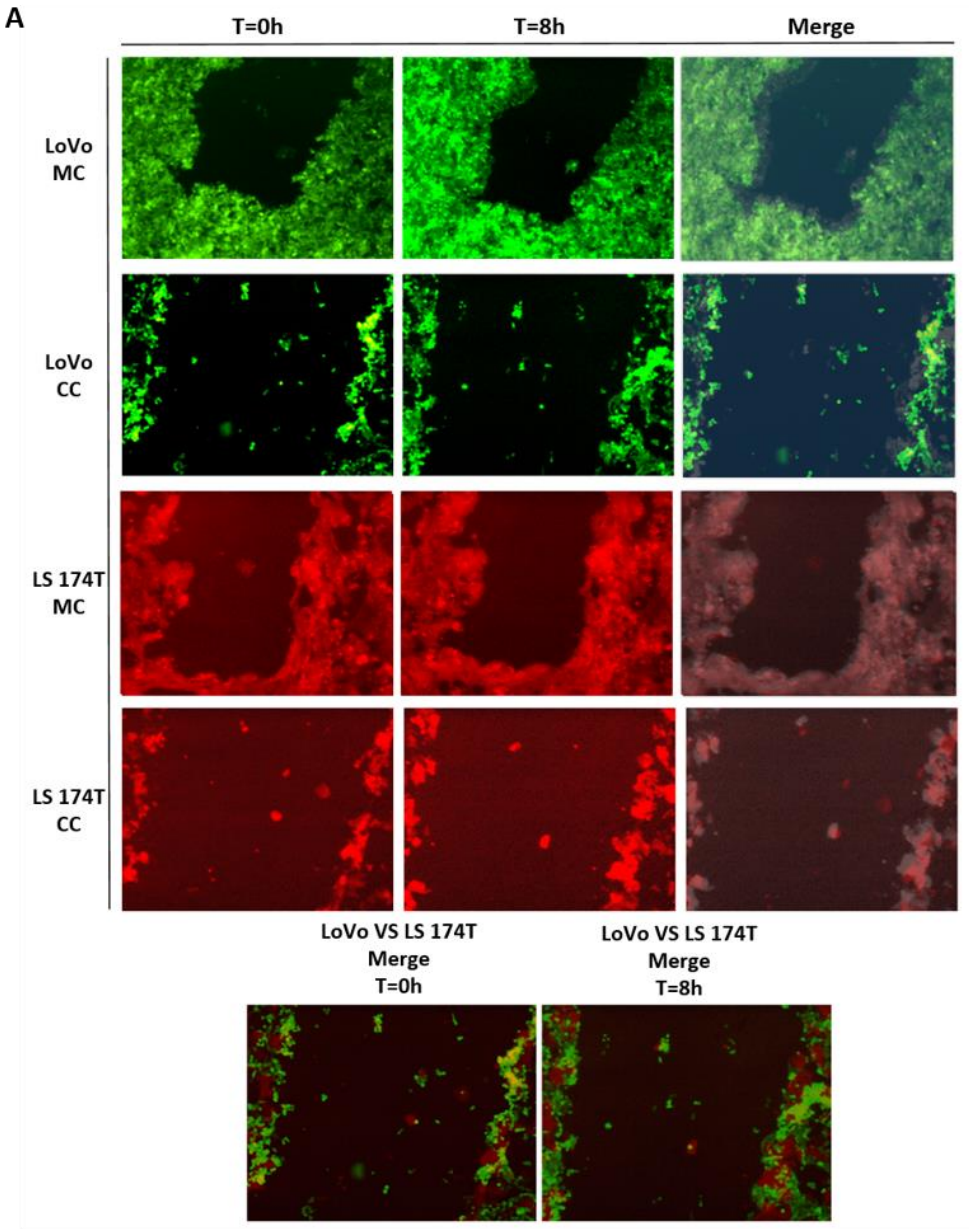

A
T=0h
T=8h
Merge
LoVo MC
LoVo CC
LS 174T MC
LS 174T CC
LoVo VS LS 174T Merge T=0h
LoVo VS LS 174T Merge T=8h
B
Wound Width(px)
Timepoint
t=0h
t=2h
t=4h
t=6h
t=8h
LoVo MC
LoVo CC
LS 174T MC
LS 174T CC
Cell Lines
Migration Rate(px/h)
Culture Type
Monoculture
Co-Culture
LoVo
LS 174T
C
Wound Width(px)
Migration Rate(px/h)

Figure 14. LoVo and LS 174T have different migration capabilities during CC. (A) Representative images of the scratch-wound assay of LoVo and LS 174T cells engaging in CC and their respective monoculture. Images shown include merges of the 0 hours' time point with the 8 hours' time point, and a merge of both cell lines, representing the co-culture. An amplification of 40X is used for the images. (B) Left -Wound width at the several time points of the different conditions of the assay. Wound width is normalized to value at t=0h time point. Right – Summary of the scratch-wound assay. (C) Scratch-wound assay using cells with inverted colours. Left - Wound width at the several time points of the different conditions of the assay. Wound width is normalized to value at t=0h time point. Right - Summary of the scratch-wound assay. All data are from lentiviral protocol. All bars have n=3. Statistical significance is not shown on the graphs on the left. Error bars correspond to ± SEM (Standard Error of the Means). *P<0,05, **P< 0,01 when compared to corresponding cells in monoculture. Image treatment was performed on ImageJ.

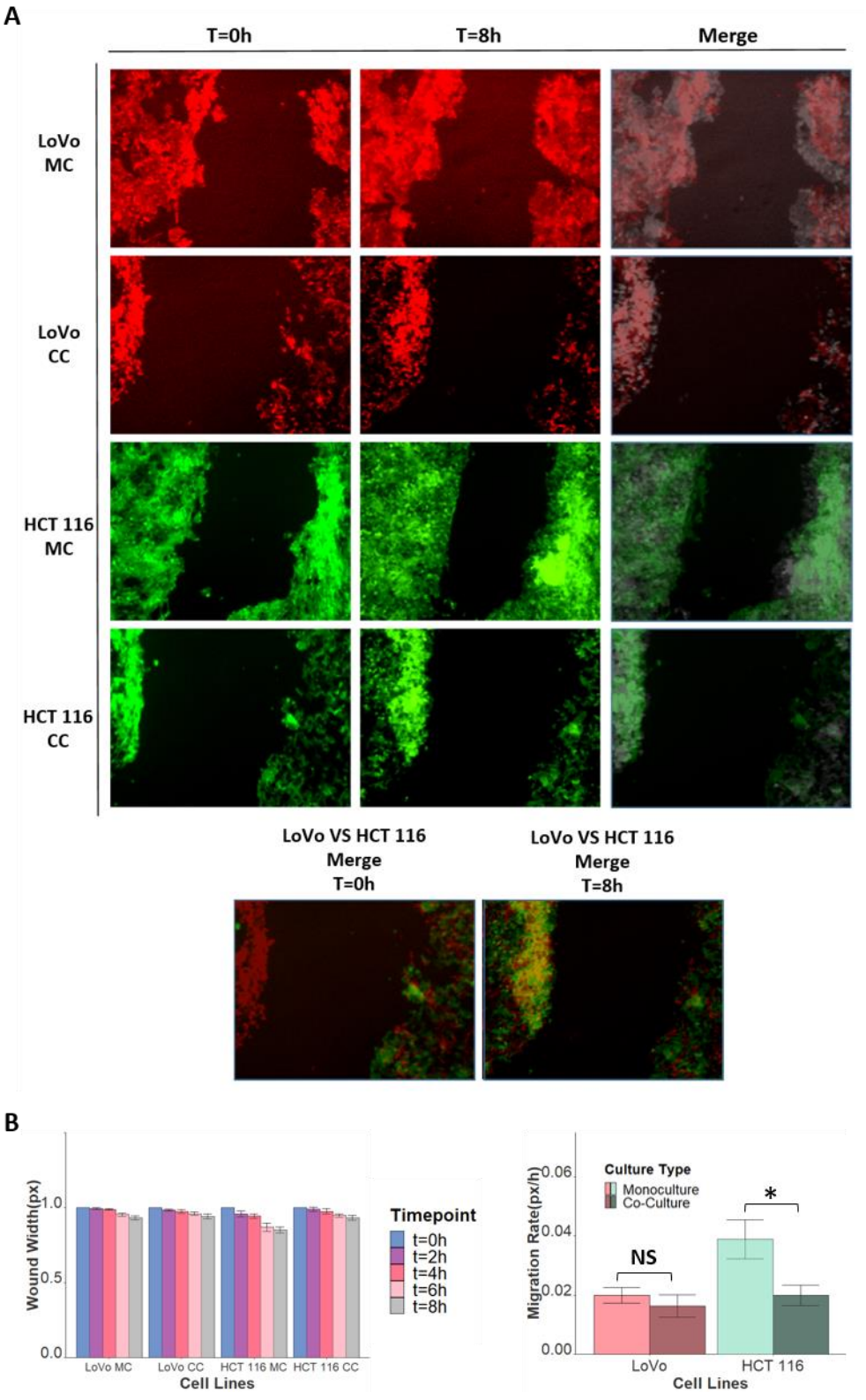


Figure 15. LoVo and HCT 116 migration profile in co-culture. (A) Representative images of the scratch-wound assay of LoVo and HCT 116 cells in co-culture and their respective monoculture. Images shown include merges of the 0 hours' time point with the 8 hours' time point, and a merge of both cell lines, representing the co-culture. An amplification of 40X is used for the images. (B) Left - Wound width at the several time points of the different conditions of the assay.

Wound width is normalized to value at t=0h time point. Right – Summary of the scratch-wound assay. All data are from lentiviral protocol. All bars have n=3. Statistical significance is not shown on the graphs on the left. Error bars correspond to ± SEM (Standard Error of the Means). NS Non-Significant, *P<0,05, when compared to corresponding cells in monoculture. Image treatment was performed on ImageJ.

4.4. Chemotherapy Affects Cell Competition Between LoVo and LS 174T

A commonly used treatment in cancer patients is chemotherapy, which is highly cytotoxic and affects cells on several levels[138]. Aiming to study the effects of chemotherapy on CC to better replicate *in vivo* conditions, LoVo and LS 174T cells were co-cultured in the presence of Oxaliplatin, used to treat colon cancer. Afterwards, cells were analysed following the apoptosis assay (Figure 16). Somehow, the experimental control contrasted with all the previous results, with both LoVo and LS 174T cells survived better in co-culture, although the reasons for such are unclear. However, the results on the seeding with oxaliplatin were striking, with an apparent "inversion" of the CC phenomenon. LoVo, the previous loser cell, now survives better in co-culture, with LS 174T, the former winners, having higher levels of apoptosis in co-culture. Assuming the control had an unexpected result owing to some mistake, and assuming that same mistake happened on their chemotherapy counterparts, a clear effect is still observed. This is therefore the first time CC has been directly implicated in chemoresistance, greatly increasing the survival of LoVo cells in detriment of the LS 174T cells.

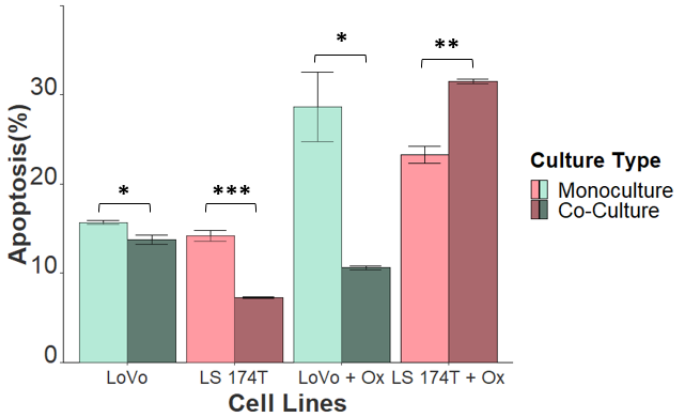


Figure 16. LoVo and LS 174T co-culture in the presence of Oxaliplatin. The graph represents LoVo and LS 174T levels of apoptosis in the presence of the cytotoxic drug, oxaliplatin. All data are from lentiviral protocol. All bars have n=3. Error bars correspond to ± SEM (Standard Error of the Means). NS Non-Significant, *P<0,05, **P<0,01, ***P<0,001 when compared to corresponding cells in monoculture.

4.5. RNA-Sequencing

To unravel the genes responsible for the transduction of the cell competition event on the loser and winner cells, LS 174T and LoVo cells were differentially membrane dyed and monocultured and co-cultured with each other. After 48 hours, cells were detached, and monocultures and co-cultures were subjected to FACS and segregated in accordance to their membrane labelling. RNA was then extracted from the sorted cells and sent for sequencing. Resulting transcript reads were then compared between co-cultured and monocultured cells, and hits were considered as genes which had at least a 150% increase/decrease on their number of reads, with an adjusted *P*-value of <0,1 being further subjected to other filters to narrow down results (Table 2).

Table 2. List of gene hits from the RNA-Seq analysis. Loser cells refers to LoVo cells in co-culture and winner cells refers to LS 174T cells in co-culture.

Alteration	*Gene IDs*
Downregulated in loser cells	*COL6A3, CYP1A1, DCN, LAMC2, MGP, MSMO1, MUC4, SLC34A2, SULF1, THBS2*
Upregulated in winner cells	*AQP3, GPX2, MYT1, NRIP1, PDK4, SKIL, SLC6A14, SPINK4, SYNPR, TRPM8, UGT1A1, UGT1A10*

In order to validate the gene hits obtained, the previously extracted RNA was used to perform a RT-qPCR. Moreover, to show that the method used to distinguish the cells did not alter gene expression, RT-qPCR validations were also conducted on cells sorted after the lentiviral transduction protocol. For these analyses, the housekeeping gene *GAPDH* was used as a reference gene. For a relative quantification the comparative ΔΔCT method was used. All melting curves were checked for and presented a single peak. A single peak on the melting curve means that no unspecific products were amplified, and that no primer-dimers occurred (Figure 17).

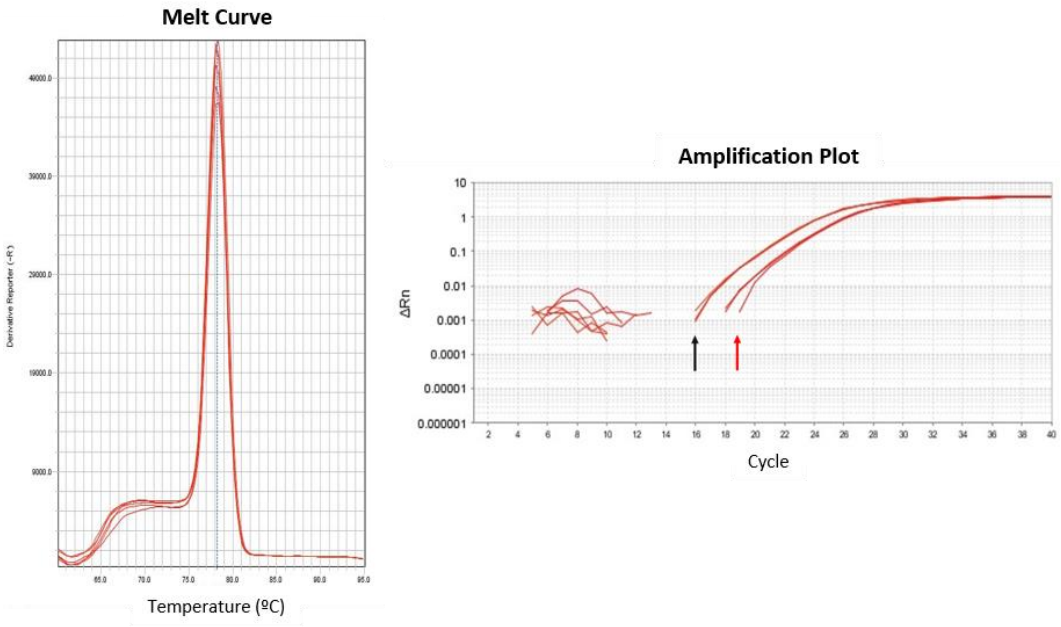


Figure 17. Representative melt curve and amplification plot of RT-qPCR experiments. Left - On the melt curve, only one peak is present, which shows that there are no unintended targets, or primer-dimer formation. Right - The gene amplification shown on the images is *UGT1A1* on LS 174T cells. The red arrow is the amplification of the transcript in monocultured cells and the black arrow is the amplification of the transcript in co-cultured cells. In this example, since the co-culture replicates appear first on the plot (left to right), the gene is therefore overexpressed in co-culture.

The results of the RT-qPCR were similar in the dye and lentiviral protocol, ensuring that both methods used to distinguish the cells are comparable with each other (Figure 18A and 18B). This ensured that only replicable genes were maintained between the two experiments, with *CYP1A1, LAMC2, MSMO1, AQP3, GPX2, MYT1, NRIP1, PDK4, SKIL, UGT1A1* and *UGT1A10* being used on all following experiments. Herein, inconclusive genes were then discarded from subsequent analysis (*SLC6A14* and *DCN* for not correlating with RNA-Seq results; *TRPM8* for having a lot of variation between the two experiments; *COL6A3, MGP, MUC4, SLC34A2, SULF1* and *THBS2* for not having amplification in any condition).

To verify if the genes could be responsible for the behaviour of the cells during CC, another RT-qPCR was run using samples from the 72 hours' co-culture (Figure 18C). Since CC is still taking place at this timepoint, it is expected for hit gene expression to still be verifiable. Unsurprisingly, most winner genes had a similar increase as the 48 hours' timepoint, with the notable exception that *AQP3* expression is much higher at

this timepoint, when comparing to the respective monoculture. The loser genes were not as evident as the winner genes, with *CYP1A1* not being as downregulated, and *LAMC2* suffering an increase in co-culture. The fact that most of the genes are identical between the two timepoints indicates that these genes are consistently altered during CC.

Moreover, to test if these same genes were involved on the CC detected between the HCT 116 vs LS 174T co-culture, a RT-qPCR was also ran for the gene hits on this experimental condition (Figure 18D). Contrasting with LoVo in co-culture, HCT 116 cells have a faint downregulation on the *MSMO1*, with *CYP1A1* and *LAMC2* being upregulated in co-culture. Referring to the winner genes, only *AQP3, PDK4, UGT1A1* and *UGT1A10* are upregulated in co-culture, and that upregulation is much less pronounced than in the LoVo vs LS 174T co-culture. The results on the gene hit validation, indicate that the higher apoptosis of HCT 116 cells in co-culture, does not appear to be regulated by the same genes as the other co-culture.

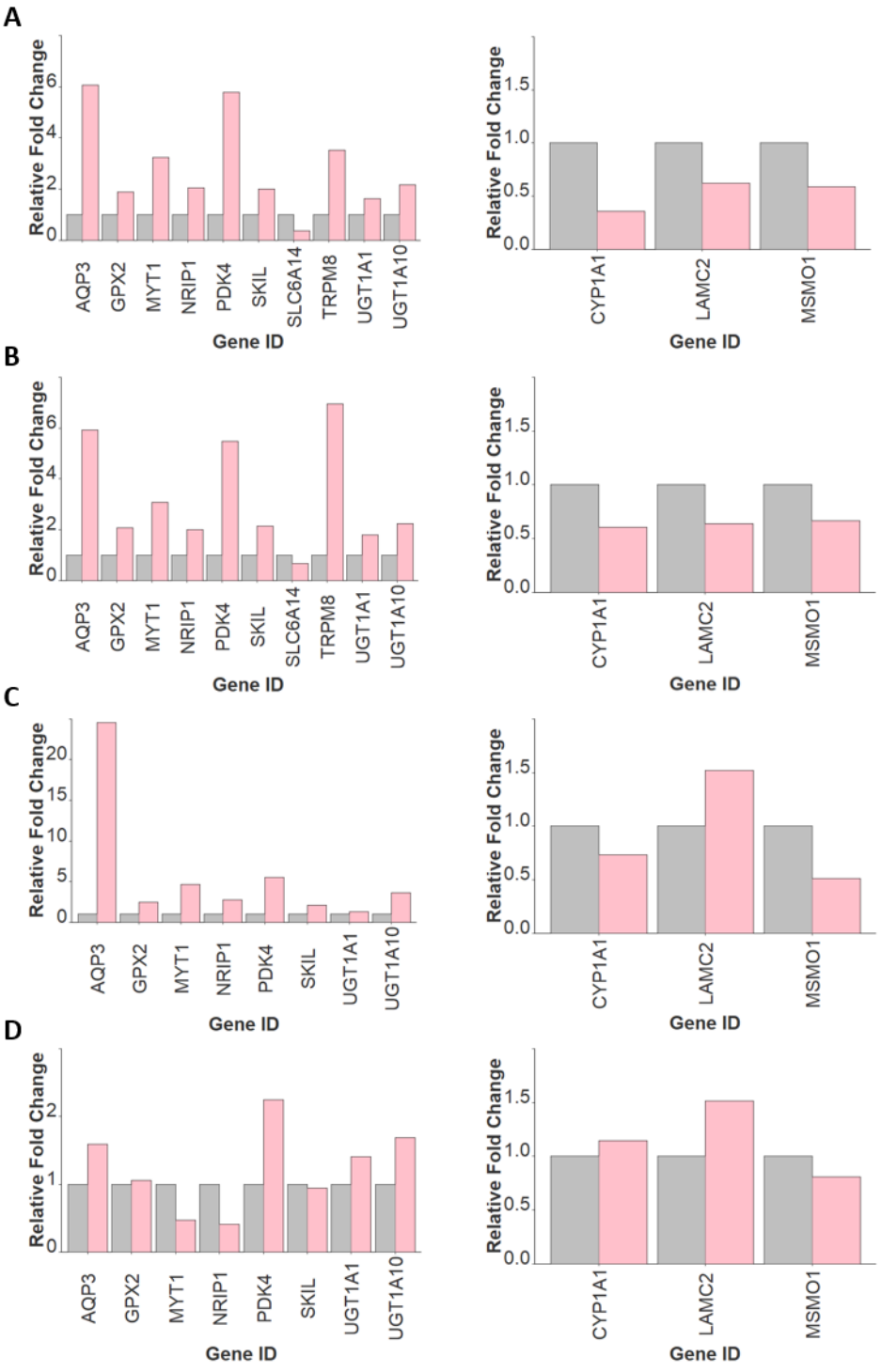


Figure 18. Graphical representation of the relative fold change of gene hits. (A) Relative Fold Change of gene hits using the dyeing protocol. Left – Gene Amplification related to winner cell genes. Right – Gene Amplification related to Loser Cell Genes. (B) Relative Fold Change of gene hits using the lentiviral transduction protocol. Left – Gene Amplification related to winner cell genes. Right – Gene Amplification related to Loser Cell Genes. (C) Relative Fold Change of gene hits using the lentiviral transduction protocol after 72 hours. Left – Gene Amplification related to winner cell genes. Right – Gene Amplification related to Loser Cell Genes. (D) Relative Fold Change of gene hits using the lentiviral transduction protocol in an HCT 116 with LS 174T co-

culture. Left – Gene Amplification related to winner cell genes. Right – Gene Amplification related to Loser Cell Genes. Grey bars represent fold change 1 (no alteration). All values were obtained through the $2^{-\Delta\Delta Ct}$ method.

4.6. Evaluation of Additional Genes

Seeing that these gene hits are likely involved in other pathways, extra RT-qPCR were run against genes of the EGFR signalling pathway on loser and winner cells. This pathway was chosen on account of encompassing other important signalling cascades, like MTOR, JAK/STAT or MAPK[125]. Moreover, it has been previously implicated in CC[51]. Adding to genes of this pathway, two common Epithelial-Mesenchymal Transition markers, CDH1 (E-cadherin) and VIM were also tested[139] (Table 3 and Figure 19).

Table 3. List of other pathway genes which were altered in the RT-qPCR. Loser cells refers to LoVo cells in co-culture and winner cells refers to LS 174T cells in co-culture. Only genes with a 25% change in either direction were considered to be differentially expressed.

Alterations	*Gene IDs*
Downregulated in loser cells	*MTOR*, *STAT3*, *EGFR*, *PTGS2*, *CDH1*, *VIM*, *MAPK1*
Downregulated in winner cells	*CDH1*, *MAPK1*
Upregulated in winner cells	*ERBB2*, *STAT3*, *MTOR*, *PTGS2*, *VIM*, *AKT*

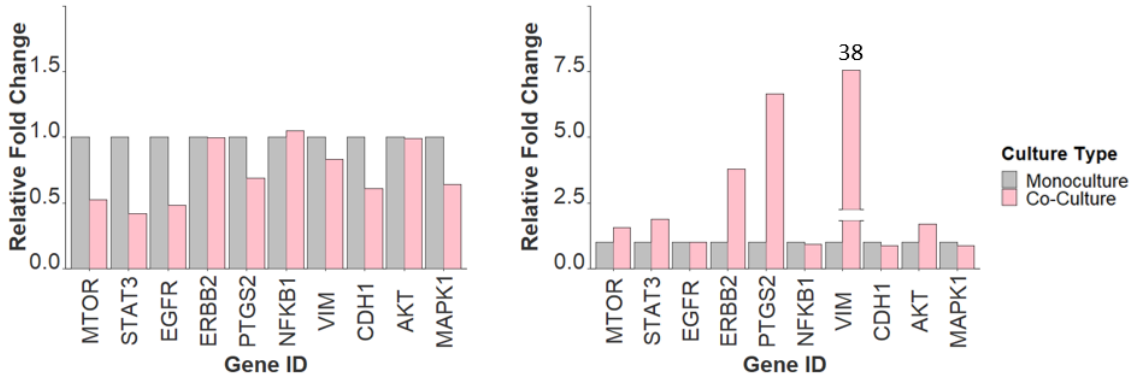


Figure 19. Graphical representation of the relative fold change of other pathway genes. Left – Gene Amplification related to loser cell. Right – Gene Amplification related to winner cells. *VIM* bar is truncated to allow representation on the graph, as it has a 38-fold increase in relation to LS 174T in monoculture. Grey bars represent fold change 1 (no alteration). All values were obtained through the $2^{-\Delta\Delta Ct}$ method.

Regarding the loser cells in co-culture, results indicate that the EGFR pathway, which is mainly linked to proliferation[125], is inhibited (Figure 19). In fact, most downstream signalling cascades have their expression diminished. *MAPK1* (Erk), mainly linked to proliferation[140], *MTOR*, mainly related to survival and proliferation[141], and *STAT3*, mainly related to proliferation and survival[142], are all diminished suggesting pathway inhibition, which can explain the higher apoptosis and lower proliferation observed in the loser cell engaging in CC (Figure 19). Additionally, *VIM*, linked to migrating cells[139], is downregulated, which also supports the lower migration seen on these cells during CC (Figure 19). Nevertheless, *CDH1*, which is commonly downregulated in migrating cells[139] is also diminished, which contraries the phenotype (Figure 19).

Concerning the winner cells in co-culture, the RT-qPCR reveal that the EGFR pathway is activated, as there is an increase in *ERBB2*, a proto-oncogene linked to a great increase in EGFR signalling[143] (Figure 19). The downstream *MTOR* and *STAT3* are also upregulated although *MAPK1* is diminished, indicating the signal is transduced mainly through MTOR and STAT cascades (Figure 19). The activation of those two cascades correlates with the lower apoptosis and higher proliferation characteristic of the LS 174T cells in co-culture. Common markers of EMT are also present, with a striking upregulation of *VIM* and downregulation of *CDH1*[144] (Figure 19). These two alterations support the hypothesis that the higher migration seen in winner cells is a direct consequence of the CC phenomenon and not a social migration phenomenon where the LS 174T cells are being dragged by the LoVo. Furthermore, *PTGS2* (COX-2), which is known to supress loser cell delimination[145], is upregulated in CC, which could signify that it is contributing to the lower winner cell apoptosis (Figure 19).

Summarising, pro-proliferating and survival genes like *EGFR*[125], *MTOR*[146] and *STAT3*[142] are active on the winner cells and inactivated on the loser cells, which is consistent with the previous results presented in this study. EMT markers[144] are also present in winner cells, which could explain the higher migration seen in these cells when in CC, whereas *VIM* downregulation in loser cells could explain the lower migration rate.

4.7. *AQP3*, *MYT1* and *NRIP1* Upregulation is Necessary for Loser Cell Elimination

To validate whether the gene hits played a functional role in CC, shRNAs were used to inhibit their expression on the winner cells. As such, LS 174T cells were lentiviral-transducted with GFP and with a shRNA against a gene hit. LoVo cells were transducted with only TdTomato. LS 174T were then selected for puromycin resistance, and afterwards, both cells lines underwent FACS. Resulting cells were then cultured for several days to increase their numbers, after which RNA was extracted to verify hit gene downregulation (Figure 20). Genes with at least a 25% decrease on their expression were considered to be downregulated. *SKIL* and *UGT1A10* showed no meaningful repression and *UGT1A1* was even increased. Since any results using these shRNAs would not be conclusive owing to their inconclusive downregulation, they were discarded from the following experiments.

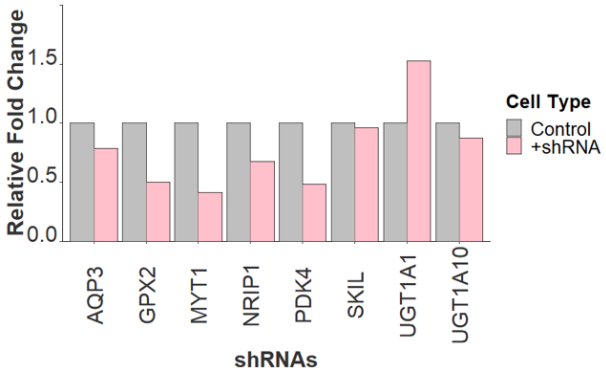


Figure 20. Graphical representation of the relative fold change of the gene hits in the presence of a targeted shRNA. Grey bars represent fold change 1 (no alteration). All values were obtained through the $2^{-\Delta\Delta Ct}$method.

Accordingly, several co-culture assays were designed to test if the genes were necessary for cell competition. LoVo cells were thus co-cultured with LS 174T control, or with LS 174T treated with shRNAs against one gene hit. Both cells' apoptosis was then measured (Figure 21). When LoVo cells were co-cultured with LS 174T treated with *AQP3* shRNA, the CC seen on the control (Figure 21A) was abrogated, with instead both cells surviving better, which is akin to the co-culture assay where loser's apoptosis is blocked with Emricasam (Figure 21B). Likewise, this seemingly blockade of CC also occurs on the co-cultures with LS 174T cells treated with shRNA against *MYT1* and with LS 174T cells treated with shRNA against *NRIP1* (Figure 21D/E). The *PDK4* shRNA had no effect on the cells as CC is apparently still in play (Figure 21F). The *GPX2* shRNA wasn't scientifically

significant (p>0,05), but the tendency seen suggests that CC is still in play, although attenuated (Figure 21C). Additionally, a direct comparison between the LoVo cells in the different co-cultures was performed (Figure 21G), with only the co-cultures with LS 174T cells treated with shRNAs against *AQP3*, *MYT1* or *NRIP1* exhibiting any difference.

Thus, this project managed to discover novel genes implicated in cell competition, with their upregulation being necessary for this process.

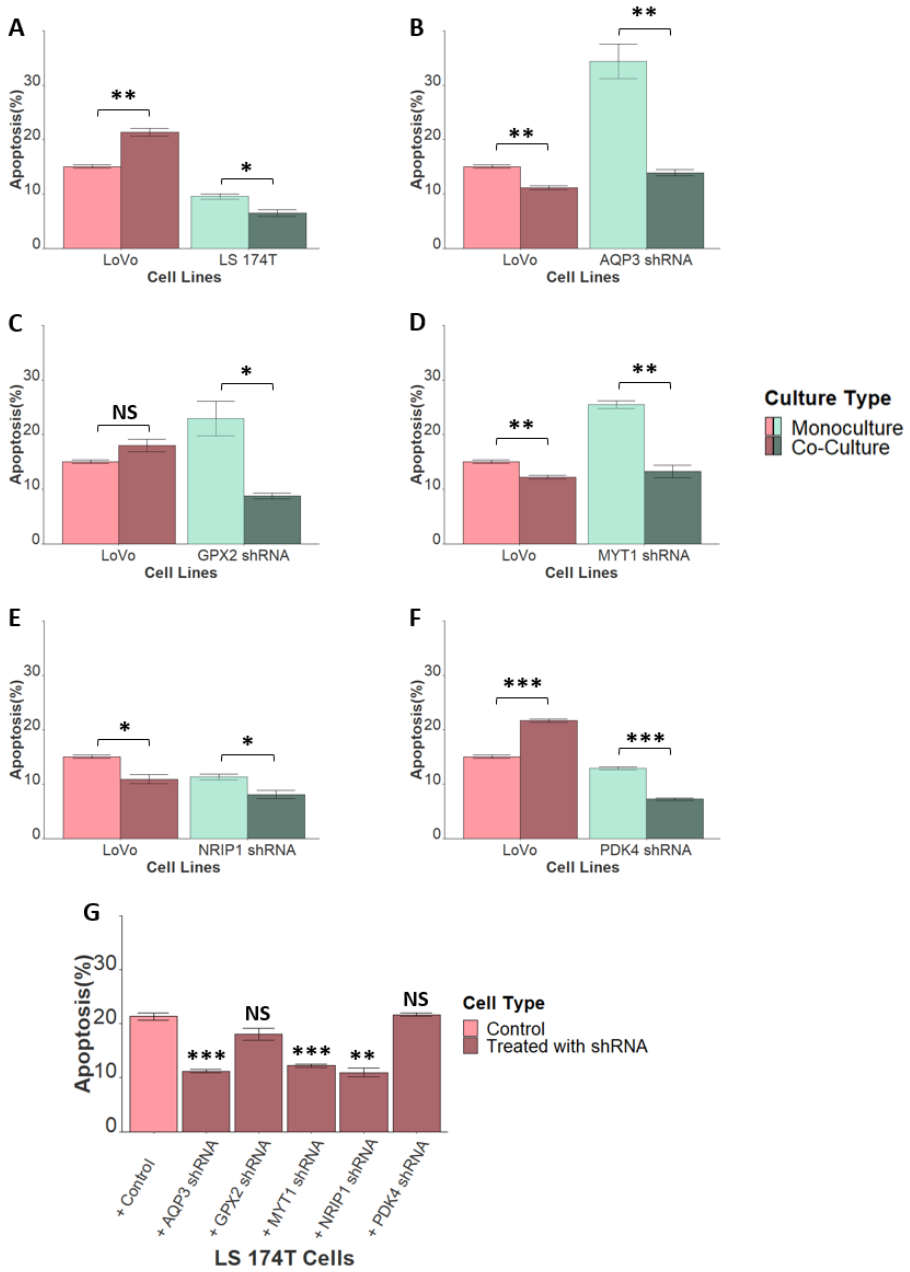


Figure 21. LoVo co-cultures with shRNA treated LS 174T cells. (A) LoVo co-culture with control LS 174T cells. (B) LoVo co-culture with LS 174T cells treated with shRNA against *AQP3*. (C) LoVo co-culture with LS 174T treated with shRNA against *GPX2*. (D) LoVo co-

culture with LS 174T cells treated with shRNA against *MYT1*. (E) LoVo co-culture with LS 174T cells treated with shRNA against *NRIP1*. (F) LoVo co-culture with LS 174T cells treated with shRNA against *PDK4*. (G) Apoptotic percentages of LoVo cells in co-culture against LS 174T cells treated with shRNA. The X axis label indicates the LS 174T cell that were cultured against the LoVo cells. All data are from lentiviral protocol. All bars have n=3. Error bars correspond to ± SEM (Standard Error of the Means). NS Non-Significant, *P<0,05, **P<0,01, ***P<0,001 when compared to corresponding cells in monoculture.

4.8. Bioinformatics Analysis

To better understand the novel CC implicated genes, a brief summary of their function and role in cancer was researched using scientific search engines and Reactome[132]. A survival analysis was also conducted using GEPIA, which compares patients with high vs low expression using TCGA data. The TCGA datasets themselves were analysed and expression levels of the gene hits were compared between normal and tumour tissue (a bioinformatics analysis on the other genes is presented as supplementary information).

4.8.1. *APQ3*

Aquaporin 3 is a membrane water channel protein that also facilitates the transport of glycerol and hydrogen peroxide[147]. In cancer, it seems to be a tumour promoter as it is overexpressed in a plethora of cancers, including lung, colon, hepatocellular carcinoma and pancreatic ductal adenocarcinoma[147]. Its overexpression has also been linked to metastasis as it upregulates matrix metalloproteases, which are necessary to break down extracellular matrix (ECM) proteins and facilitate invasion of cancer cells[148,149]. Curiously, disruption of the *AQP3* gene was found to prevent skin tumorigenesis in a murine study[150]. An important pathway in which it has been implicated is the PI3K/AKT pathway[151]. Despite these published results, expression analysis reveals that *AQP3* has lower expression in kidney clear cell carcinoma, hepatocellular carcinoma and prostate adenocarcinoma, which could be owed to different patient samples being used (GDC TCGA data – Figure 22). Although being described as tumour promoter, in pancreatic ductal adenocarcinoma, overexpression is linked with higher DFS (Gepia analysis – Figure 23).

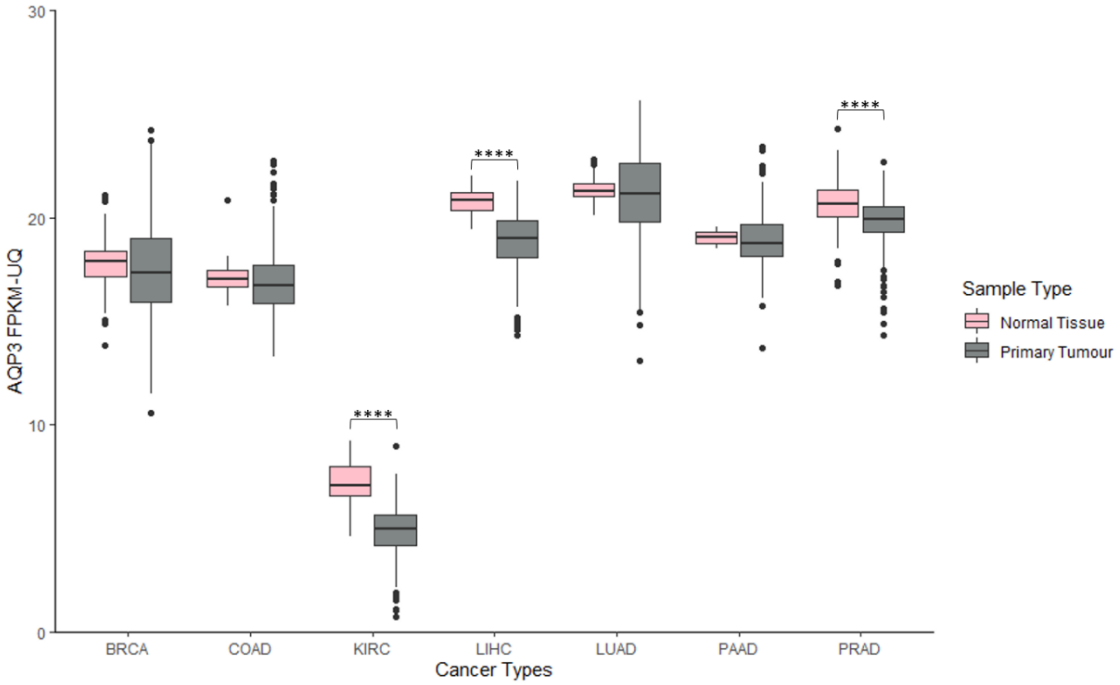


Figure 22. Expression analysis of *AQP3* on several GDC TCGA datasets. *AQP3* expression is lower in kidney clear cell carcinoma, hepatocellular carcinoma and prostate adenocarcinoma. Dots represent outliers. Non-significance not shown. ****P<0,0001 when compared to normal tissue expression.

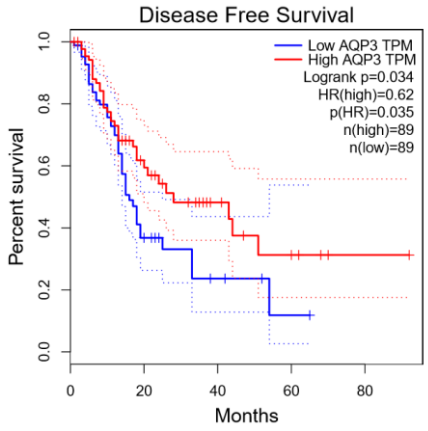


Figure 23. Disease free survival analysis of pancreatic adenocarcinoma patients stratified by *AQP3* expression levels. *AQP3* expression is correlated with a better DFS in these patients.

4.8.2. *MYT1*

Myelin transcription factor 1 is a cell cycle-regulated kinase that inhibits Cdc2, thus stopping cell cycle progression[152-154]. The knockdown of this gene has been found to lead to apoptosis[155]. It has been found to downregulate YAP1, improving survival in glioblastoma[156]. Expression analysis reveals that *MYT1* has higher expression in breast invasive carcinoma and hepatocellular carcinoma. (GDC TCGA data – Figure 24). Lower

expression has been associated with worse OS in pancreatic adenocarcinoma (Gepia Analysis – Figure 25).

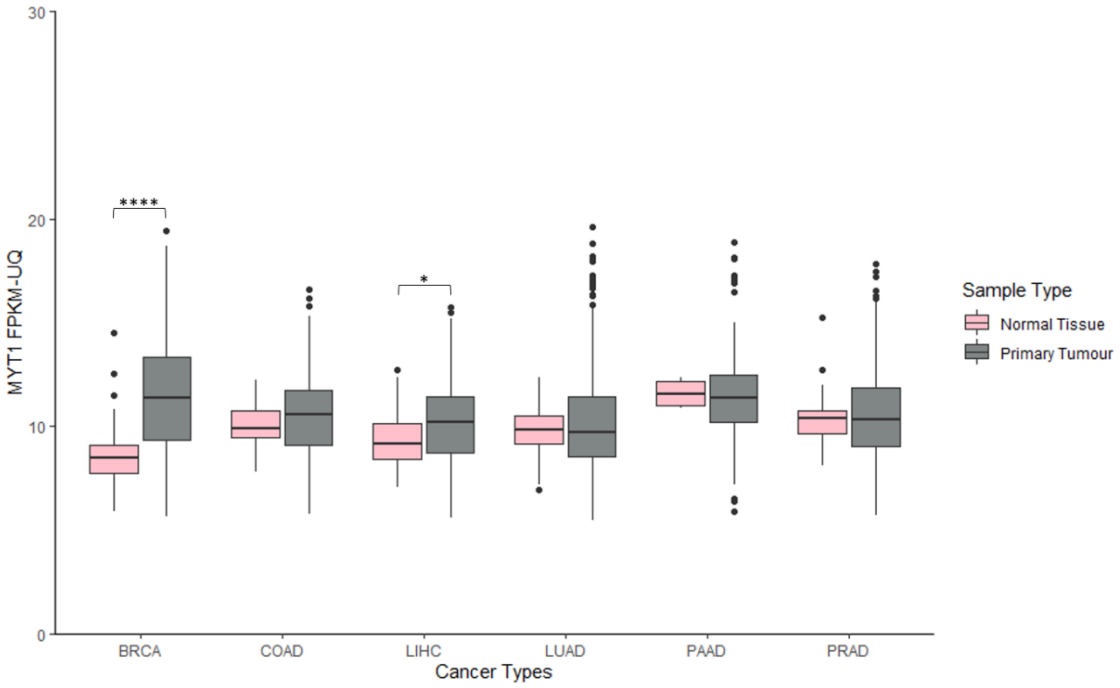


Figure 24. Expression analysis of *MYT1* on several GDC TCGA datasets. *MYT1* expression is higher in breast invasive carcinoma and hepatocellular carcinoma. KIRC database was not used because *MYT1* expression levels were very low in both tumour and normal tissue. Dots represent outliers. Non-significance not shown. *P<0,05, ****P<0,0001 when compared to normal tissue expression.

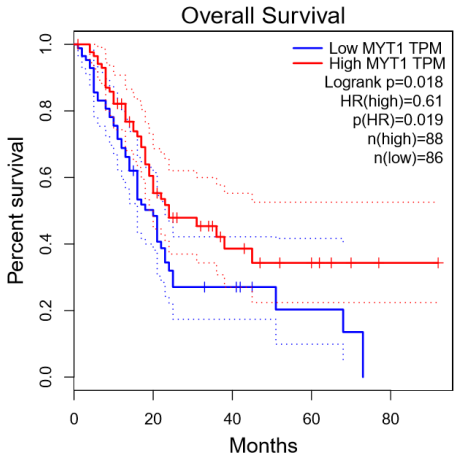


Figure 25. Overall survival analysis of pancreatic adenocarcinoma patients stratified by *MYT1* expression levels. *MYT1* overexpression is linked to better OS in these patients.

4.8.3. *NRIP1*

Nuclear receptor interacting protein 1 plays a role in regulating both lipid and glucose metabolism[157]. Suppressing of its activity inhibits breast cancer cell growth[158], but promotes growth and migration of hepatocellular carcinoma cells[159]. It has been found to inhibit WNT signalling through up-regulation of APC, thus stopping colon tumourogenesis[160]. A notable pathway where it is a key regulator is estrogen receptor signal transduction[161]. Expression analysis reveals that *NRIP1* has lower expression in colon adenocarcinoma, kidney clear cell carcinoma and hepatocellular carcinoma (GDC TCGA data – Figure 26). Its overexpression is linked to better OS and DFS in kidney renal clear cell carcinoma and to worse DFS in pancreatic adenocarcinoma (Gepia Analysis – Figure 27).

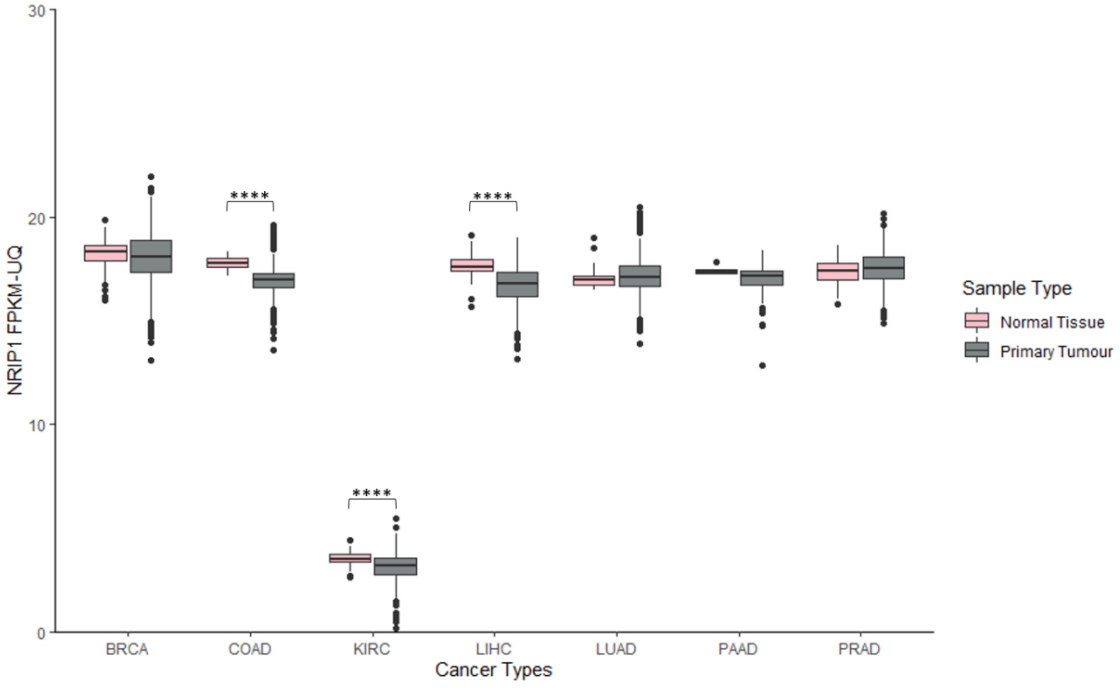


Figure 26. Expression analysis of *NRIP1* on several GDC TCGA datasets. *NRIP1* expression is lower in colon adenocarcinoma, kidney clear cell carcinoma and hepatocellular carcinoma. Dots represent outliers. Non-significance not shown. ****P<0,0001 when compared to normal tissue expression.

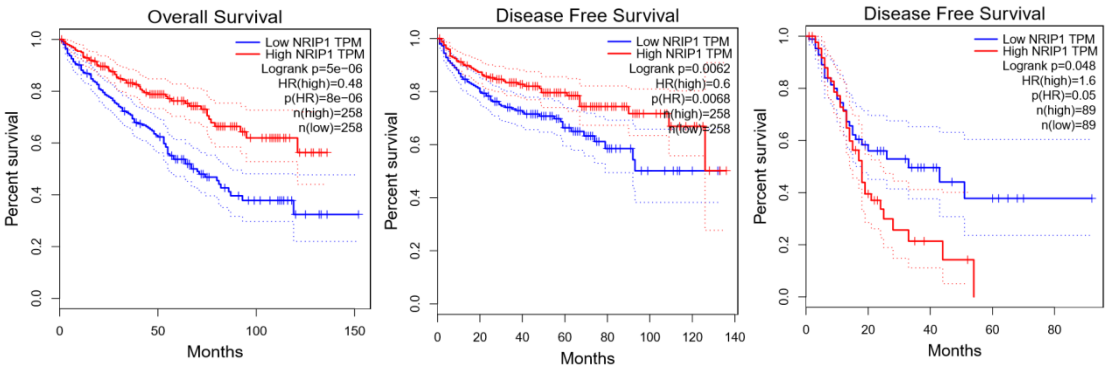


Figure 27. *NRIP1* stratified OS and DFS analysis. (A and B) OS and DFS analysis of kidney clear cell carcinoma patients stratified by *NRIP1* expression levels, relating overexpression with better prognosis. (C) DFS of pancreatic adenocarcinoma patients stratified by *NRIP1* expression levels, relating overexpression with worse DFS.

5. Discussion

In spite of having been discovered more than 40 years ago, research efforts on CC only started on the early 2000s, resulting in a myriad of discoveries linking this mechanism to several processes, with likely many more to discover. It plays a role in normal organ development in *Drosophila* through Dpp signaling and that same pathway likely plays a role in vertebrate development. BMP signaling, the vertebrate homologue of Dpp, is required for interdigital cell death, a process where certain cells suffer apoptosis to allow for normal digit formation[162]. As it stands, it could be that CC is implicated in the "selective" interdigital apoptosis, although no objective study has been conducted in this particular regard, and this is just one of many events where CC could participate that have yet to be investigated.

In addition to natural events, however, CC has also been implicated in several disorders, like cancer, where it plays a dual role as a tumour suppressor and promotor[44,46]. As a tumour suppressor, it serves as a cell extrinsic mechanism to prevent tumour development by killing cells that harbour potentially carcinogenic mutations like *HRAS-V12* [75], *ERBB2* overexpression[78], and *TP53*[74]. Alas, although CC was found to occur with DNA mutations, no study has yet revealed if epigenetic alterations that alter gene expression are sufficient to cause the same phenotype. This elimination of pre-cancer cells suggests that CC is a defence that tumours need to surmount to grow[44,46], thus being a vital step in carcinogenesis. Integrating CC in Vogelstein's theory of tumourogenesis[41] would put CC evasion as one of the earliest steps, as either mutations in oncogenes or tumour suppressors can be eliminated by the mechanism, and so the cells must first evade it[44,46]. CC can also fit into the field cancerization theory[163], as the first mutations that prevail on a tissue are those which permit CC evasion. Those cells then proliferate, and their daughter cells have the same evasion phenotype, therefore spreading throughout the tissue, laying the ground for any future oncogenic mutations to have higher potential to develop into cancer. Nevertheless, despite the initial cells requiring to escape CC to grow, it is likely that this ability is lost during cancer progression, as several cancer cell lines are able to engage in CC with each other[60,67]. After the initial onset of carcinogenic mutations are in place, the tumour then seizes CC to support its growth, by killing WT cells to garner space for its own proliferation. This also leads to more aggressive cancers, as cancer cells which are fitter can eliminate

others. Following the more recent neutral growth theory[36], this intra cancer competition only occurs on the first steps of carcinogenesis while cells have vastly different carcinogenic potential. Yet, the exact genes and other alterations that define the pathway of CC are not yet elucidated.

Hence, this project aimed to uncover novel genes responsible for CC, and other ways this process altered the behaviour of the competing cells. To that end, two colon cancer cell lines, LoVo and LS 174T, which had been previously described to engage in CC[60] were co-cultured together and their apoptosis was measured (Figure 10). LoVo cells suffered more apoptosis in co-culture and LS 174T suffered less. Following CC definitions, LoVo behaves as the loser cell and LS 174T as the winner. The co-culture was repeated with double initial seeding, and although CC was still in play, the effect was greatly diminished on the loser cells, which suggests that the confluency/space restriction inhibits this type of cell competition. This can be explained by the existence of stronger mechanical restrictions at a higher confluency, thus impeding cell mixing, which has been found to play a role in CC[100]. This hypothesis that CC is inhibited by confluency goes in accordance with observations that found it obeys organ boundaries and even regulates organ size in *Drosophila*[122]. Wing discs homozygous for *dMyc* overexpression overgrow, but in heterozygous discs, where cells are engaging in CC, wings instead grow to normal size[122], indicating a fine tuning on CC's function.

Furthermore, different proportions were also tested, both with more winners and more losers on the mix (Figure 10). When the loser cells were at a minority their elimination was even greater. These results support previous observations that state that the rate of elimination is directly proportional to the surface of cell contact between winner and loser cells[100,123], as in this condition, it is more likely that all loser cells present will be in contact with one or more winner cells. In this same proportion, the winner cells experience no difference in apoptotic levels when compared to their monoculture, which could be explained by the fact that most winners are in contact with themselves, and not with the loser cells. Supporting this, on the inverse seeding where losers are on the majority, the winners suffer lower apoptosis, which could be explained by the higher contact with the losers. Still, the loser cells continue to die more than their monoculture analogues, even though most of them are in contact with themselves. This result indicates that a single winner cell is capable of inducing apoptosis on several

surrounding loser cells, although at a diminished rate, owing to the lower overall winner cell to loser cell surface contact. The same experiment was repeated at 72 hours and similar results were found for the losers except that they no longer suffer more apoptosis when they are in the majority. As the loser cells which are not in contact with the winner cells continue to divide and proliferate, the total number of loser cells might have increased to a point where the effect of the cells which are in contact with the winner cells and die, is diluted within the now larger population, which was not seen at the 48 hours' time point. It could also be that cells have now grown to a stage where the confluency is enough to inhibit CC, as visualized on the above double seeding results. Curiously, all effects on the winner cells appear to be exacerbated at the 72 hours' time point, with all conditions featuring lower apoptosis. This suggests that there could be some sort of temporal sequence. The effects on the loser cells would come into play shortly as there is visible elimination at 24 hours, but the effects on the winner cells take some time to reach their full extent as they are more striking at the 72 than at the 48 hours' time point.

This phenomenon of CC was examined for caspase dependency, as it was expected it would be consistent with previous research[60]. Co-cultures and respective monocultures were incubated with a powerful pan-caspase inhibitor, and apoptosis was measured (Figure 12). In the presence of the pan-caspase inhibitor, loser cells no longer had an increase of apoptosis, and in fact had a lower apoptosis in co-culture when compared to monoculture, allowing to infer that loser cell elimination is indeed caspase-dependent. The winner cells maintained their lower apoptotic levels, which suggests that the effect on the winners is not loser-apoptosis dependent. The fact that both winners and losers have lower apoptosis in co-culture with emricasam, when compared to their monocultures in the same condition, proposes that this "apoptosis protection" is dependent on secreted factors. These factors would protect winner cells against noxious substances released from the loser dying cells, but would not be strong enough to protect the loser cells from winner-induced apoptosis. Since the caspases are inhibited, both cells survive and experience the "protective" substances that were secreted. SPARC, which is a secreted protein, was found to protect loser cells from apoptosis[98], and since it's extracellular, it could have a similar effect of the winners. Although a direct experiment to investigate whether or not survival factors were in play

is outside the scope of this project, a future experiment that could ascertain this rationale is the conditioned medium protocol[99]. Cells would be grown in accordance to the co-culture assay, but after 48 hours their medium would be added to newly seeded monocultures, with monoculture medium serving as control. If there are survival factors present on the co-culture medium and they are indeed responsible for the phenotype, an effect should be seen, with cells showing higher survival rates when cultured with that medium. A technique which would allow to ascertain these secreted factors is mass spectrometry, which permits the identification of all proteins present in the cell medium. The usage of stable isotope labelling by amino acids in cell culture by previously incubating one of the cell lines with an heavy isotope[164], would even permit to discern the origin of these survival factors, that is, if they are secreted by the losers or by the winners.

Another colon cancer cell line was tested to see if this CC event was conserved (Figure 11). This new cell line had no altered phenotype when co-cultured with the previous loser (LoVo), but behaved as a loser when co-cultured with the previous winner (LS 174T). The same different proportions experiment was also repeated, and the results were akin to the ones obtained in the LoVo vs LS 174T experiment. Yet, the LS 174T cells in this co-culture have an unaltered apoptosis percentage, presenting an even higher percentage when in the minority, contrasting to the previous experiment, where they always have an equal or lower percentage of cell death. The results from the co-cultures between the three cell lines indicate that the fittest cells are the LS 174T, as they always behave as winner. It is necessary to note that this fitness is not something that is easily quantified because, unlike the studies where competition relies on just one mutated gene, cancer cells have innumerous mutations between them, making it difficult to pinpoint the mutation that is responsible for CC in these cell lines. It might even be owed to epistasis, where not one but a multitude of genes contribute to the phenotype. Nonetheless, there does not seem to be any relation between fitness and cancer stage or proliferation, as LoVo have their origin on a later cancer stage than LS 174T (Stage according to ATCC) and exhibited a higher proliferation rate. Still, as a pan-caspase inhibitor completely blocked the competition between LoVo and LS 174T, but had no effect on the competition between HCT 116 and LS 174T (Figure 12), this could signify

that the quantity of the inhibitor used wasn't sufficient, or that the interaction between LS 174T and HCT 116 belongs to a different pathway of cell competition.

To investigate if apoptosis-induced compensatory proliferation was in play[135,136], a proliferation essay was conducted at 48 and 72 hours after seeding (Figure 13). At the 48 hours' timepoint, winner cells showed higher proliferation whilst losers showed lower proliferation, and at the 72 hours' timepoint only the higher proliferation on the winners is seen. Once more, this seems to fit the temporal sequence hypothesis, with the effects on the losers starting early but wearing off, and the effects on the winners taking longer to take effect but lasting for longer. To confirm if this was apoptosis-induced compensatory proliferation, the co-cultures were repeated with the pan-caspase inhibitor. It was observed that the winner cells still increased proliferation, hinting that it is not dependent on loser cell apoptosis, and that it is instead a process intrinsic to the winners, merely reliant on cell contact. The apoptosis assay results supported this, as the protective effect on the winner cells is observed even when loser cell elimination is blocked. This seems to go against published results on the same cell lines, which stated that the increase on proliferation is dependent on loser apoptosis, although this could be due to the fact that a different method was used to calculate proliferation (Phospho-Histone H3 positive cells)[60].

The cells migration profile was also measured through a scratch-wound assay (Figure 14). On the co-culture, loser cells migrated less while winner cells migrated more. This implied that CC itself aggressively transforms cells, not just through selection of the fittest. However, in co-culture both cells reach a similar migration rate, which can raise the possibility that they were "meeting halfway", maybe due to social migration[137], where different cells migrate at the same rate or one cell line drags the other along. To exclude this possibility, the blockade of social migration would be necessary, but this wasn't feasible during this study. Emricasam usage to block CC was decided against because the results showed that the effects on the winner cells are not loser apoptosis dependent. As an alternative, the two cell lines that don't engage in CC were co-cultured and the assay was repeated (Figure 15). In this co-culture, HCT 116 appears to have less migration, but LoVo seems to maintain the same level as the monoculture. In this result, cells only migrate as fast as the slowest cell present, which contradicts the "meeting halfway" phenotype of the other culture. Since these results only measure migration, a

good follow up experiment would be to use transwell plates, which allows for a better evaluation of cell invasiveness, more similar to what is observed *in vivo,* where ECM digestion is particularly important.

The results obtained from the succeeding RT-qPCR to discern genes that might be involved in CC allowed to investigate this migration change (Table 3 and Figure 19). Although not confirmed by a direct association, RT-qPCR for the EMT markers *CDH1* and *VIM* greatly support the higher migration seen on the winner cells. *VIM* upregulation and *CDH1* downregulation are commonly found on more invasive cells[139,144], and both of these alterations are seen on the winner cells in co-culture, with a striking *VIM* overexpression. Several of the gene hits obtained also seem to reinforce the idea that CC influences migration, as several of them have been linked to migration and invasion. Overexpression of *AQP3*[148,149], *GPX2*[165] or *NRIP1*[159] has been linked to migration and all of those have augmented expression in the winners in co-culture, although *TRPM8*, also augmented in the winners, has been reported to lower migration[166]. On the loser cells, one of the genes which was downregulated, *LAMC2*, was found to be negatively related to invasion[167]. Therefore, the EMT markers, the upregulation of these genes in the winner and the downregulation of *LAMC2* on the loser, could explain the results obtained, with winners migrating more, and losers migrating less in co-culture, independently of those genes being direct effectors in CC or not.

Since CC is theoretically taking place throughout the entirety of cancer progression, it would be useful to determine how it is modulated on the onset of chemotherapy regimens, known to be greatly cytotoxic[138]. On this account, the co-culture was repeated in the presence of oxaliplatin with surprising results (Figure 16). The CC event was apparently inverted, with LoVo cells behaving as winners, with lower apoptosis, and LS 174T as losers, with higher apoptosis in co-culture. Regardless, seeing as the control of this experiment did not exhibit CC, results are still preliminary and warrant repetitions, possibly with chemotherapy dosage to determine if the effect observed is proportional. This higher apoptosis seen on the winner cells could be explained by their higher proliferation in co-culture, which would make them more susceptible to an intercalating agent like oxaliplatin, linking CC to chemotherapy response for the first time.

Notwithstanding, it is unlikely that changes in proliferation are responsible for the effect observed on the loser cells, that remarkably survive almost three times as better

in co-culture in the presence of oxaliplatin. This higher LoVo viability, when taken together with the higher LS 174T apoptosis, suggests that these effects could be due to an inversion of the CC event. A possible explanation is that the addition of oxaliplatin to the co-culture changes the relative fitness, which serves as the basis for CC. In the presence of oxaliplatin, LoVo cells are now considered to be the fitter, and could therefore be acting as winners, with LS 174T behaving as the new losers. Although not confirmed in this project, this could be owing to the fact that the LoVo cell line has been noted to be more resistant to oxaliplatin than the LS 174T cell line[168,169], but the opposite has also been described[170], warranting further elucidation. Regardless if this is or not the case, to conclude if CC is indeed in play, this same co-culture needs to be repeated with Emricasam. If this were to be confirmed, it would reveal an exciting new side of CC. While its context dependent nature was already described[46], it was mostly due to cell intrinsic factors. That is, the same cells can behave as winner or losers depending on their competing counterparts. However, the same exact pair of cells had seldom been described to alter their phenotype. These results, although preliminary, suggest that chemotherapy, similar to what was observed with obesity[115] and hyperinsulinemia[117], might constitute an extrinsic factor that is also able to influence CC, opening a door for potential adjustment of CC through microenvironment modulation. If this novel implication of CC in chemotherapy is verified, it may also improve chemotherapy regiments, as CC's intensification would augment the tumoural cells proliferation, which in turn would lead to a higher susceptibility to certain cytotoxic agents like oxaliplatin, improving patient response. Furthermore, CC may play a critical role in acquired treatment resistance. According to the neutral growth theory, clonal selection only occurs when there are no space constraints, or when cells are under a great selective pressure like chemotherapy[35,36]. This could explain why intermittent therapies thwart the development of resistances[171]. During chemotherapy treatment, resistant cancer cells survive, resulting in a population of cells that no longer respond to the therapy[171]. However, if the therapy is halted midway, the non-resistant cells resume their proliferation. In a cell competition scenario, if the loser cells are resistant to the chemotherapy, the initial round of treatment kills susceptible cells, conferring free space, while pausing the treatment allows the winner cells to proliferate and engage in CC, thus outcompeting and eliminating the resistant cells now that spatial limitations

are no longer in play. Therefore, the resistant population never becomes the bulk of the tumour, forestalling the emergence of therapeutic resistances.

Finally, RNA-Seq was done to unravel the genes responsible for CC (Table 2 and Figure 18). After validation through RT-qPCR, the final list of gene hits was obtained: *CYP1A1*, *LAMC2* and *MSMO1*, downregulated on the loser cells; *AQP3*, *GPX2*, *MYT1*, *NRIP1*, *PDK4*, *SKIL*, *TRPM8*, *UGT1A1* and *UGT1A10*, upregulated on the winner cells. Adding to this, these genes showed a similar pattern of upregulation and downregulation at the 72 hours' timepoint. To elucidate if the same genes were responsible for the CC phenomenon that occurs in the HCT 116 vs LS 174T co-culture, a RT-qPCR was ran for these samples. Contrariwise to LS 174T with LoVo co-cultures, the genes either exhibited an inverse alteration, an attenuated upregulation/dowregulation, or no change in their expression whatsoever. This indicates that what happens in the LS 174T with HCT 116 co-culture is not the same exact CC visible on the other co-culture, which goes in accordance to the caspase-independent apoptosis of this CC.

To determine additional pathways which could explain the phenotypic changes observed on the LoVo and LS 174T cell lines, a RT-qPCR was run for genes central to other signalling pathways (Table 3 and Figure 19). This analysis allowed to infer that winner cells do migrate more during CC, owing to the presence of EMT markers, and that the reason the winner cells survive better and proliferate more could be due to the activation of the EGFR, MTOR and JAK/STAT pathways. Conversely, the same pathways being inactive in loser cells would explain the lower proliferation and migration rate present in these cells. A big caveat of these RT-qPCR is that it's only measuring the transcription of the gene, and many pathways simply require the modification of proteins to be activated. To prove that the pathways described are really activated, further experiments using, for example, western blot with antibodies against phosphorylated and non-phosphorylated proteins, are required.

To confirm whether or not the gene hits are necessary for cell competition, winner cell genes were downregulated using shRNAs, which are able to silence gene expression through target mRNA cleavage[172]. After lentiviral transduction with the selected shRNAs, RNA extraction followed by RT-qPCR was done to determine if the genes were downregulated (Figure 20). Afterwards, the LS 174T cells treated with the validated

shRNAs were cultured with LoVo cells, and apoptosis was measured after 48 hours (Figure 21). Of those experiments, three shRNAs led to a blockade of CC, resulting in both cells surviving better in co-culture, reiterating what was seen on the presence of the pan-caspase inhibitor. Since CC was apparently blocked, it can be stated that the upregulation of either of these 3 genes, *AQP3*, *MYT1* and *NRIP1*, is shown for the first time to be necessary for the mechanism. It is not clear if this upregulation is sufficient, as that would require the overexpression of the gene on a cell line, and a posterior co-culture with the same cell line used. If CC were to occur in that scenario, it would mean that the gene upregulation is necessary and sufficient. The loser genes were not tested yet, but a similar rationale follows. Loser cells with upregulation of the gene would be co-cultured with the winner cell line. If CC was blocked, then that gene is necessary for the transduction of the apoptotic signal. To check if it was sufficient, the gene would be downregulated in a cell line, and that modified cell line would be co-cultured with its WT self. If CC occurs, then the gene downregulation is sufficient to initiate the mechanism.

Although this shRNA mediated blockade of CC was enough to prevent loser elimination, it did not block the enhanced winner survival. In fact, loser cells also survive better in co-culture when CC is blocked, which strongly suggests the presence of soluble factors, mirroring the results seen with the caspase inhibitor. This indicates that the obtained genes only play a role in the elimination of loser cells, but not in the production of survival factors, suggesting that the upregulation of these genes is upstream of loser cell elimination, but downstream of the theorized survival factor secretion. The results suggest that, when winner cells contact with loser cells, survival factors are produced. The winner cells then uptake these survival factors and use them to improve their own survival. When loser elimination is inhibited, survival factors are still produced, but since loser cells are not eliminated, they are also under the effects of the survival factors, resulting in lower apoptosis for both cell lines. The tests described above are proposed as valuable follow up studies to further unravel the mechanism.

A bioinformatics analysis was also conducted to find out if these gene hits were implicated in cancer biology (Figure 22 to Figure 27). Datasets were procured and analysed for expression and differential survival. However, it is necessary to note that, even if an effect is seen, it cannot be attributed to CC, as theoretically it is mostly taking place at the boundary between tumour and “healthy” tissue. As such, to visualize any

alteration, one must compare the non-boundary tumour with the boundary tumour, or non-adjacent normal tissue with adjacent normal tissue. Since there were no sufficient public datasets (of high output sequencing) to permit comparison between the two different areas, only the mentioned expression analysis was conducted. The genes did not have consistently higher or lower expression across cancer types, with *AQP3* and *NRIP1* downregulated in some types, and *MYT1* overexpressed on others. The overexpression of winner genes was, however, consistently associated with pancreatic adenocarcinoma prognosis, with *AQP3* and *MYT1* overexpressing linked to better prognosis, albeit *NRIP1* overexpression lead to worse DFS. This cancer is known for having an hypoxic microenvironment[173], but since hypoxia cannot be directly connected to the objective of this study, the information presented here is simply descriptive. Yet, how these genes contribute to CC is still not understood, and any hypothesis to explain so was deemed to be too speculative to be included here.

All in all, this study fulfilled its aim of revealing novel genes implicated in CC, a process which can be implicated in several steps of carcinogenesis. The development of inhibitors against the upregulated proteins on the winner cells can stop cancer from seizing this mechanism to fuel its growth, thus impeding cancer progression *in situ*. As a membrane channel, *AQP3* constitutes a logical target for the development of antibody therapies.

A suggestion for follow-up research would be to repeat these same experiments using different cells from different cancers, which would allow to determine if a similar phenomenon, and these same genes, are conserved in other neoplasms. The importance of context in CC also indicates that it could also play a role in the seed and soil hypothesis[174], and why certain tumours exhibit a predilection to form metastasis on specific organs. A cancer cell can be extruded from the epithelium, where it behaves as a loser, through EDAC[52], afterwards undergoing anoikis[175]. Yet, if the cancer cell evades anoikis[175], it may then migrate into another tissue, and if in that tissue the cell is a winner, it will “seed” and create a metastasis. A similar process occurs with transplanted cells, which rely on cell competition to engraft[57,73].

The implication of CC in response to chemotherapy is also rather exciting, as on one hand, it can improve therapeutic response. Since CC leads to a higher proliferation of the winner cells, it can also make these cells more vulnerable to chemotherapy. The

upregulation of the gene targets described on the tumour through gene therapy would enhance CC, leading to higher proliferation of these cells, culminating in a better response to chemotherapy. On the other hand, although still mere conjecture, it is thought provoking to consider what would happen if the loser cells were more resistant to therapy. If the above-mentioned theory of CC inversion is confirmed, then under chemotherapy, these resistant losers could use CC to further enhance non-resistant winner cell elimination, using it as a "parasitic" process to augment their proliferation capabilities. This would eventually result in a tumour comprised of only the resistant cells. If it were feasible to stop this process, resistant loser cells would still survive, but wouldn't have their proliferation increased, which could lead to a better DFS. The discovery of new molecular targets in this project reveals novel ways in which CC could be blocked, without the severe secondary effects of pan-caspase inhibitors[176], thus preventing this chemoresistance from emerging.

As such, fully understanding the phenomenon of cell competition would be ground-breaking, permitting to tackle cancer in all its steps. Manipulating CC would provide novel approaches to prevent the evading of the tumour suppressor mechanism of CC, to stop the seizing of this mechanism to grow in situ and to hinder metastasis establishment. Lastly, by using CC as a prevention mechanism against therapy resistance, or as a therapy enhancer, treatment efficiency can be improved.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
2. Cancer Incidence. 2017. (Accessed August, 25th, 2020, at https://ourworldindata.org/grapher/cancer-incidence.)
3. Institute for Health Metrics and Evaluation (IHME). Findings from the Global Burden of Disease Study 2017. The Lancet 2018.
4. Human Development Index (HDI). 2020. (Accessed September, 7th, 2020, at http://hdr.undp.org/en/content/statistical-data-tables-7-15.)
5. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69:7-34.
6. The top 10 causes of death. 2018. (Accessed August, 25th, 2020, at https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.)
7. What do people die from? 2018. (Accessed August, 25th, 2020, at https://ourworldindata.org/what-does-the-world-die-from.)
8. Hankey BF, Feuer EJ, Clegg LX, et al. Cancer surveillance series: interpreting trends in prostate cancer--part I: Evidence of the effects of screening in recent prostate cancer incidence, mortality, and survival rates. J Natl Cancer Inst 1999;91:1017-24.
9. Vaccarella S, Franceschi S, Bray F, Wild CP, Plummer M, Dal Maso L. Worldwide Thyroid-Cancer Epidemic? The Increasing Impact of Overdiagnosis. N Engl J Med 2016;375:614-7.
10. Rebbeck TR. Handbook for cancer research in Africa: WHO Regional Office for Africa; 2013.
11. Rebbeck TR, Devesa SS, Chang BL, et al. Global patterns of prostate cancer incidence, aggressiveness, and mortality in men of african descent. Prostate Cancer 2013;2013:560857.
12. Cancer. 2015. (Accessed September, 8th, 2020, at https://ourworldindata.org/cancer.)
13. Peacock A, Leung J, Larney S, et al. Global statistics on alcohol, tobacco and illicit drug use: 2017 status report. Addiction 2018;113:1905-26.
14. Organization WH. WHO global report on trends in prevalence of tobacco smoking 2000-2025: World Health Organization; 2018.
15. Daily meat consumption per person. 2013. (Accessed September, 8th, 2020, at https://ourworldindata.org/grapher/daily-meat-consumption-per-person?time=1961..2013.)
16. Chan DS, Lau R, Aune D, et al. Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. PLoS One 2011;6:e20456.
17. Hackshaw AK, Law MR, Wald NJ. The accumulated evidence on lung cancer and environmental tobacco smoke. BMJ 1997;315:980-8.
18. Chang Y, Moore PS, Weiss RA. Human oncogenic viruses: nature and discovery. Philos Trans R Soc Lond B Biol Sci 2017;372.
19. Bruni L, Diaz M, Barrionuevo-Rosas L, et al. Global estimates of human papillomavirus vaccination coverage by region and income level: a pooled analysis. The Lancet Global Health 2016;4:e453-e63.
20. Ott J, Stevens G, Groeger J, Wiersma S. Global epidemiology of hepatitis B virus infection: new estimates of age-specific HBsAg seroprevalence and endemicity. Vaccine 2012;30:2212-9.
21. Wu S, Powers S, Zhu W, Hannun YA. Substantial contribution of extrinsic risk factors to cancer development. Nature 2016;529:43-7.
22. Kwan W, Jackson J, Weir LM, Dingee C, McGregor G, Olivotto IA. Chronic arm morbidity after curative breast cancer treatment: prevalence and impact on quality of life. J Clin Oncol 2002;20:4242-8.

23. Feiten S, Dunnebacke J, Heymanns J, et al. Breast cancer morbidity: questionnaire survey of patients on the long term effects of disease and adjuvant therapy. Dtsch Arztebl Int 2014;111:537-44.
24. Nayak MG, George A, Vidyasagar MS, et al. Quality of Life among Cancer Patients. Indian J Palliat Care 2017;23:445-50.
25. Cancer. 2018. (Accessed September, 9th, 2020, at https://www.who.int/health-topics/cancer.)
26. Bill-Axelson A, Holmberg L, Garmo H, et al. Radical Prostatectomy or Watchful Waiting in Prostate Cancer - 29-Year Follow-up. N Engl J Med 2018;379:2319-29.
27. Kleeff J, Korc M, Apte M, et al. Pancreatic cancer. Nat Rev Dis Primers 2016;2:16022.
28. Zheng H-C. The molecular mechanisms of chemoresistance in cancers. Oncotarget 2017;8:59950.
29. Housman G, Byler S, Heerboth S, et al. Drug resistance in cancer: an overview. Cancers 2014;6:1769-92.
30. Zahreddine H, Borden K. Mechanisms and insights into drug resistance in cancer. Frontiers in pharmacology 2013;4:28.
31. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. Nature reviews Clinical oncology 2018;15:81.
32. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell 2017;168:613-28.
33. Allemani C, Weir HK, Carreira H, et al. Global surveillance of cancer survival 1995–2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). The Lancet 2015;385:977-1010.
34. Five year survival rates by cancer type, Portugal, 2009. 2009. (Accessed September, 14th, 2020, at https://ourworldindata.org/grapher/five-year-survival-rates-by-cancer-type?year=latest&country=~PRT.)
35. Sottoriva A, Kang H, Ma Z, et al. A Big Bang model of human colorectal tumor growth. Nature genetics 2015;47:209-16.
36. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. Nature genetics 2016;48:238-44.
37. Grabocka E, Bar-Sagi D. Mutant KRAS Enhances Tumor Cell Fitness by Upregulating Stress Granules. Cell 2016;167:1803-13 e12.
38. Bozic I, Reiter JG, Allen B, et al. Evolutionary dynamics of cancer in response to targeted combination therapy. Elife 2013;2:e00747.
39. Darwin C, Bynum WF. The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life: AL Burt New York; 2009.
40. Nowell PC. The clonal evolution of tumor cell populations. Science 1976;194:23-8.
41. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759-67.
42. Ling S, Hu Z, Yang Z, et al. Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution. Proc Natl Acad Sci U S A 2015;112:E6496-505.
43. Almendro V, Cheng YK, Randles A, et al. Inference of tumor evolution during chemotherapy by computational modeling and in situ analysis of genetic and phenotypic cellular diversity. Cell Rep 2014;6:514-27.
44. Moreno E. Is cell competition relevant to cancer? Nat Rev Cancer 2008;8:141-7.
45. Moreno E, Basler K, Morata G. Cells compete for decapentaplegic survival factor to prevent apoptosis in Drosophila wing development. Nature 2002;416:755-9.
46. Merino MM, Levayer R, Moreno E. Survival of the Fittest: Essential Roles of Cell Competition in Development, Aging, and Cancer. Trends Cell Biol 2016;26:776-88.
47. Morata G, Ripoll P. Minutes: mutants of drosophila autonomously affecting cell division rate. Dev Biol 1975;42:211-21.

48. Kolahgar G, Suijkerbuijk SJ, Kucinski I, et al. Cell Competition Modifies Adult Stem Cell and Tissue Population Dynamics in a JAK-STAT-Dependent Manner. Dev Cell 2015;34:297-309.
49. Tamori Y, Deng WM. Tissue repair through cell competition and compensatory cellular hypertrophy in postmitotic epithelia. Dev Cell 2013;25:350-63.
50. Coelho DS, Schwartz S, Merino MM, et al. Culling Less Fit Neurons Protects against Amyloid-beta-Induced Brain Damage and Cognitive and Motor Decline. Cell Rep 2018;25:3661-73 e3.
51. Eichenlaub T, Cohen SM, Herranz H. Cell Competition Drives the Formation of Metastatic Tumors in a Drosophila Model of Epithelial Tumor Formation. Curr Biol 2016;26:419-27.
52. Kajita M, Fujita Y. EDAC: Epithelial defence against cancer-cell competition between normal and transformed epithelial cells in mammals. J Biochem 2015;158:15-23.
53. Martins VC, Busch K, Juraeva D, et al. Cell competition is a tumour suppressor mechanism in the thymus. Nature 2014;509:465-70.
54. Petrova E, Soldini D, Moreno E. The expression of SPARC in human tumors is consistent with its role during cell competition. Commun Integr Biol 2011;4:171-4.
55. Kongsuwan K, Yu Q, Vincent A, et al. A Drosophila Minute gene encodes a ribosomal protein. Nature 1985;317:555-8.
56. Lambertsson A. 3 The Minute Genes in Drosophila and Their Molecular Functions. Advances in genetics: Elsevier; 1998:69-134.
57. Oertel M, Menthena A, Dabeva MD, Shafritz DA. Cell competition leads to a high level of normal liver reconstitution by transplanted fetal liver stem/progenitor cells. Gastroenterology 2006;130:507-20; quiz 90.
58. Oliver ER, Saunders TL, Tarle SA, Glaser T. Ribosomal protein L24 defect in belly spot and tail (Bst), a mouse Minute. Development 2004;131:3907-20.
59. Tamori Y, Bialucha CU, Tian AG, et al. Involvement of Lgl and Mahjong/VprBP in cell competition. PLoS Biol 2010;8:e1000422.
60. Di Giacomo S, Sollazzo M, de Biase D, et al. Human Cancer Cells Signal Their Competitive Fitness Through MYC Activity. Sci Rep 2017;7:12568.
61. Brumby AM, Richardson HE. scribble mutants cooperate with oncogenic Ras or Notch to cause neoplastic overgrowth in Drosophila. EMBO J 2003;22:5769-79.
62. Norman M, Wisniewska KA, Lawrenson K, et al. Loss of Scribble causes cell competition in mammalian cells. J Cell Sci 2012;125:59-66.
63. Wagstaff L, Goschorska M, Kozyrska K, et al. Mechanical cell competition kills cells via induction of lethal p53 levels. Nat Commun 2016;7:11373.
64. Thompson BJ, Mathieu J, Sung HH, Loeser E, Rorth P, Cohen SM. Tumor suppressor properties of the ESCRT-II complex component Vps25 in Drosophila. Dev Cell 2005;9:711-20.
65. Rhiner C, Lopez-Gay JM, Soldini D, et al. Flower forms an extracellular code that reveals the fitness of a cell to its neighbors in Drosophila. Dev Cell 2010;18:985-98.
66. Moreno E, Basler K. dMyc transforms cells into super-competitors. Cell 2004;117:117-29.
67. Patel MS, Shah HS, Shrivastava N. c-Myc-Dependent Cell Competition in Human Cancer Cells. J Cell Biochem 2017;118:1782-91.
68. Neto-Silva RM, de Beco S, Johnston LA. Evidence for a growth-stabilizing regulatory feedback mechanism between Myc and Yorkie, the Drosophila homolog of Yap. Dev Cell 2010;19:507-20.
69. Chiba T, Ishihara E, Miyamura N, et al. MDCK cells expressing constitutively active Yes-associated protein (YAP) undergo apical extrusion depending on neighboring cell status. Sci Rep 2016;6:28383.
70. Rodrigues AB, Zoranovic T, Ayala-Camargo A, et al. Activated STAT regulates growth and induces competitive interactions independently of Myc, Yorkie, Wingless and ribosome biogenesis. Development 2012;139:4051-61.

71. Vincent JP, Kolahgar G, Gagliardi M, Piddini E. Steep differences in wingless signaling trigger Myc-independent competitive cell interactions. Dev Cell 2011;21:366-74.
72. Suijkerbuijk SJ, Kolahgar G, Kucinski I, Piddini E. Cell Competition Drives the Growth of Intestinal Adenomas in Drosophila. Curr Biol 2016;26:428-38.
73. Bondar T, Medzhitov R. p53-mediated hematopoietic stem and progenitor cell competition. Cell Stem Cell 2010;6:309-22.
74. Watanabe H, Ishibashi K, Mano H, et al. Mutant p53-Expressing Cells Undergo Necroptosis via Cell Competition with the Neighboring Normal Epithelial Cells. Cell Rep 2018;23:3721-9.
75. Hogan C, Dupre-Crochet S, Norman M, et al. Characterization of the interface between normal and transformed epithelial cells. Nat Cell Biol 2009;11:460-7.
76. Kajita M, Hogan C, Harris AR, et al. Interaction with surrounding normal epithelial cells influences signalling pathways and behaviour of Src-transformed cells. J Cell Sci 2010;123:171-80.
77. Grieve AG, Rabouille C. Extracellular cleavage of E-cadherin promotes epithelial cell extrusion. J Cell Sci 2014;127:3331-46.
78. Leung CT, Brugge JS. Outgrowth of single oncogene-expressing cells from suppressive epithelial environments. Nature 2012;482:410-3.
79. HomoloGene. 2020. (Accessed 30th October, 2020, at https://www.ncbi.nlm.nih.gov/homologene.)
80. Fly Base. 2020. (Accessed 30th October, 2020, at https://flybase.org/.)
81. Martin-Castellanos C, Edgar BA. A characterization of the effects of Dpp signaling on cell growth and proliferation in the Drosophila wing. Development 2002;129:1003-13.
82. Adachi-Yamada T, O'Connor MB. Morphogenetic apoptosis: a mechanism for correcting discontinuities in morphogen gradients. Dev Biol 2002;251:74-90.
83. Giraldez AJ, Cohen SM. Wingless and Notch signaling provide cell survival cues and control cell proliferation during wing development. Development 2003;130:6533-43.
84. Johnston LA, Sanders AL. Wingless promotes cell survival but constrains growth during Drosophila wing development. Nat Cell Biol 2003;5:827-33.
85. Moreno E. Cancer: Darwinian tumour suppression. Nature 2014;509:435-6.
86. Orr AW, Helmke BP, Blackman BR, Schwartz MA. Mechanisms of mechanotransduction. Dev Cell 2006;10:11-20.
87. Baratchi S, Khoshmanesh K, Woodman OL, Potocnik S, Peter K, McIntyre P. Molecular Sensors of Blood Flow in Endothelial Cells. Trends Mol Med 2017;23:850-68.
88. Lyon RC, Zanella F, Omens JH, Sheikh F. Mechanotransduction in cardiac hypertrophy and failure. Circ Res 2015;116:1462-76.
89. Holick MF. Perspective on the impact of weightlessness on calcium and bone metabolism. Bone 1998;22:105S-11S.
90. Bras-Pereira C, Moreno E. Mechanical cell competition. Curr Opin Cell Biol 2018;51:15-21.
91. Shraiman BI. Mechanical feedback as a possible regulator of tissue growth. Proc Natl Acad Sci U S A 2005;102:3318-23.
92. Legoff L, Rouault H, Lecuit T. A global pattern of mechanical stress polarizes cell divisions and cell shape in the growing Drosophila wing disc. Development 2013;140:4051-9.
93. Mao Y, Tournier AL, Hoppe A, Kester L, Thompson BJ, Tapon N. Differential proliferation rates generate patterns of mechanical tension that orient tissue growth. EMBO J 2013;32:2790-803.
94. Eisenhoffer GT, Loftus PD, Yoshigi M, et al. Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. Nature 2012;484:546-9.
95. Marinari E, Mehonic A, Curran S, Gale J, Duke T, Baum B. Live-cell delamination counterbalances epithelial growth to limit tissue overcrowding. Nature 2012;484:542-5.
96. Levayer R, Dupont C, Moreno E. Tissue Crowding Induces Caspase-Dependent Competition for Space. Curr Biol 2016;26:670-7.

97. Vincent JP, Fletcher AG, Baena-Lopez LA. Mechanisms and mechanics of cell competition in epithelia. Nat Rev Mol Cell Biol 2013;14:581-91.
98. Portela M, Casas-Tinto S, Rhiner C, et al. Drosophila SPARC is a self-protective signal expressed by loser cells during cell competition. Dev Cell 2010;19:562-73.
99. Senoo-Matsuda N, Johnston LA. Soluble factors mediate competitive and cooperative interactions between cells expressing different levels of Drosophila Myc. Proc Natl Acad Sci U S A 2007;104:18543-8.
100. Levayer R, Hauert B, Moreno E. Cell mixing induced by myc is required for competitive tissue invasion and destruction. Nature 2015;524:476-80.
101. Penzo-Mendez AI, Chen YJ, Li J, Witze ES, Stanger BZ. Spontaneous Cell Competition in Immortalized Mammalian Cell Lines. PLoS One 2015;10:e0132437.
102. Coelho DS, Moreno E. Emerging links between cell competition and Alzheimer's disease. J Cell Sci 2019;132.
103. Coelho DS, Moreno E. Neuronal Selection Based on Relative Fitness Comparison Detects and Eliminates Amyloid-beta-Induced Hyperactive Neurons in Drosophila. iScience 2020;23:101468.
104. Merino MM, Rhiner C, Lopez-Gay JM, Buechel D, Hauert B, Moreno E. Elimination of unfit cells maintains tissue health and prolongs lifespan. Cell 2015;160:461-76.
105. Igaki T. Correcting developmental errors by apoptosis: lessons from Drosophila JNK signaling. Apoptosis 2009;14:1021-8.
106. Bos JL. ras oncogenes in human cancer: a review. Cancer Res 1989;49:4682-9.
107. Kon S, Ishibashi K, Katoh H, et al. Cell competition with normal epithelial cells promotes apical extrusion of transformed cells through metabolic changes. Nat Cell Biol 2017;19:530-41.
108. Hunter T, Sefton BM. Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc Natl Acad Sci U S A 1980;77:1311-5.
109. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 1987;235:177-82.
110. Zanconato F, Cordenonsi M, Piccolo S. YAP/TAZ at the Roots of Cancer. Cancer Cell 2016;29:783-803.
111. Kamai T, Yamanishi T, Shirataki H, et al. Overexpression of RhoA, Rac1, and Cdc42 GTPases is associated with progression in testicular cancer. Clin Cancer Res 2004;10:4799-805.
112. Calle EE, Thun MJ. Obesity and cancer. Oncogene 2004;23:6365-78.
113. Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nat Rev Cancer 2004;4:579-91.
114. Avgerinos KI, Spyrou N, Mantzoros CS, Dalamaga M. Obesity and cancer risk: Emerging biological mechanisms and perspectives. Metabolism 2019;92:121-35.
115. Sasaki A, Nagatake T, Egami R, et al. Obesity Suppresses Cell-Competition-Mediated Apical Elimination of RasV12-Transformed Cells from Epithelial Tissues. Cell Rep 2018;23:974-82.
116. Giovannucci E, Harlan DM, Archer MC, et al. Diabetes and cancer: a consensus report. Diabetes Care 2010;33:1674-85.
117. Sanaki Y, Nagata R, Kizawa D, Leopold P, Igaki T. Hyperinsulinemia Drives Epithelial Tumorigenesis by Abrogating Cell Competition. Dev Cell 2020;53:379-89 e5.
118. Muller PA, Vousden KH. p53 mutations in cancer. Nat Cell Biol 2013;15:2-8.
119. Menendez J, Perez-Garijo A, Calleja M, Morata G. A tumor-suppressing mechanism in Drosophila involving cell competition and the Hippo pathway. Proc Natl Acad Sci U S A 2010;107:14651-6.
120. Igaki T, Pastor-Pareja JC, Aonuma H, Miura M, Xu T. Intrinsic tumor suppression and epithelial maintenance by endocytic activation of Eiger/TNF signaling in Drosophila. Dev Cell 2009;16:458-65.
121. Dang CV. MYC on the path to cancer. Cell 2012;149:22-35.

122. de la Cova C, Abril M, Bellosta P, Gallant P, Johnston LA. Drosophila myc regulates organ size by inducing cell competition. Cell 2004;117:107-16.
123. Levayer R, Moreno E. How to be in a good shape? The influence of clone morphology on cell competition. Commun Integr Biol 2016;9:e1102806.
124. Morin PJ, Sparks AB, Korinek V, et al. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. Science 1997;275:1787-90.
125. Sigismund S, Avanzato D, Lanzetti L. Emerging functions of the EGFR in cancer. Mol Oncol 2018;12:3-20.
126. Petrova E, Lopez-Gay JM, Rhiner C, Moreno E. Flower-deficient mice have reduced susceptibility to skin papilloma formation. Dis Model Mech 2012;5:553-61.
127. Ryoo HD, Bergmann A. The role of apoptosis-induced proliferation for regeneration and cancer. Cold Spring Harb Perspect Biol 2012;4:a008797.
128. Camacho D, Jesus JP, Palma AM, et al. SPARC-p53: The double agents of cancer. Adv Cancer Res 2020;148:171-99.
129. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. Genome research 2002;12:996-1006.
130. Primer-BLAST. (Accessed 3rd November, 2020, at https://www.ncbi.nlm.nih.gov/tools/primer-blast/.)
131. RNA-Seq Quantification. (Accessed November, 20th, 2020, at https://gdc.cancer.gov/about-data/data-harmonization-and-generation/genomic-data-harmonization/high-level-data-generation/rna-seq-quantification.)
132. Fabregat A, Jupe S, Matthews L, et al. The reactome pathway knowledgebase. Nucleic acids research 2018;46:D649-D55.
133. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic acids research 2017;45:W98-W102.
134. Vermes I, Haanen C, Steffens-Nakken H, Reutelingsperger C. A novel assay for apoptosis. Flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled Annexin V. J Immunol Methods 1995;184:39-51.
135. Ryoo HD, Gorenc T, Steller H. Apoptotic cells can induce compensatory cell proliferation through the JNK and the Wingless signaling pathways. Dev Cell 2004;7:491-501.
136. Maeda S, Kamata H, Luo JL, Leffert H, Karin M. IKKbeta couples hepatocyte death to cytokine-driven compensatory proliferation that promotes chemical hepatocarcinogenesis. Cell 2005;121:977-90.
137. Vedel S, Tay S, Johnston DM, Bruus H, Quake SR. Migration of cells in a social context. Proc Natl Acad Sci U S A 2013;110:129-34.
138. Miller KD, Nogueira L, Mariotto AB, et al. Cancer treatment and survivorship statistics, 2019. CA Cancer J Clin 2019;69:363-85.
139. Brabletz T, Kalluri R, Nieto MA, Weinberg RA. EMT in cancer. Nat Rev Cancer 2018;18:128-34.
140. McCubrey JA, Steelman LS, Chappell WH, et al. Roles of the Raf/MEK/ERK pathway in cell growth, malignant transformation and drug resistance. Biochim Biophys Acta 2007;1773:1263-84.
141. Saxton RA, Sabatini DM. mTOR Signaling in Growth, Metabolism, and Disease. Cell 2017;168:960-76.
142. Johnson DE, O'Keefe RA, Grandis JR. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. Nat Rev Clin Oncol 2018;15:234-48.
143. Harari D, Yarden Y. Molecular mechanisms underlying ErbB2/HER2 action in breast cancer. Oncogene 2000;19:6102-14.
144. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol 2014;15:178-96.
145. Sato N, Yako Y, Maruyama T, et al. The COX-2/PGE2 pathway suppresses apical elimination of RasV12-transformed cells from epithelia. Commun Biol 2020;3:132.

146. Hayashi Y, Yokota A, Harada H, Huang G. Hypoxia/pseudohypoxia-mediated activation of hypoxia-inducible factor-1alpha in cancer. Cancer Sci 2019;110:1510-7.
147. Marlar S, Jensen HH, Login FH, Nejsum LN. Aquaporin-3 in Cancer. Int J Mol Sci 2017;18.
148. Xu H, Xu Y, Zhang W, Shen L, Yang L, Xu Z. Aquaporin-3 positively regulates matrix metalloproteinases via PI3K/AKT signal pathway in human gastric carcinoma SGC7901 cells. J Exp Clin Cancer Res 2011;30:86.
149. Chen J, Wang Z, Xu D, Liu Y, Gao Y. Aquaporin 3 promotes prostate cancer cell motility and invasion via extracellular signal-regulated kinase 1/2-mediated matrix metalloproteinase-3 secretion. Mol Med Rep 2015;11:2882-8.
150. Hara-Chikuma M, Verkman AS. Prevention of skin tumorigenesis and impairment of epidermal cell proliferation by targeted aquaporin-3 gene disruption. Mol Cell Biol 2008;28:326-32.
151. Satooka H, Hara-Chikuma M. Aquaporin-3 Controls Breast Cancer Cell Migration by Regulating Hydrogen Peroxide Transport and Its Downstream Cell Signaling. Mol Cell Biol 2016;36:1206-18.
152. Mueller PR, Coleman TR, Kumagai A, Dunphy WG. Myt1: a membrane-associated inhibitory kinase that phosphorylates Cdc2 on both threonine-14 and tyrosine-15. Science 1995;270:86-90.
153. Liu F, Stanton JJ, Wu Z, Piwnica-Worms H. The human Myt1 kinase preferentially phosphorylates Cdc2 on threonine 14 and localizes to the endoplasmic reticulum and Golgi complex. Mol Cell Biol 1997;17:571-83.
154. Booher RN, Holman PS, Fattaey A. Human Myt1 is a cell cycle-regulated kinase that inhibits Cdc2 but not Cdk2 activity. J Biol Chem 1997;272:22300-6.
155. Wang Y, Decker SJ, Sebolt-Leopold J. Knockdown of Chk1, Wee1 and Myt1 by RNA interference abrogates G2 checkpoint and induces apoptosis. Cancer Biol Ther 2004;3:305-13.
156. Melhuish TA, Kowalczyk I, Manukyan A, et al. Myt1 and Myt1l transcription factors limit proliferation in GBM cells by repressing YAP1 expression. Biochim Biophys Acta Gene Regul Mech 2018;1861:983-95.
157. Rosell M, Jones MC, Parker MG. Role of nuclear receptor corepressor RIP140 in metabolic syndrome. Biochim Biophys Acta 2011;1812:919-28.
158. Aziz MH, Chen X, Zhang Q, et al. Suppressing NRIP1 inhibits growth of breast cancer cells in vitro and in vivo. Oncotarget 2015;6:39714-24.
159. Zhang D, Wang Y, Dai Y, et al. Downregulation of RIP140 in hepatocellular carcinoma promoted the growth and migration of the cancer cells. Tumour Biol 2015;36:2077-85.
160. Lapierre M, Bonnet S, Bascoul-Mollevi C, et al. RIP140 increases APC expression and controls intestinal homeostasis and tumorigenesis. J Clin Invest 2014;124:1899-913.
161. Carroll JS, Meyer CA, Song J, et al. Genome-wide analysis of estrogen receptor binding sites. Nat Genet 2006;38:1289-97.
162. Zou H, Niswander L. Requirement for BMP signaling in interdigital apoptosis and scale formation. Science 1996;272:738-41.
163. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer 1953;6:963-8.
164. Ong SE, Blagoev B, Kratchmarova I, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 2002;1:376-86.
165. Li F, Dai L, Niu J. GPX2 silencing relieves epithelial-mesenchymal transition, invasion, and metastasis in pancreatic cancer by downregulating Wnt pathway. J Cell Physiol 2020;235:7780-90.
166. Noyer L, Grolez GP, Prevarskaya N, Gkika D, Lemonnier L. TRPM8 and prostate: a cold case? Pflugers Arch 2018;470:1419-29.

167. Moon YW, Rao G, Kim JJ, et al. LAMC2 enhances the metastatic potential of lung adenocarcinoma. Cell Death Differ 2015;22:1341-52.
168. Li XX, Peng JJ, Liang L, et al. RNA-seq identifies determinants of oxaliplatin sensitivity in colorectal cancer cell lines. Int J Clin Exp Pathol 2014;7:3763-70.
169. Toscano F, Parmentier B, Fajoui ZE, et al. p53 dependent and independent sensitivity to oxaliplatin of colon cancer cells. Biochem Pharmacol 2007;74:392-406.
170. Perez M, Lucena-Cacace A, Marin-Gomez LM, et al. Dasatinib, a Src inhibitor, sensitizes liver metastatic colorectal carcinoma to oxaliplatin in tumors with high levels of phospho-Src. Oncotarget 2016;7:33111-24.
171. Das Thakur M, Salangsang F, Landman AS, et al. Modelling vemurafenib resistance in melanoma reveals a strategy to forestall drug resistance. Nature 2013;494:251-5.
172. Rao DD, Vorhies JS, Senzer N, Nemunaitis J. siRNA vs. shRNA: similarities and differences. Adv Drug Deliv Rev 2009;61:746-59.
173. Danhier F, Feron O, Preat V. To exploit the tumor microenvironment: Passive and active tumor targeting of nanocarriers for anti-cancer drug delivery. J Control Release 2010;148:135-46.
174. Fidler IJ. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. Nat Rev Cancer 2003;3:453-8.
175. Simpson CD, Anyiwe K, Schimmer AD. Anoikis resistance and tumor metastasis. Cancer Lett 2008;272:177-85.
176. Elbekai RH, Paranjpe MG, Contreras PC, Spada A. Carcinogenicity assessment of the pan-caspase inhibitor, emricasan, in Tg.rasH2 mice. Regul Toxicol Pharmacol 2015;72:169-78.
177. Brigelius-Flohe R, Kipp AP. Physiological functions of GPx2 and its role in inflammation-triggered carcinogenesis. Ann N Y Acad Sci 2012;1259:19-25.
178. Yan W, Chen X. GPX2, a direct target of p63, inhibits oxidative stress-induced apoptosis in a p53-dependent manner. J Biol Chem 2006;281:7856-62.
179. Florian S, Krehl S, Loewinger M, et al. Loss of GPx2 increases apoptosis, mitosis, and GPx1 expression in the intestine of mice. Free Radic Biol Med 2010;49:1694-702.
180. Banning A, Florian S, Deubel S, et al. GPx2 counteracts PGE2 production by dampening COX-2 and mPGES-1 expression in human colon cancer cells. Antioxid Redox Signal 2008;10:1491-500.
181. Chu FF, Esworthy RS, Chu PG, et al. Bacteria-induced intestinal cancer in mice with disrupted Gpx1 and Gpx2 genes. Cancer Res 2004;64:962-8.
182. Kolobova E, Tuganova A, Boulatnikov I, Popov KM. Regulation of pyruvate dehydrogenase activity through phosphorylation at multiple sites. Biochem J 2001;358:69-77.
183. Weng Y, Shen Y, He Y, et al. The miR-15b-5p/PDK4 axis regulates osteosarcoma proliferation through modulation of the Warburg effect. Biochem Biophys Res Commun 2018;503:2749-57.
184. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science 2009;324:1029-33.
185. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. Radiother Oncol 2009;92:329-33.
186. Liu Z, Chen X, Wang Y, et al. PDK4 protein promotes tumorigenesis through activation of cAMP-response element-binding protein (CREB)-Ras homolog enriched in brain (RHEB)-mTORC1 signaling cascade. J Biol Chem 2014;289:29739-49.
187. Wallden K, Nyman T, Hallberg BM. SnoN Stabilizes the SMAD3/SMAD4 Protein Complex. Sci Rep 2017;7:46370.
188. Liu C, Zhang H, Zang X, Wang C, Kong Y, Zhang H. The influence of SnoN gene silencing by siRNA on the cell proliferation and apoptosis of human pancreatic cancer cells. Diagn Pathol 2015;10:30.

189. Zhu Q, Le Scolan E, Jahchan N, Ji X, Xu A, Luo K. SnoN Antagonizes the Hippo Kinase Complex to Promote TAZ Signaling during Breast Carcinogenesis. Dev Cell 2016;37:399-412.
190. Bautista DM, Siemens J, Glazer JM, et al. The menthol receptor TRPM8 is the principal detector of environmental cold. Nature 2007;448:204-8.
191. Yang ZH, Wang XH, Wang HP, Hu LQ. Effects of TRPM8 on the proliferation and motility of prostate cancer PC-3 cells. Asian J Androl 2009;11:157-65.
192. Zhu G, Wang X, Yang Z, et al. Effects of TRPM8 on the proliferation and angiogenesis of prostate cancer PC-3 cells in vivo. Oncol Lett 2011;2:1213-7.
193. Nagar S, Blanchard RL. Pharmacogenetics of uridine diphosphoglucuronosyltransferase (UGT) 1A family members and its role in patient response to irinotecan. Drug Metab Rev 2006;38:393-409.
194. Strassburg CP, Oldhafer K, Manns MP, Tukey RH. Differential expression of the UGT1A locus in human liver, biliary, and gastric tissue: identification of UGT1A7 and UGT1A10 transcripts in extrahepatic tissue. Mol Pharmacol 1997;52:212-20.
195. Marcuello E, Altes A, Menoyo A, Del Rio E, Gomez-Pardo M, Baiget M. UGT1A1 gene variations and irinotecan treatment in patients with metastatic colorectal cancer. Br J Cancer 2004;91:678-82.
196. Danielson PB. The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. Curr Drug Metab 2002;3:561-97.
197. Androutsopoulos VP, Tsatsakis AM, Spandidos DA. Cytochrome P450 CYP1A1: wider roles in cancer progression and prevention. BMC Cancer 2009;9:187.
198. Huang D, Du C, Ji D, Xi J, Gu J. Overexpression of LAMC2 predicts poor prognosis in colorectal cancer patients and promotes cancer cell proliferation, migration, and invasion. Tumour Biol 2017;39:1010428317705849.
199. Takahashi S, Hasebe T, Oda T, et al. Cytoplasmic expression of laminin gamma2 chain correlates with postoperative hepatic metastasis and poor prognosis in patients with pancreatic ductal adenocarcinoma. Cancer 2002;94:1894-901.
200. Koshikawa N, Moriyama K, Takamura H, et al. Overexpression of laminin gamma2 chain monomer in invading gastric carcinoma cells. Cancer Res 1999;59:5596-601.
201. Pei YF, Liu J, Cheng J, Wu WD, Liu XQ. Silencing of LAMC2 Reverses Epithelial-Mesenchymal Transition and Inhibits Angiogenesis in Cholangiocarcinoma via Inactivation of the Epidermal Growth Factor Receptor Signaling Pathway. Am J Pathol 2019;189:1637-53.
202. He M, Smith LD, Chang R, Li X, Vockley J. The role of sterol-C4-methyl oxidase in epidermal biology. Biochim Biophys Acta 2014;1841:331-5.
203. Sukhanova A, Gorin A, Serebriiskii IG, et al. Targeting C4-demethylating genes in the cholesterol pathway sensitizes cancer cells to EGF receptor inhibitors via increased EGF receptor degradation. Cancer Discov 2013;3:96-111.
204. Liang Y, Chen X, Wu Y, et al. LncRNA CASC9 promotes esophageal squamous cell carcinoma metastasis through upregulating LAMC2 expression by interacting with the CREB-binding protein. Cell Death Differ 2018;25:1980-95.
205. Manzella C, Singhal M, Alrefai WA, Saksena S, Dudeja PK, Gill RK. Serotonin is an endogenous regulator of intestinal CYP1A1 via AhR. Sci Rep 2018;8:6103.
206. Liu P, Yang H, Zhang J, et al. The lncRNA MALAT1 acts as a competing endogenous RNA to regulate KRAS expression by sponging miR-217 in pancreatic ductal adenocarcinoma. Sci Rep 2017;7:5186.
207. Chang CH, Yen MC, Liao SH, et al. Secreted Protein Acidic and Rich in Cysteine (SPARC) Enhances Cell Proliferation, Migration, and Epithelial Mesenchymal Transition, and SPARC Expression is Associated with Tumor Grade in Head and Neck Cancer. Int J Mol Sci 2017;18.

208. Brasil da Costa FH, Lewis MS, Truong A, Carson DD, Farach-Carson MC. SULF1 suppresses Wnt3A-driven growth of bone metastatic prostate cancer in perlecan-modified 3D cancer-stroma-macrophage triculture models. PLoS One 2020;15:e0230354.
209. Liu M, Chen S, Yueh MF, et al. Cadmium and arsenic override NF-kappaB developmental regulation of the intestinal UGT1A1 gene and control of hyperbilirubinemia. Biochem Pharmacol 2016;110-111:37-46.

Supplementary Data

Complete list of materials and reagents used throughout this study:

DMSO – Sigma-Aldrich, D2438-5X10ML

PBS – Biowest, without Calcium and without Magnesium, L0615-500

Cell culture Flasks -Sigma-Aldrich, Nunc EasYFlask 25cm2 (156367) and 75cm2 (156499)

DMEM – Biowest, with L-Glutamine and Sodium Pyruvate, L0104-500

FBS - Gibco, 10270106

Penicillin/Streptomycin - ABM, G255.

Opti-MEM Reduced Serum Medium – Gibco, 51985034.

TrypLE Express Enzyme (1X) – Gibco, 12605028

Incubator–PHCbi, MCO-170AICUV-PE IncuSafe CO2 Incubator

Cell Culture Centrifuge –Eppendorf, Centrifuge 5702

Non-Cell Culture Centrifuge – Eppendorf, Microcentrifuge 5415R

Cell Counting Slides – Bio-Rad, 1450011

Automated Cell Counter – Bio-Rad, TC20

Membrane Labeling - Sigma-Aldrich, MINI67-1KT and MINCLARET-1KT

Laminar Flow Hood – Heal Force, Hfsafe 1200LC. Where all cell related protocols were performed.

Annexin V Binding Buffer – BioLegend, 422201

Pacific Blue Annexin V – BioLegend, 640918

Propidium Iodide - Sigma-Aldrich, P4170

Cell Culture Microscope Camera – Leica, Leica DFC3000 G

LED Illumination – CoolLed, pE-300Lite

24-Well Plates – Thermo Scientific, Nunclon Delta Surface 142475

Flow Cytometer - BD Biosciences, BD LSRFortessa X-20. Champalimaud's Flow Cytometry Platform.

Cell Strainers – Fisher Scientific, 22363548

Cell Sorter -BD Biosciences, BD FACSAria Fusion Cell Sorter. Champalimaud's Flow Cytometry Platform.

RNA extraction Kit – Qiagen, 74106-RNeasy Mini Kit

Nuclease-Free Water – Invitrogen, AM9932

Spectrophotometer – Thermo Scientific, NanoDrop 2000

Puromycin – Sigma-Aldrich, P8833-10MG

Reverse Transcription Kit –Qiagen, 205313-QuantiTect Reverse Transcription Kit

SYBR green kit -Applied Biosystems, PowerUp SYBR Green Master Mix

Thermal Cycler – Bio-Rad, T100 Thermal Cycler

RT-qPCR equipment - Applied Biosystems, StepOnePlus Real-Time PCR System

RT-qPCR Plates – Applied Biosystems, 4346906

LB Broth – Prepared by Champalimaud's Glass Wash Plataform.

Midiprep kit – Zymo Research, ZymoPURE II Plasmid Midiprep Kit

EdU Kit – Invitrogen, Click-iT Plus EdU Flow Cytometry Assay Kit

Emricasam – Selleckchem, S7775

Oxaliplatin - Sigma-Aldrich, Y0000271

HEPES Buffer – Corning, 25-060-CI

EDTA – Corning, 46-034-CI

FACS Buffer - 2% FBS + 25mM HEPES + 1mM EDTA, diluted in PBS.

Lipofectamine 3000 – Invitrogen, L3000015

Bioinformatics analysis

GPX2

Glutathione peroxidase 2 is involved in the reduction of hydrogen peroxide[177]. In cancer, it seems to be a tumour promoter as its expression inhibits apoptosis[178,179], and promotes metastatic capabilities through the WNT pathway[165]. It was also found to downregulate COX-2, thus hindering inflammation which could stop carcinogenesis[180]. Supporting that protective role, KO mice for GPX-2 are highly susceptible to bacteria-associated inflammation, which can lead to cancer[181]. An important pathway to which it has been implicated is the NRF2 pathway[177]. Expression analysis reveals that *GPX2* has lower expression in breast invasive carcinoma, kidney clear cell carcinoma and prostate adenocarcinoma, and higher expression in colon adenocarcinoma, lung adenocarcinoma and pancreatic adenocarcinoma. (GDC TCGA data – Figure S1). Its overexpression is related to a better DFS in prostate adenocarcinoma (Gepia analysis – Figure S2).

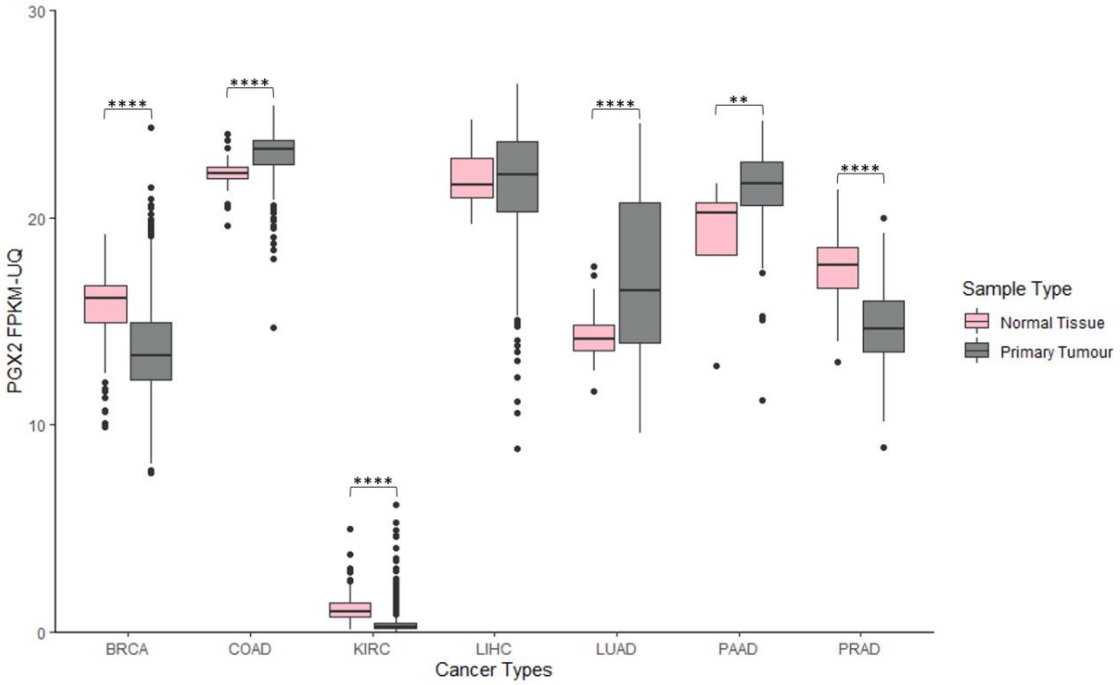


Figure S1. Expression analysis of *GPX2* on several GDC TCGA datasets. *GPX2* expression is lower in breast invasive carcinoma, kidney clear cell carcinoma and prostate adenocarcinoma, and higher expression in colon adenocarcinoma, lung adenocarcinoma and pancreatic adenocarcinoma. Dots represent outliers. Non-significance not shown. **P<0,01, ****P<0,0001 when compared to normal tissue expression.

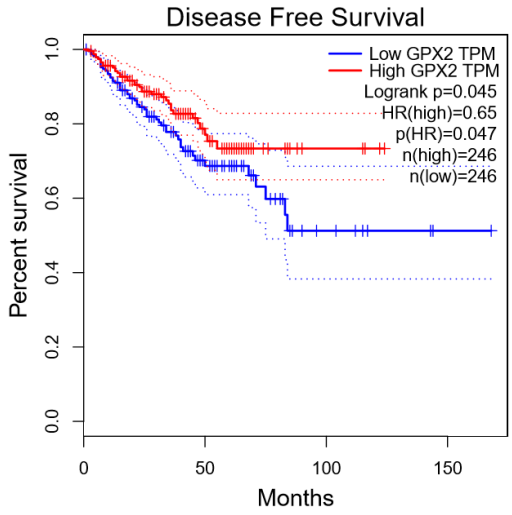


Figure S2. Disease free survival analysis of prostate adenocarcinoma patients stratified by *GPX2* expression levels. *GPX2* Overexpression is linked to better DFS in these patients.

PDK4

Pyruvate dehydrogenase kinase 4 inhibits the pyruvate dehydrogenase complex, thus preventing the progression into aerobic respiration[182]. Its overexpression leads cells to the Warburg Effect[183], where cells maintain a high rate of lactic acid fermentation

instead of aerobic respiration, which is commonly found in cancer[184,185]. An important pathway to which it relates is mTORC1, through which it increases tumourogenesis[186]. Expression analysis reveals that *PDK4* has lower expression in invasive breast carcinoma, colon adenocarcinoma, hepatocellular carcinoma, lung adenocarcinoma and prostate adenocarcinoma (GDC TCGA data – Figure S3). Low expression leads to worse OS in kidney clear cell carcinoma and worse DFS in liver hepatocellular carcinoma, pancreatic adenocarcinoma and prostate adenocarcinoma (Gepia Analysis – Figure S4).

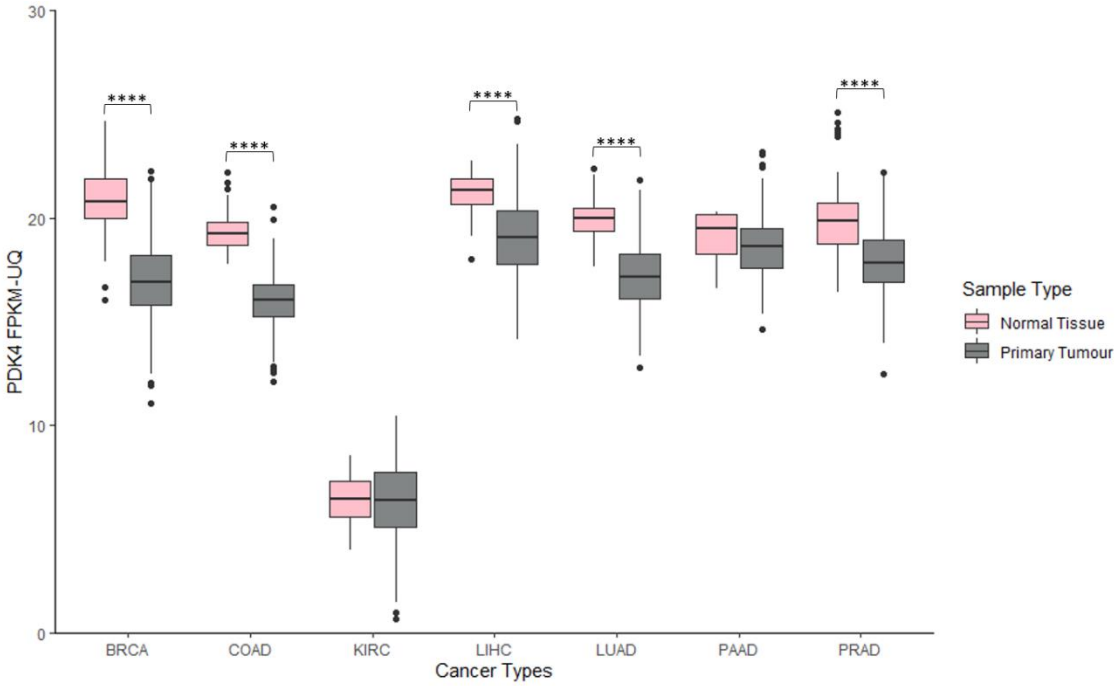
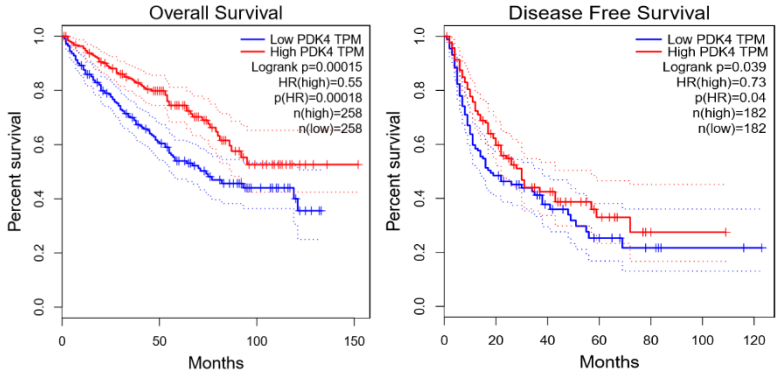


Figure S3. Expression analysis of *PDK4* on several GDC TCGA datasets. *PDK4* expression is lower in invasive breast carcinoma, colon adenocarcinoma, hepatocellular carcinoma, lung adenocarcinoma and prostate adenocarcinoma. Dots represent outliers. Non-significance not shown. ****P<0,0001 when compared to normal tissue expression.



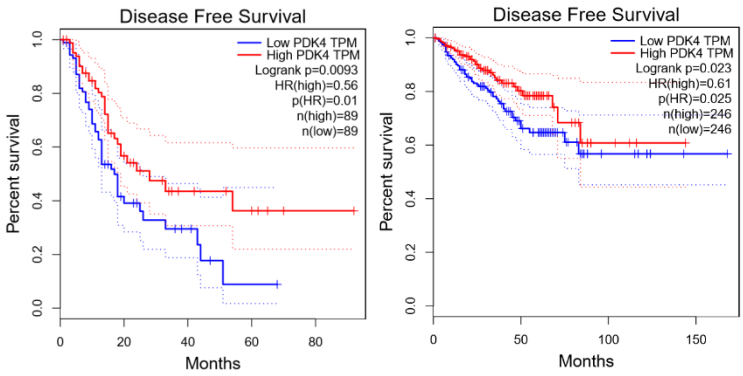


Figure S4. *PDK4* stratified OS and DFS analysis (A) OS analysis of kidney clear cell carcinoma patients stratified by *PDK4* expression levels. (B, C and D) DFS analysis of hepatocellular carcinoma, pancreatic adenocarcinoma, and prostate adenocarcinoma patients stratified by *PDK4* expression levels, respectively.

SKIL

Ski-like protein plays a role in regulating the SMAD3/SMAD4 complex, which are key transducers in the TGF-β pathway[187]. In cancer, its silencing leads to lower proliferation and higher apoptosis of pancreatic cancer cells[188]. A noteworthy pathway in which it is implicated is the Hippo-Pathway, where it promotes WWTR1 signaling in breast cancer[189]. Expression analysis reveals that *SKIL* has higher expression in invasive breast carcinoma, colon adenocarcinoma, and lower expression in hepatocellular carcinoma and prostate adenocarcinoma (GDC TCGA data – Figure S5). Its overexpression in pancreatic adenocarcinoma is linked to a worse DFS (Gepia Analysis – Figure S6).

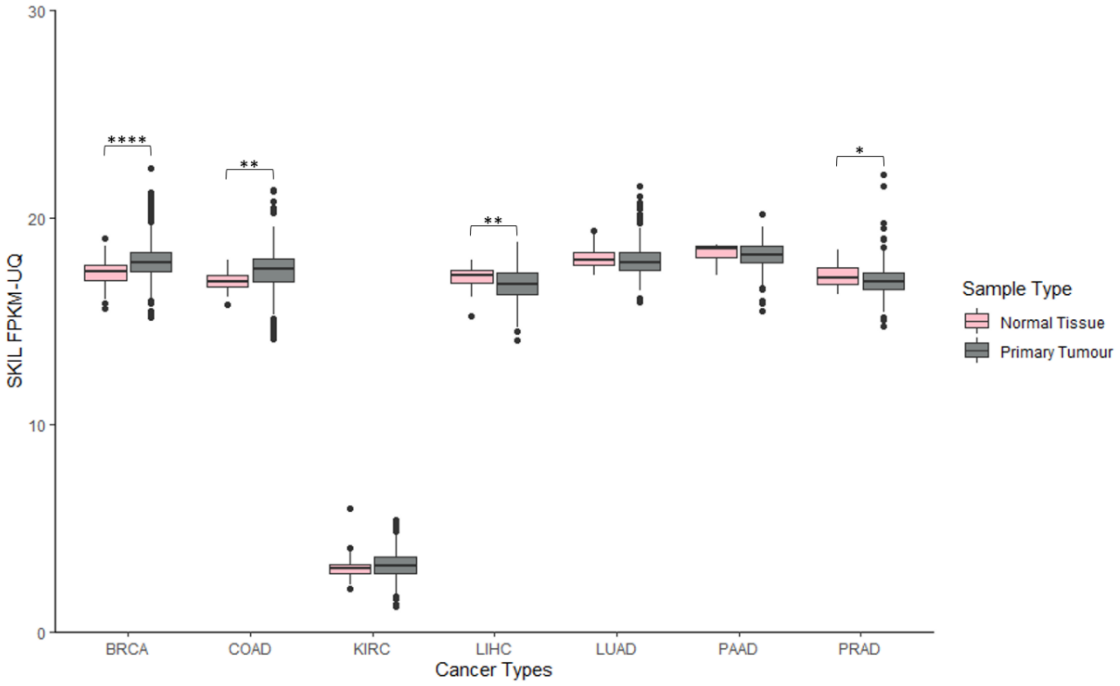


Figure S5. Expression analysis of *SKIL* on several GDC TCGA datasets. *SKIL* expression is higher in invasive breast carcinoma and colon adenocarcinoma, and lower in hepatocellular carcinoma and prostate adenocarcinoma. Dots represent outliers. Non-significance not shown. *P<0,05, **P<0,01, ****P<0,0001 when compared to normal tissue expression.

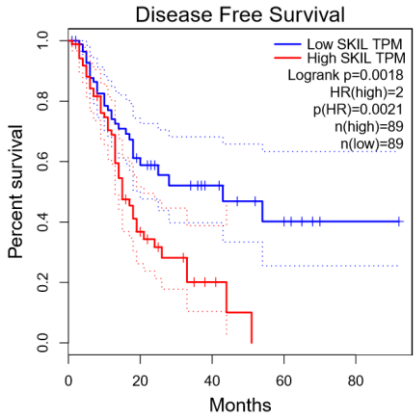


Figure S6. Disease free survival of pancreatic adenocarcinoma patients stratified by *SKIL* expression levels. *SKIL* overexpression is linked to lower DFS in these patients.

TRPM8

Transient receptor potential cation channel subfamily melastatin member 8 is the principal ion channel responsible in cold detection[190]. In cancer, it is upregulated in prostate and pancreatic cancer, but substantially downregulated in metastatic prostate cancer[166]. When overexpressed, it has been found to inhibit cancer cell migration, and tumour angiogenesis[191,192]. Expression analysis reveals that *TRPM8* has higher

expression in colon adenocarcinoma, lung adenocarcinoma and prostate adenocarcinoma, and lower expression in hepatocellular carcinoma (GDC TCGA data – Figure S7). When it has low expression levels in breast cancer, it links to a worse DFS (Gepia Analysis – Figure S8).

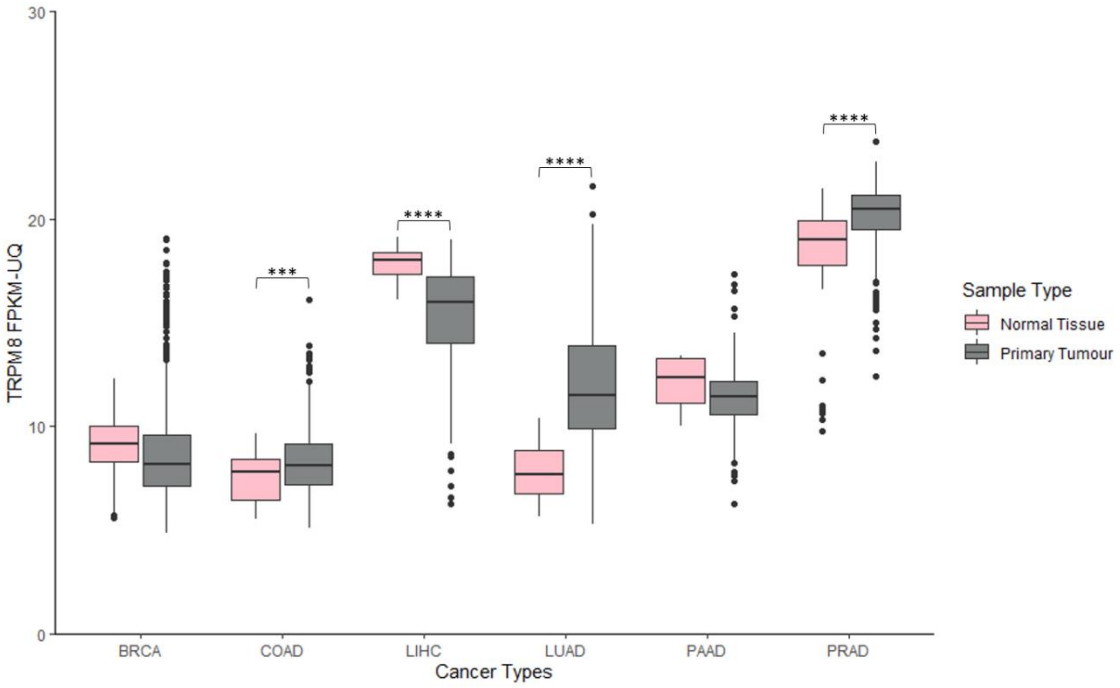


Figure S7. Expression analysis of *TRPM8* on several GDC TCGA datasets. *TRPM8* expression is higher in colon adenocarcinoma, lung adenocarcinoma and prostate adenocarcinoma, and lower expression in hepatocellular carcinoma. KIRC database was not used because *TRPM8* expression levels were very low in both tumour and normal tissue. Dots represent outliers. Non-significance not shown. ***P<0,001, ****P<0,0001 when compared to normal tissue expression.

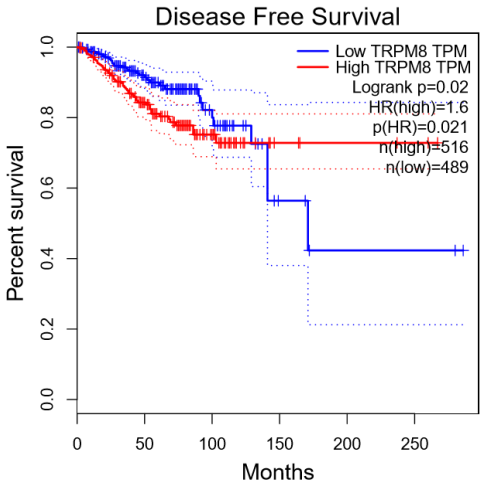


Figure S8. Disease free survival analysis of breast cancer patients stratified by *TRPM8* expression levels. *TRPM8* overexpression is linked to better DFS in these patients.

UGT1A1 and *UGT1A10*

Uridine diphosphate family 1 member A1 and Uridine diphosphate family 1 member A10 are enzymes of the glucuronidation pathway that transform small lipophilic molecules, such as steroids, bilirubin, hormones, and drugs, into water-soluble, excretable metabolites[193,194]. In cancer, deficiency in UGT1A1, leads to higher toxicity amidst treatment with Irinotecan in colon cancer patients[195]. Expression analysis reveals that *UGT1A1* has higher expression in lung adenocarcinoma, and lower expression in colon adenocarcinoma and hepatocellular carcinoma (GDC TCGA data – Figure S9). *UGT1A1* overexpression is linked to worse OS in pancreatic adenocarcinoma and in prostate adenocarcinoma (Gepia analysis – Figure S10). Expression analysis reveals that *UGT1A10* has lower expression in colon adenocarcinoma and hepatocellular carcinoma (GDC TCGA data – Figure S11). *UGT1A10* overexpression in kidney clear cell carcinoma and pancreatic adenocarcinoma is linked to worse OS (Gepia analysis – Figure S12).

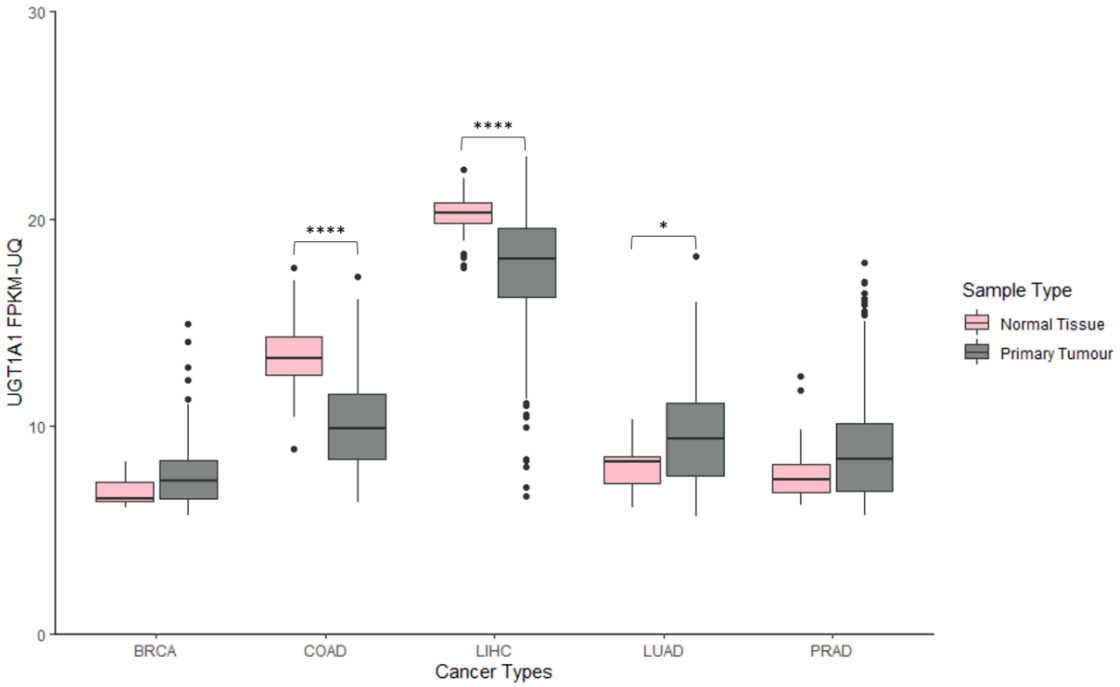


Figure S9. Expression analysis of *UGT1A1* on several GDC TCGA datasets. *UGT1A1* expression is higher in lung adenocarcinoma, and lower in colon adenocarcinoma and hepatocellular carcinoma. KIRC and PAAD database were not used because there was only one "normal tissue"

sample. Dots represent outliers. Non-significance not shown. *P<0,05, ****P<0,0001 when compared to normal tissue expression.

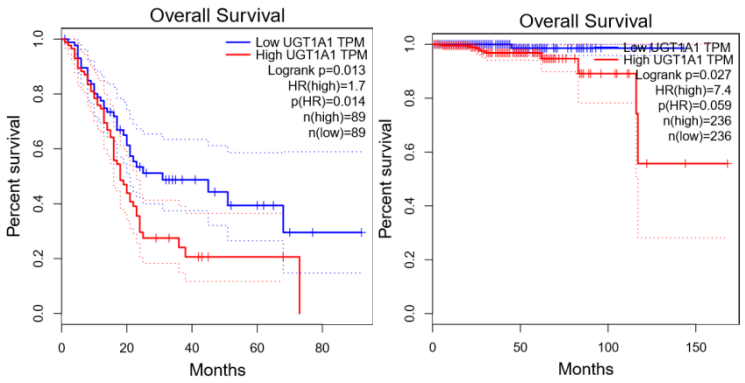


Figure S10. *UGT1A1* stratified OS analysis. OS analysis of pancreatic adenocarcinoma (A) and prostate adenocarcinoma (B) patients stratified by *UGT1A1* expression levels.

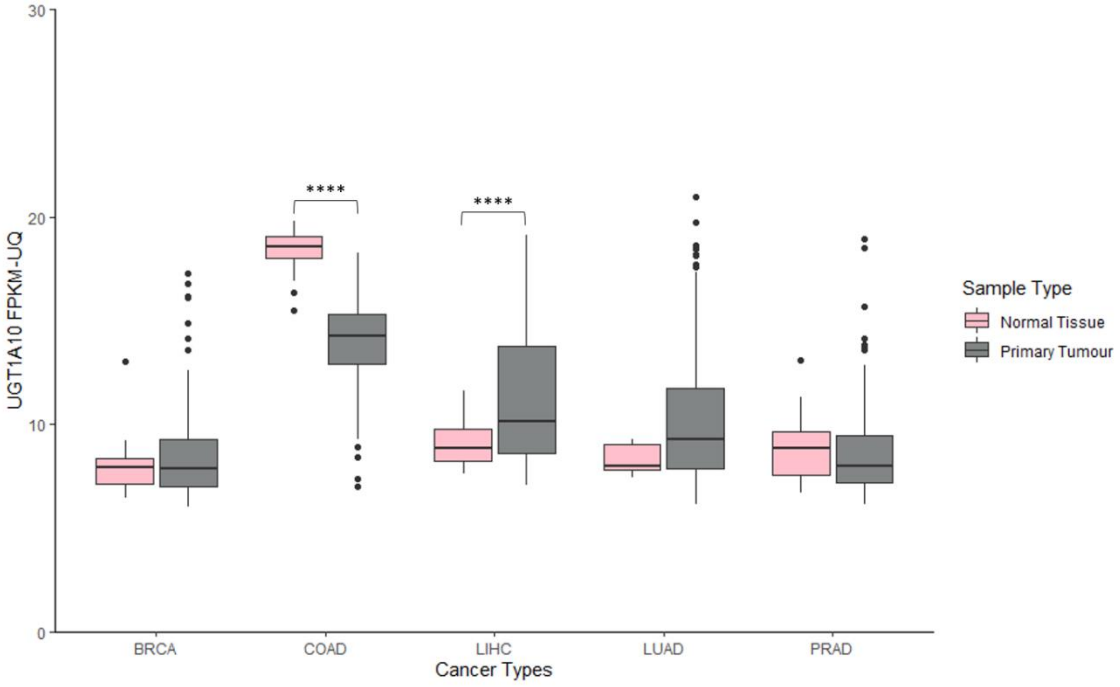


Figure S11. Expression analysis of *UGT1A10* on several GDC TCGA datasets. *UGT1A10* expression is lower in colon adenocarcinoma and hepatocellular carcinoma. KIRC and PAAD database were not used because there was only one "normal tissue" sample. Dots represent outliers. Non-significance not shown. ****P<0,0001 when compared to normal tissue expression.

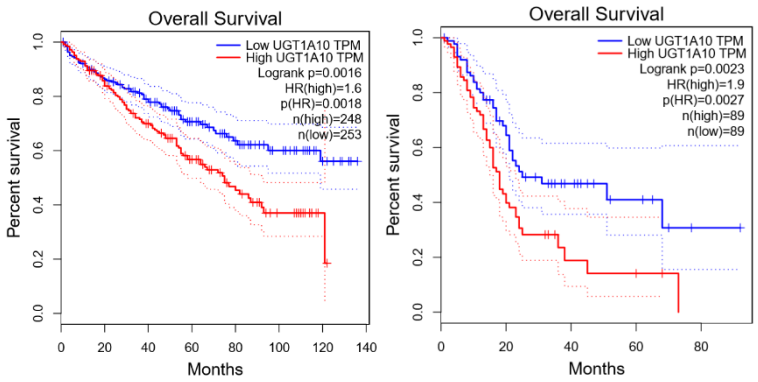


Figure S12. *UGT1A10* stratified OS analysis. OS analysis of kidney clear cell carcinoma (A) and prostate adenocarcinoma (B) patients stratified by *UGT1A10* expression levels.

CYP1A1

Cytochrome P450 family 1 subfamily A member 1 is part of a complex which catalyses many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids[196]. Its role in cancer is still rather controversial, and CYP1A1 can work both in tumour progression and prevention depending on the context[197]. Expression analysis reveals that *CYP1A1* has lower expression in breast invasive carcinoma, kidney clear cell carcinoma, hepatocellular carcinoma and lung adenocarcinoma (GDC TCGA data – Figure S13). High expression correlates with better DFS in lung adenocarcinoma (Gepia Analysis – Figure S14).

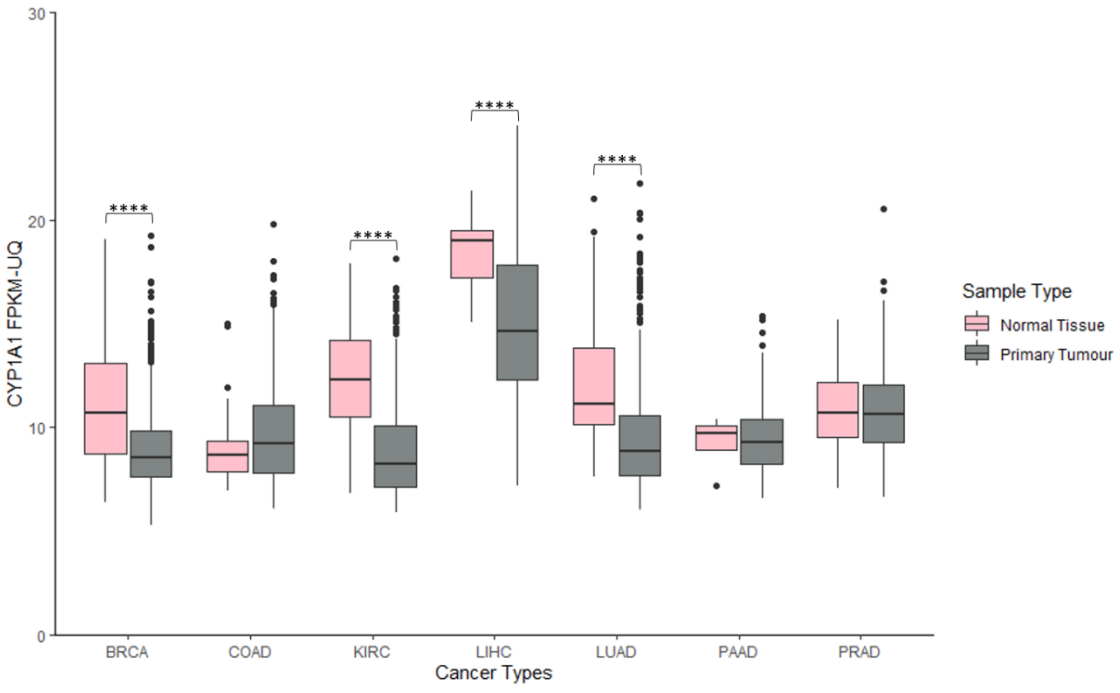


Figure S13. Expression analysis of *CYP1A1* on several GDC TCGA datasets. *CYP1A1* expression is lower in breast invasive carcinoma, kidney clear cell carcinoma, hepatocellular carcinoma and lung adenocarcinoma. Dots represent outliers. Non-significance not shown. ****P<0,0001 when compared to normal tissue expression.

LAMC2

Laminin subunit gamma 2 is an essential component of epithelial basement membranes and regulates cell motility and adhesion[198]. It is overexpressed in a wide array of cancers, including gastric cancer and pancreatic adenocarcinoma[198-200]. Silencing of the gene was found to inhibit epithelial-mesenchymal transition through silencing of the EGFR pathway[201], and its upregulation enhances metastatic capabilities in lung adenocarcinoma[167]. Expression analysis reveals that *LAMC2* has lower expression in breast invasive carcinoma, kidney clear cell carcinoma, hepatocellular carcinoma and higher expression in colon adenocarcinoma (GDC TCGA data – Figure S15). Its overexpression is linked to poorer DFS in colon adenocarcinoma and pancreatic adenocarcinoma, and to worse OS in lung adenocarcinoma (Gepia Analysis – Figure S16).

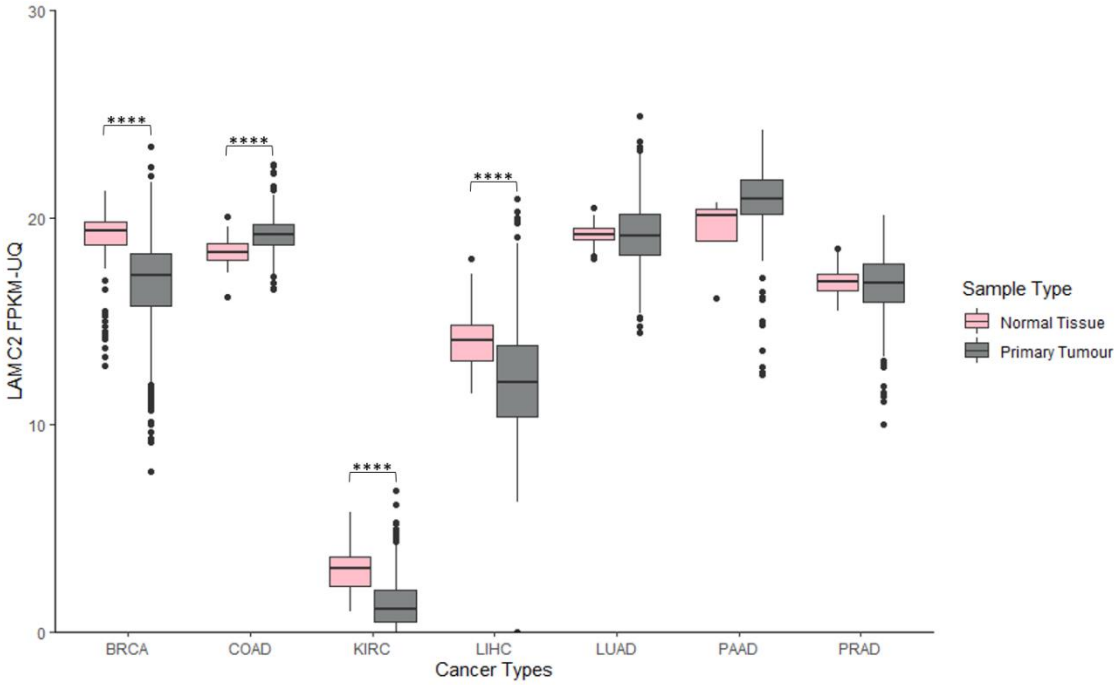


Figure S14. Expression analysis of *LAMC2* on several GDC TCGA datasets. *LAMC2* expression is lower in breast invasive carcinoma, kidney clear cell carcinoma, hepatocellular carcinoma and higher in colon adenocarcinoma. Dots represent outliers. Non-significance not shown. ****P<0,0001 when compared to normal tissue expression.

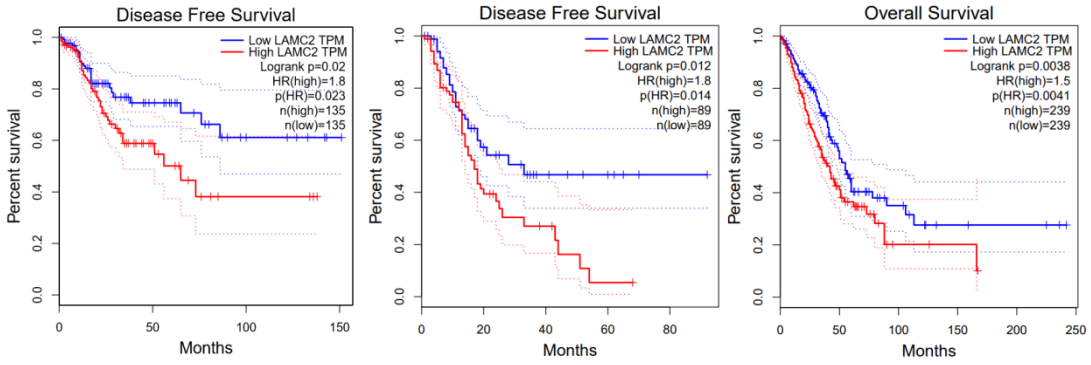


Figure S15. *LAMC2* stratified DFS and OS analysis. DFS analysis of colon adenocarcinoma (A) and pancreatic adenocarcinoma (B) patients stratified by *LAMC2* expression levels. (C) OS analysis of lung adenocarcinoma patients stratified by *LAMC2* expression levels.

MSMO1

Methylsterol monooxygenase 1 is an enzyme important in cholesterol biosynthesis[202]. Deficiency in this gene has been found to inhibit EGFR signaling[203]. Expression analysis reveals that *LAMC2* has lower expression in kidney clear cell carcinoma, hepatocellular carcinoma, lung adenocarcinoma and prostate adenocarcinoma, and higher expression in colon adenocarcinoma (GDC TCGA data – Figure S17). Its overexpression is linked to poorer DFS in colon adenocarcinoma, and to worse OS in kidney clear cell carcinoma, and to worse OS and DFS in pancreatic adenocarcinoma (Gepia Analysis – Figure S18).

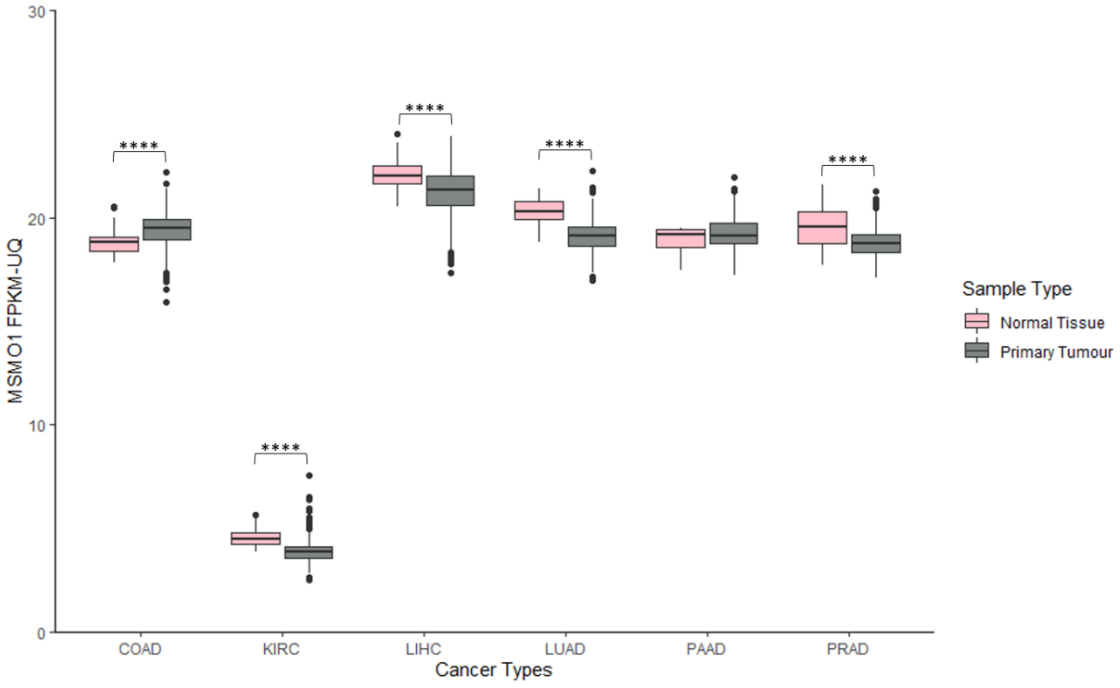


Figure S16. Expression analysis of *MSMO1* on several GDC TCGA datasets. *MSMO1* expression is lower in kidney clear cell carcinoma, hepatocellular carcinoma, lung adenocarcinoma and

prostate adenocarcinoma, and higher in colon adenocarcinoma. No breast invasive carcinoma data is available for *MSMO1* expression. Dots represent outliers. Non-significance not shown. ****P<0,0001 when compared to normal tissue expression.

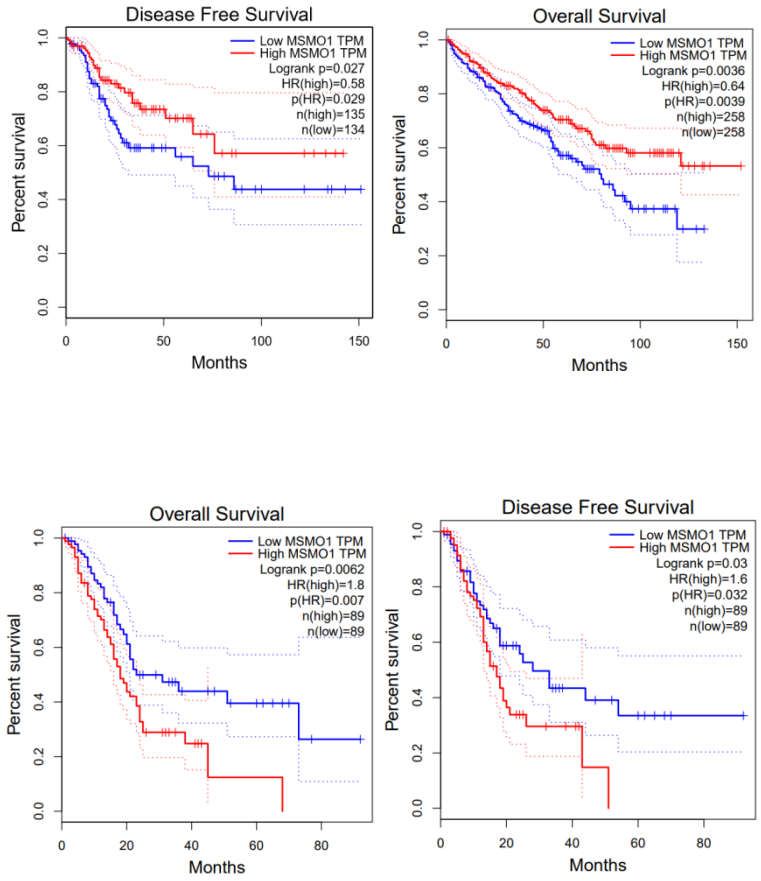


Figure S17. *MSMO1* stratified OS and DFS analysis. (A) DFS analysis of colon adenocarcinoma patients stratified by *MSMO1* expression levels. (B) OS analysis of kidney clear cell carcinoma stratified by *MSMO1* expression levels. (C) OS analysis of pancreatic adenocarcinoma patients stratified by *MSMO1* expression levels. (D) DFS analysis of pancreatic adenocarcinoma patients stratified by *MSMO1* expression levels.

Supplementary Tables
RT-qPCR primers and shRNA sequences

Table S1.**Full list of primers used in this study, and their respective origin.**

Primer Name	*Sequence 5' 3'*	*Source*
AKT1-Fw	TGGACTACCTGCACTCGGAGAA	Designed by the author
AKT1-Rv	GTGCCGCAAAAGGTCTTCATGG	Designed by the author
AQP3-Fw	CCGTGACCTTTGCCATGTG	Taken from REF[204]
AQP3-Rv	CGAAGTGCCAGATTGCATCATAA	Taken from REF[204]
CDH1-Fw	AGAACGCATTGCCACATACACTCTC	Designed by the author
CDH1-Rv	CGGTTACCGTGATCAAAATCTCCA	Designed by the author
COL6A3-Fw	ACAGAGTGGAGGAGTCTGGTTT	Designed by the author
COL6A3-Rv	CAGGAGAACTGAACGGCCT	Designed by the author
CYP1A1-Fw	TCGGCCACGGAGTTTCTTC	Taken from REF[205]
CYP1A1-Rv	GGTCAGCATGTGCCCAATCA	Taken from REF[205]
DCN-Fw	CTGTTGTTCGATTATGGGTGCT	Designed by the author
DCN-Rv	CCAGGAGGGGTCAGAGACAA	Designed by the author
EGFR-Fw	CCAGTATTGATCGGGAGAGCC	Designed by the author
EGFR-Rv	GTGCCTTGGCAAACTTTCTT	Designed by the author
ERBB2-Fw	GGGGAGGAGCTTAGGAGTATGAA	Designed by the author

ERBB2-Rv	TTCCCTTAACCCTATCGCCT	Designed by the author
GAPDH-Fw	GCACCGTCAAGGCTGAGAAC	Taken from REF[206]
GAPDH-Rv	GCCTTCTCCATGGTGGTGAA	Taken from REF[206]
GPX2-Fw	AACCAATTTGGACATCAGGAGA	Designed by the author
GPX2-Rv	GATGCTCGTTCTGCCCATTC	Designed by the author
LAMC2-Fw	GCCTTTTGGCACCTGTATTC	Taken from REF[204]
LAMC2-Rv	CAGGATTCTCATCCCCTGAA	Taken from REF[204]
MAPK1-Fw	ACACCAACCTCTCGTACATCGG	Designed by the author
MAPK1-Rv	TGGCAGTAGGTCTGGTGCTCAA	Designed by the author
MGP-Fw	TTTGTGTTATGAATCACATGAAAGC	Designed by the author
MGP-Rv	AGCGTTCTCGGATCCTCTCT	Designed by the author
MSMO1-Fw	TCTCGCGGCCGTTCAGAATTA	Designed by the author
MSMO1-Rv	ACAGCCAAGGATGCTGAACTAA	Designed by the author
MTOR-Fw	AAGCCGCGCGAACCTC	Designed by the author
MTOR-Rv	CTGGTTTCCTCATTCCGGCT	Designed by the author
MUC4-Fw	GGACAAGAACCAGCAGAGGG	Designed by the author
MUC4-Rv	GGGTGGCGTTGCTGATG	Designed by the author

MYT1-Fw	TCCAGGCACCGAAGTTTACA	Designed by the author
MYT1-Rv	GGGGGTGTGACTTCCTCTTG	Designed by the author
NFKB1-Fw	CGCCGCTTAGGAGGGAGA	Designed by the author
NFKB1-Rv	CCATTCTGAAGCCGGGTGG	Designed by the author
NRIP1-Fw	GGGGAAGTGTTTGGATTGTGAG	Designed by the author
NRIP1-Rv	TCAAGTGTGCATCTTCTGGC	Designed by the author
PDK4-Fw	CATACTCCACTGCACCAACG	Designed by the author
PDK4-Rv	CCGTAACCAAAACCAGCCAAAG	Designed by the author
PTGS2-Fw	CTCCCTTGGGTGTCAAAGGTAAA	Designed by the author
PTGS2-Rv	AACTGATGCGTGAAGTGCTG	Designed by the author
SKIL-Fw	ATATGCAGGACAGTTGGCAGA	Designed by the author
SKIL-Rv	GTCTTGCTTCCCGTTCCTGT	Designed by the author
SLC34A2-Fw	AAGCGGCCTCCAGCTA	Designed by the author
SLC34A2-Rv	GGTACTTATCGGGGTTGGGC	Designed by the author
SLC6A14-Fw	GGGGAGGTGCCTTCTTCATC	Designed by the author

SLC6A14-Rv	GCATAGCCAATGCCCTCAAA	Designed by the author
SPARC-Fw	GTGCAGAGGAAACCGAAGAG	Taken from REF[207]
SPARC-Rv	AGTGGCAGGAAGAGTCGAAG	Taken from REF[207]
SPINK4-Fw	GGGATGCTTCGCACACAGAC	Designed by the author
SPINK4-Rv	CCAGGTTGGACATCTGGGAAC	Designed by the author
STAT3-Fw	GATTGGCTGAAGGGGCTGTAA	Designed by the author
STAT3-Rv	GGATCCGGTTGGGGCTTG	Designed by the author
SULF1-Fw	AGACCTAAGTCTTGATGTTGGAA	Taken from REF[208]
SULF1-Rv	CCATCCCATAACTGTCCTCTG	Taken from REF[208]
SYNPR-Fw	GCTCCCGAAGGGGCT	Designed by the author
SYNPR-Rv	ATGCAAAGATTGCAAAAAGCAATTC	Designed by the author
TBP-Fw	GAGCTGTGATGTGAAGTTTCC	Designed by the author
TBP-Rv	TCTGGGTTTGATCATTCTGTAG	Designed by the author
THBS2-Fw	GACGGCATCCAGTACAGAGGG	Designed by the author
THBS2-Rv	GCTTTATATGCCGCTTCTCCTG	Designed by the author
TRPM8-Fw	AATGACACTCTGGACAGCACC	Designed by the author
TRPM8-Rv	TTGCACACATTCTCCGTGGC	Designed by the author

UGT1A10-Fw	CCTCTTTCCTATGTCCCCAATGA	Designed by the author
UGT1A10-Rv	GTTCGCAAGATTCGATGGTCG	Designed by the author
UGT1A1-Fw	CCATCATGCCCAATATGGTT	Taken from REF[209]
UGT1A1-Rv	CCACAATTCCATGTTCTCCA	Taken from REF[209]
VIM-Fw	TGCAGGAGGAGATGCTTCAG	Designed by the author
VIM-Rv	ATTCCACTTTGCGTTCAAGG	Designed by the author

Table S2.**Full list of shRNAs used in this study.**

shRNA Name	*Sequence 5' – 3'*	*MISSION Reference ID*
shRNA_AQP3	CCGGCGGTGGTTTCCTCACCATCAACTCGAGTTGATGGTGAGGAAACCACCGTTTTTG	TRCN0000059590
shRNA_CYP1A1	CCGGCAAATGCAGCTGCGCTCTTAGCTCGAGCTAAGAGCGCAGCTGCATTTGTTTTTTG	TRCN0000431185
shRNA_GPX2	CCGGCCGATCCCAAGCTCATCATTTCTCGAGAAATGATGAGCTTGGGATCGGTTTTTG	TRCN0000046239
shRNA_LAMC2	CCGGGCCCTGCAATTGTAACTCCAACTCGAGTTGGAGTTACAATTGCAGGGCTTTTTG	TRCN0000083391
shRNA_MSMO1	CCGGGCATAGACTCTTACACCACAACTCGAGTTGTGGTGTAAGAGTCTATGCTTTTTG	TRCN0000046245
shRNA_MYT1	CCGGTCGGTCACATCAGCGGGAAATCTCGAGATTTCCCGCTGATGTGACCGATTTTTTG	TRCN0000424938
shRNA_NRIP1	CCGGGCTGCAAGATTACAGGCTGTTCTCGAGAACAGCCTGTAATCTTGCAGCTTTTT	TRCN0000019782
shRNA_PDK4	CCGGCCGCCTCTTTAGTTATACATACTCGAGTATGTATAACTAAAGAGGCGGTTTTT	TRCN0000006265

shRNA_SKIL	CCGGCATTGGCACAATTCCATTTAACTCGAGTTAA ATGGAATTGTGCCAATGTTTTTTG	TRCN00004 24201
shRNA_UGT1A1	CCGGCTTTCTGTGCGACGTGGTTTACTCGAGTAAA CCACGTCGCACAGAAAGTTTTTTG	TRCN00004 14335
shRNA_UGT1A10	CCGGGCACAGGCACAAAGTATATTTCTCGAGAAAT ATACTTTGTGCCTGTGCTTTTTG	TRCN00000 36404

RT-qPCR analysis data

Table S3. Validation of gene hits through RT-qPCR. RNA used from a co-culture assay using the dying protocol. RNA used was the same one that was sent to RNA-Sequencing. N=3.

LS 174T MC	*Gene average C_T ±SD*	*Δ C_TGene-GAPDH ±SD*	*Fold Expression Change*
AQP3	27,515±0,065	9,508±0,093	6,062
GPX2	20,909±0,043	2,903±0,075	1,888
MYT1	29,111±0,086	11,103±0,121	3,230
NRIP1	27,359±0,017	9,352±0,036	2,060
PDK4	29,975±0,086	11,968±0,074	5,777
SKIL	24,647±0,027	6,639±0,059	2,012
TRPM8	32,929±0,626	15,359±0,643	3,522
UGT1A1	21,840±0,019	3,833±0,055	1,622
UGT1A10	25,363±0,012	7,355±0,041	2,166
SLC6A14	34,592±0,885	16,584±0,910	0,376
GAPDH	18,007±0,041	----------	----------
LS 174T CC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
AQP3	24,852±0,088	6,908± 0,045	6,062
GPX2	19,930±0,014	1,985± 0,053	1,888
MYT1	27,356±0,059	9,412±0,102	3,230
NRIP1	26,253±0,043	8,309± 0,026	2,060
PDK4	27,381±0,052	9,437± 0,094	5,777
SKIL	23,575±0,031	5,631± 0,015	2,012

TRPM8	31,049±0,221	13,251± 0,242	3,522
UGT1A1	21,079±0,008	3,135± 0,051	1,622
UGT1A10	24,184±0,051	6,240± 0,019	2,166
SLC6A14	35,938±0,118	17,994± 0,063	0,376
GAPDH	17,944±0,045	----------	----------
LoVo MC	***Gene average CT ±SD***	***Δ CT Gene-GAPDH ±SD***	***Fold Expression Change***
CYP1A1	28,359±0,074	9,056±0,077	0,358
DCN	38,072±0,755	18,754±1,124	8,457
LAMC2	26,787±0,063	7,436±0,087	0,622
MSMO1	24,419±0,178	5,194±0,225	0,590
GAPDH	19,351±0,057	----------	----------
LoVo CC	***Gene average CT ±SD***	***Δ CT Gene-GAPDH ±SD***	***Fold Expression Change***
CYP1A1	29,490±0,068	10,488±0,053	0,358
DCN	34,642±0,293	15,640±0,335	8,457
LAMC2	27,121±0,081	8,173±0,065	0,622
MSMO1	24,831±0,044	5,829±0,077	0,590
GAPDH	19,002±0,040	----------	----------

Table S4. Validation of gene hits through RT-qPCR part 2. RNA used from a co-culture assay using the lentiviral protocol. N=3.

LS 174T MC	*Gene average C_T ±SD*	*Δ C_TGene-GAPDH ±SD*	*Fold Expression Change*
AQP3	27,082±0,027	10,350±0,059	5,929
GPX2	19,341±0,069	2,609±0,104	2,079
MYT1	28,418±0,074	11,686±0,104	3,080
NRIP1	26,367±0,037	9,635±0,041	2,007
PDK4	28,355±0,139	11,719±0,141	5,476
SKIL	24,314±0,034	7,582±0,064	2,135
TRPM8	34,276±0,431	17,832±0,448	6,943
UGT1A1	22,105±0,047	5,373±0,027	1,787

UGT1A10	23,638±0,038	6,906±0,058	2,239
SLC6A14	30,662±0,542	13,930±0,569	0,668
GAPDH	16,731±0,035	----------	----------
LS 174T CC	***Gene average CT ±SD***	***Δ CT Gene-GAPDH ±SD***	***Fold Expression Change***
AQP3	24,725±0,106	7,782±0,113	5,929
GPX2	18,498±0,023	1,553±0,032	2,079
MYT1	27,008±0,036	10,063±0,013	3,080
NRIP1	25,575±0,055	8,630±0,021	2,007
PDK4	26,115±0,049	9,170±0,086	5,476
SKIL	23,433±0,015	6,488±0,061	2,135
TRPM8	31,693±0,048	14,749±0,017	6,943
UGT1A1	21,481±0,025	4,536±0,028	1,787
UGT1A10	22,689±0,043	5,744±0,026	2,239
SLC6A14	31,457±0,286	14,512±0,307	0,668
GAPDH	16,945±0,049	----------	----------
LoVo MC	***Gene average CT ±SD***	***Δ CT Gene-GAPDH ±SD***	***Fold Expression Change***
CYP1A1	27,573±0,217	9,144±0,488	0,605
DCN	----------	----------	----------
LAMC2	25,649±0,032	7,220±0,381	0,640
MSMO1	23,387±0,032	4,958±0,373	0,663
GAPDH	18,145±0,403	----------	----------
LoVo CC	***Gene average CT ±SD***	***Δ CT Gene-GAPDH ±SD***	***Fold Expression Change***
CYP1A1	30,795±0,198	10,152±0,246	0,605
DCN	----------	----------	----------
LAMC2	28,790±0,041	8,147±0,014	0,640
MSMO1	26,476±0,035	5,833±0,074	0,664
GAPDH	20,643±0,049	----------	----------

Table S5. Validation of gene hits through RT-qPCR 72 hours. RNA used from a 72 hours co-culture assay using the lentiviral protocol. N=3.

LS 174T MC	*Gene average C_T ±SD*	*Δ C_TGene-GAPDH ±SD*	*Fold Expression Change*
AQP3	30,197±0,191	11,615±0,181	24,572
GPX2	21,781±0,009	3,198±0,185	2,489
MYT1	30,323±0,154	11,741±0,153	4,628
NRIP1	29,851±0,092	11,269±0,094	2,799
PDK4	31,361±0,096	12,778±0,238	5,555
SKIL	27,671±0,144	9,088±0,075	2,139
UGT1A1	23,043±0,085	2,625±0,565	1,338
UGT1A10	28,28±0,14	7,82±0,63	3,647
GAPDH	18,582±0,175	----------	----------
LS 174T CC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
AQP3	25,109±0,142	6,996±0,016	24,572
GPX2	19,996±0,089	1,883±0,188	2,489
MYT1	27,644±0,069	9,531±0,162	4,628
NRIP1	27,897±0,054	9,784±0,173	2,799
PDK4	28,418±0,253	10,305±0,253	5,555
SKIL	26,105±0,185	7,992±0,091	2,139
UGT1A1	21,42±0,037	2,163±0,094	1,338
UGT1A10	26,47±0,187	5,953±0,159	3,647
GAPDH	18,114±0,150	----------	----------
LoVo MC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
CYP1A1	30,677±0,091	9,53±0,071	0,734
LAMC2	27,653±0,068	6,507±0,076	1,519
MSMO1	21,397±0,034	0,25±0,100	0,514
GAPDH	21,147±0,070	----------	----------
LoVo CC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*

CYP1A1	31,123±0,241	10,213±0,259	0,734
LAMC2	27,05±0,070	6,14±0,102	1,519
MSMO1	22,357±0,048	1,447±0,079	0,514
GAPDH	20,91±0,033	----------	----------

Table S6. Validation of gene hits through RT-qPCR HCT vs LS. RNA used from a co-culture assay using the lentiviral protocol between HCT 116 and LS 174T. N=3.

HCT 116 MC	Gene average C_T ±SD	Δ C_TGene-GAPDH ±SD	Fold Expression Change
AQP3	28,723±0,097	10,754±0,011	1,587
GPX2	20,459±0,108	2,491±0,047	1,055
MYT1	28,297±0,048	10,329±0,148	0,473
NRIP1	27,330±0,076	9,362±0,126	0,410
PDK4	32,326±0,078	14,358±0,035	2,242
SKIL	26,910±0,084	8,942±0,033	0,941
UGT1A1	21,993±0,034	2,587±0,111	1,411
UGT1A10	26,47±0,155	7,063±0,065	1,689
GAPDH	17,968±0,107	----------	----------
HCT 116 CC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
AQP3	27,534±0,156	10,088±0,116	1,587
GPX2	19,860±0,003	2,414±0,044	1,055
MYT1	28,854±0,159	11,408±0,179	0,473
NRIP1	28,094±0,088	10,648±0,110	0,410
PDK4	30,638±0,073	13,193±0,099	2,242
SKIL	26,475±0,044	9,029±0,082	0,941
UGT1A1	21,407±0,069	2,09±0,037	1,411
UGT1A10	25,623±0,0411	6,307±0,025	1,689
GAPDH	17,445±0,042	----------	----------
HCT 116 MC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
CYP1A1	32,03±0,236	12,843±0,302	1,146

LAMC2	27,97±0,017	8,777±0,099	1,512
MSMO1	21,03±0,106	1,843±0,186	0,810
GAPDH	19,19±0,087	----------	----------
HCT 116 CC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
CYP1A1	31,83±0,012	13,117±0,104	1,146
LAMC2	27,37±0,016	8,653±0,094	1,512
MSMO1	21,33±0,217	2,617±0,124	0,810
GAPDH	18,72±0,096	----------	----------

Table S7. Relative expression analysis of central pathway genes through RT-qPCR. RNA used from a co-culture assay using the lentiviral protocol. N=3.

LS 174T MC	Gene average C_T ±SD	Δ C_TGene-GAPDH ±SD	Fold Expression Change
MTOR	27,071± 0,179	8,398± 0,257	1,573
STAT3	34,125± 0,900	15,453± 0,919	1,883
EGFR	27,064± 0,089	8,392± 0,180	1,009
ERBB2	32,982± 0,305	14,094±0,253	3,778
PTGS2	33,047±0,086	14,375±0,182	6,646
NFKB1	28,577±0,223	9,746±0,208	0,933
VIM	36,014±0,536	17,341±0,451	38,271
CDH1	24,801±0,104	6,129±0,125	0,871
AKT	26,080±0,106	7,407±0,202	1,708
MAPK1	26,028±0,094	7,355±0,163	0,877
GAPDH	18,673±0,115	----------	----------
LS 174T CC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
MTOR	26,502± 0,158	8,087± 0,059	1,573
STAT3	33,298± 0,135	14,883± 0,222	1,883
EGFR	27,136± 0,408	8,721± 0,306	1,009
ERBB2	31,150± 0,129	12,735±0,035	3,778
PTGS2	30,057± 0,276	11,642± 0,184	6,646

NFKB1	28,419± 0,080	10,003± 0,149	0,933
VIM	30,498±0,399	12,083±0,409	38,271
CDH1	24,744±0,036	6,329±0,138	0,871
AKT	25,050±0,265	6,819±0,174	1,708
MAPK1	25,960±0,065	7,545±0,131	0,877
GAPDH	18,415±0,102	----------	----------
LoVo MC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
MTOR	27,919± 0,352	8,518± 0,304	0,527
STAT3	34,766± 1,272	15,567± 1,02	0,421
EGFR	26,464± 0,094	7,265± 0,060	0,483
ERBB2	32,717±0,514	13,507±0,374	0,997
PTGS2	24,992±0,130	4,649±0,132	0,690
NFKB1	30,256±0,213	9,983±0,047	1,051
VIM	25,917±0,118	5,498±0,078	0,833
CDH1	27,056±0,180	6,577±0,059	0,607
AKT	26,962±0,258	6,803±0,278	0,994
MAPK1	26,715±0,068	6,320±0,092	0,640
GAPDH	20,404±0,038	----------	----------
LoVo CC	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
MTOR	29,341± 0,351	9,512± 0,292	0,527
STAT3	36,511± 1,494	16,899±1,184	0,421
EGFR	28,013±0,913	7,850±0,216	0,483
ERBB2	33,220±0,857	13,038±0,698	0,997
PTGS2	25,306±0,042	5,124±0,070	0,690
NFKB1	29,961±0,286	9,779±0,196	1,051
VIM	25,959±0,018	5,777±0,036	0,833
CDH1	27,554±0,044	7,372±0,056	0,607
AKT	27,153±0,162	6,971±0,163	0,994
MAPK1	27,136±0,044	6,953±0,073	0,640

GAPDH	20,182±0,047	----------	----------

Table S8. Validation of shRNAs against the gene hits through RT-qPCR. RNA used from a co-culture assay using the lentiviral protocol between LS 174T and LS 174T treated with an shRNA. Each cell treated with shRNA has their own set of reference gene values, which are not represented in this table. N=3.

LS 174T Control	Gene average CT ±SD	Δ CT Gene-GAPDH ±SD	Fold Expression Change
AQP3	31,592± 0,436	6,776±0,397	0,788
GPX2	21,205± 0,031	-3,911± 0,060	0,499
MYT1	31,112± 0,177	5,995± 0,148	0,411
NRIP1	27,294±0,045	2,177±0,052	0,678
PDK4	30,764±0,258	5,820±0,248	0,482
SKIL	26,992±0,063	1,876±0,032	0,962
UGT1A1	26,358±0,230	1,241±0,192	1,526
UGT1A10	26,527±0,291	1,613±0,261	0,875
TBP	25,117±0,040	----------	----------
LS 174T + shRNA	*Gene average CT ±SD*	*Δ CT Gene-GAPDH ±SD*	*Fold Expression Change*
AQP3	31,942±0,187	6,820±0,314	0,788
GPX2	22,069±0,118	-2,897±0,047	0,499
MYT1	32,191±0,415	7,277±0,487	0,411
NRIP1	28,095±0,282	2,926±0,236	0,678
PDK4	31,607±0,263	6,885±0,278	0,482
SKIL	26,357±0,061	1,932±0,061	0,962
UGT1A1	25,325±0,125	0,632±0,098	1,526
UGT1A10	26,971±0,089	1,602±0,054	0,875
TBP	----------	----------	----------