

**Maria Constança Gomes Redinha Pais do
Amaral**

Bachelor Degree in Biology

***In vitro* and *in vivo* models to assess
cancer metastasis**

Dissertation to obtain the Master Degree in Biochemistry for Health

Supervisor: Professor Paula Videira, PhD, FCT/UNL

Cosupervisor: Professor José Ramalho, PhD, FCM/UNL

September 2017

**Maria Constança Gomes Redinha Pais do
Amaral**

Bachelor Degree in Biology

***In vivo* and *in vitro* models to assess
cancer metastasis**

Dissertation to obtain the Master Degree in Biochemistry for Health

Supervisor: Professor Paula Videira, PhD, FCT/UNL

Cossupervisor: Professor José Ramalho, PhD, FCM/UNL

Jury:

President: Prof. Dra. Teresa Catarino, FCT-UNL

Arguer: Dra. Sofia Coroadinha, ITQB, FCT-UNL

**Faculdade de Ciências e Tecnologia da Universidade Nova de
Lisboa**

September 2017

Copyright

***In vivo* and *in vitro* models to assess cancer metastasis**

Maria Constança Gomes Redinha Pais do Amaral

FCT-UNL, UNL

The Faculty of Sciences and Technology and the NOVA University of Lisbon have the right, forever and without geographical limits, to file and publish this dissertation through printed copies reproduced in paper or by digital means, or by any other mean known or that is invented, and to disclose it through scientific repositories and to allow its copying and distribution for non-commercial educational or research purposes, provided that the author and editor are credited.

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

The work developed until the present date has originated:

- **One Oral Presentation:**

Constança Amaral, Roberta Zoppi, Fanny Deschepper, José Ramalho, Paula Videira (2017). *Genetic cell models to assess the role of colorectal cancer-associated glycans*. Jornadas Portuguesas da Genética, June 8th-9th, Aveiro, Portugal.

- **One Paper:**

Carrascal, Mylène; Silva, Mariana; Ramalho, José; Pen, Cláudia; Pascoal, Carlota; Amaral, Constança; Martins, Manuela; Serrano, Isabel; Oliveira, Maria; Sackstein, Robert; Videira, Paula (2017). *Inhibition of fucosylation in human invasive ductal carcinoma reduces E-selectin ligand expression, cell proliferation and ERK1/2 and p38 MAPK activation*.

“BELIEVE IN YOUR INFINITE POTENTIAL. YOUR ONLY LIMITATIONS ARE THOSE YOU
SET UPON YOURSELF”

Roy T. Bennett

Acknowledgements

To my supervisor, Professor Paula Videira, for the opportunity of working in her lab, for all the knowledge, guidance, trust, and advices, but mainly for the words of support when most needed.

To Professor José Ramalho, for accepting to co-supervise me and for always opening his lab door to me.

To Roberta Zoppi (Mommy) for being the best mentor and friend someone can ask for, for never saying no to my cries for help but always letting me fly, for all the peeking while I was in the flow chamber alone for the first times, for the screams of not eating spaghetti properly and mainly for the friendship that is going to follow me for life (even if I must drag her).

To Carlota Pascoal whom I'm very proud of, for helping me every time, for the companionship, friendship, and for the great strength example that she is, always with a smile in her face and a music note.

To all the Glycoimmunology Group for the knowledge transmitted, advices, for all the laughs in the lab, companionship and for our common food problem: Fanny Deschepper, Tiago Ferro, Diana Sousa, Liliana Loureiro, Mylène Carrascal e Zélia Silva.

To our collaborators for providing us the colorectal cancer cell lines.

To Mafalda Sobral for always putting a smile on my face, cheering me up and having my back, for all the craziness and pure profound conversations, that may follow us in eld.

To my family for all the love, patience, care and help for me to proceed with my dreams, putting behind their own needs for mine, in particular to my mom Margarida, my dad José, my grandfather José, my aunt Regina, my sister Joana and my B.

To João, my partner for life, for all the late nights, cheers, support, for the more adorable distractions and friendship.

And last, to the Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa for having me and for the opportunity of developing this work.

Abstract

Cancer metastasis is the major cause of cancer-related deaths. Some proteins/lipids-attached glycans are aberrantly expressed in cancer and this pattern is highly associated with malignancy and tumor progression. Thus, aberrant glycans are promising candidates as cancer biomarkers and therapeutic targets. Glycosyltransferases overexpression is the main mechanism behind tumor-associated glycans such as Sialyl-Tn (STn), α 2,6-sialyl lactosamines and Sialyl Lewis X (SLe^x) antigens.

Colorectal cancer (CRC), the second most common cancer, is described to present STn and SLe^x, which correlate with disease progression. Although no glycan biomarkers yet reached clinical application due to the scarce understanding of the molecular mechanisms and the lack of relevant *in vitro/vivo* models to address them.

Therefore, this study main goal was to establish experimental models of CRC cell lines expressing altered glycans and to evaluate their behavior *in vitro* through: CRC cell lines 1) overexpressing STn and 2) expressing Luciferase. In the first, the LS174T cell line was transduced with *ST6GALNaC I* gene to induce STn expression. In the second, SW48, SW948, SW620 and HT29 cell lines that overexpressed, or not, relevant sialyl and fucosyltransferases were used: ST6Gal 1 responsible for increased α 2,6 sialylation; and FUT 6, responsible for increased SLe^x expression. These cell lines were then genetically modified to express LUCIFERASE (LUC) gene to allow cell tracking.

After *ST6GALNaC I* or *LUC* genes transduction, the cells were phenotyped by flow cytometry to assess glycan expression and Luciferase was assessed by bioluminescence. Results showed that LS174T did not induce STn expression after transduction, since its precursor was absent. Regarding the remaining cell lines transduced with *LUC*, only HT29 *FUT VI* and SW948 *ST6GAL I*, and respective controls, emitted bioluminescence when tested.

This study contributed to the comprehension of cancer-related glycan expression in CRC cell lines. These established cell lines may open the way for future investigation lines of metastasis development *in vivo*.

Keywords: Colorectal Cancer; Sialyl-Tn; Sialyl Lewis^x; *ST6GALNaC I*; *ST6GAL I*; *FUT VI*; *LUCIFERASE*

Resumo

Metástases tumorais são a maior causa de mortes relacionados com cancro. Alguns glicanos, ligados a proteínas/lípidos, estão aberrantemente expressos em cancro e este padrão está altamente associado a malignidade e progressão tumoral. Assim, glicanos aberrantes tornam-se promissores candidatos para biomarcadores tumorais e alvos terapêuticos. A sobre expressão de dadas glicosiltransferases é o principal mecanismo adjacente a glicanos tumorais como é o caso dos antígenos Sialil-Tn (STn), α 2,6-sialil lactosaminas e Sialil Lewis X (SLe^x).

Cancro colorretal (CRC) é o segundo tipo de cancro mais comum e expressa STn e SLe^x, o que se correlaciona com progressão da doença. Apesar de nenhum dos glicanos propostos ter chegado ainda à clínica devido à falta de compreensão dos mecanismos moleculares adjacentes e de modelos *in vitro* e *in vivo*.

Assim sendo, o principal objetivo deste estudo foi estabelecer modelos de linhas celulares de CRC que expressavam glicanos alterados e avaliar o seu comportamento *in vitro* por: estabelecer linhas de CRC que 1) sobre expressassem STn e 2) que expressassem Luciferase. Na primeira, a linha LS174T foi transduzida com o gene *ST6GALNaC I* de forma a induzir a expressão de STn. Na segunda, as linhas SW48, SW948, SW620 e HT29, a expressar ou não relevantes sialil e fucosiltransferases foram usadas; nomeadamente ST6Gal 1, responsável pelo aumento de α 2,6-sialilação; e FUT 6, responsável pelo aumento de expressão de SLe^x. Estas linhas foram geneticamente modificadas para expressar o gene da *LUCIFERASE (LUC)* para permitir o seguimento celular por bioluminescência.

Após transdução com os genes *ST6GALNaC I* ou *LUC*, as células foram fenotipadas por citometria de fluxo para aceder à expressão de glicanos e a expressão de Luciferase foi testada por bioluminescência. Os resultados mostraram que as LS174T, não expressavam STn após transdução, uma vez que o seu precursor estava ausente. Relativamente às restantes linhas quando transduzidas com *LUC*, só as HT29 *FUT VI* e SW948 *ST6GAL I* e respetivos controlos, emitiram bioluminescência.

Este estudo contribuiu para a compreensão da expressão de glicanos-tumorais em linhas celulares de CRC. Estas linhas estabelecidas podem ter aberto o caminho para novas linhas de investigação para o desenvolvimento metastático *in vivo*.

Palavras-chave: Cancro Colorretal; Sialil-Tn; Sialil Lewis^x; *ST6GALNaC I*; *ST6GAL I*; *FUT VI*; *LUCIFERASE*

1 TABLE OF CONTENTS

1.	Introduction	1
1.1	Cancer	1
1.1.1	Colorectal Cancer	2
1.2	Glycosylation	2
1.2.1	Sialic Acids	3
1.2.2	Glycosyltransferases and Glycosidases.....	4
1.2.2.1	Sialyltransferases and Sialidases.....	4
1.2.2.2	Fucosyltransferases	5
1.3	Glycosylation and Cancer	5
1.3.1	Thomsen-Friedenreich Antigens	7
1.3.2	Sialyl-Tn Antigen	8
1.4	Study of glycans in vitro and Luciferase enzyme	8
1.5	Introduction to the aims of this thesis	8
2	Materials and Methods	11
2.1	Culture of Colorectal Cancer Cell Lines	11
2.2	Techniques	11
2.2.1	Transduction.....	11
2.2.2	Flow Cytometry.....	12
2.2.3	Fluorescence Microscopy.....	13
2.2.4	Real Time Polymerase Chain Reaction	14
2.2.5	Luciferase Detection Assay	14
2.3	Colorectal Cancer Cell Lines Characterization	15
2.3.1	Transduction Protocol.....	15
2.3.2	Cell Phenotype Analysis by Flow Cytometry.....	15
2.3.2.1	Flow Cytometry Protocol	15
2.3.2.2	Sialidase Treatment Protocol	16
2.3.2.3	Cell Lines Staining with Antibodies	16
2.3.3	Cell Phenotype Analysis by Fluorescence Microscopy.....	17
2.3.3.1	Fluorescence Microscopy Protocol	17
2.3.4	Gene Expression Analysis of ST6GALNACI by RT-PCR	17
2.3.4.1	RT-PCR Protocol.....	17

2.3.5	Luciferase Expression Assessment	19
2.3.5.1	Luciferase Detection Assay Protocol.....	19
2.3.5.2	Colony Picking for SW48 and SW948 Cell Lines	20
2.3.6	Cell Growth Curves	20
2.4	Statistical Analysis.....	21
3	Results.....	23
3.1	Establishment of a CRC cell line overexpressing STn Antigen	24
3.2	Luciferase Expression and Luciferase Cell Lines Characterization	27
3.2.1	SW948 MOCK LUC Cell Line.....	28
3.2.2	SW620 MOCK, SW620 FUT VI, HT29 MOCK, HT29 FUT VI, SW948 ST6GAL I, SW48 MOCK and SW48 ST6GAL I Cell lines	32
4	Discussion of Results	35
5	Conclusion and Future perspectives	38
6	Bibliography	40
7	Appendixes	45

List of Figures:

Figure 1.1 Hallmarks of Cancer. Adapted from Hanahan, 2011 ³	1
Figure 1.2 N-linked (A) and O-linked (B) glycans representation. From Varki, 2009 ¹¹	3
Figure 1.3 Sialic acids structure and position in the terminal branches of N-glycans, O-glycans and glycosphingolipids at cell surface. A) Usual position of sialic acids. Adapted from Pinho, 2015 ⁴⁹ B) Structure of N-acetyl-neuraminic-acid. Adapted from Park, 2013 ⁵⁰	4
Figure 1.4 Tumor associated carbohydrate antigens structure in normal mammalian cell surface mucins and in tumor cells surface. In normal healthy cells, the glycosylation pattern is more elaborated than in cancer cells. Adapted from Roy, 2013 ⁵¹	6
Figure 1.5 Cancer associated glycans. A) N-glycans are covalently bound to Asn residues of protein membranes, with two or more branches, associated with tumor invasion; B) O-linked glycans are covalently bounded to Ser/Thr residues of glycoproteins or mucins, as SLe ^{x/a} that are overexpressed in cancer-cell mucins; STn and Tn are tumor antigens correlated with invasion, that are truncated forms of O-glycans; C) Glycosphingolipids are sugar linked to the lipidic layer of cell membranes that includes some sialylated structures associated with tumor growth: GM ₃ , GM ₁ , GD ₂ and GD ₃ ; D) Glycosylphosphatidylinositol (GPI)-linked proteins are anchored to the membrane by a glycan covalently linked to a phosphatidylinositol. Glycosaminoglycans can also appear as free chains (E); F) Proteoglycans can be related with cell-growth activation and cell adhesion; G) O-linked N-acetylglucosamine (O-GlcNAc) can be associated with cytoplasmic or nuclear proteins and can even be released to the extracellular matrix. (Glc – glucose, Gal – galactose, Man – mannose, GlcNAc - N-acetylglucosamine, GalNAc - N-acetylgalactosamine, GlcA - glucuronic acid, Fuc – fucose, Xyl – xylose, Sia - sialic acid). Adapted from Fuster, 2005 ²¹	6
Figure 1.6 Thomsen-Friedenreich Antigens biosynthesis pathway. After the linkage of GalNAc to a Ser/Thr residue, Tn antigen can be converted to core 1 structure, T antigen, and further in core 2, in normal cells. In cancer cells, Sialyl-Tn antigen can be produced due to core 1 synthesis inability. Adapted from Pavão, 2011 ⁵²	7
Figure 2.1 Three plasmid packing system. Adapted from System Bio, ⁵³	12
Figure 2.2 Gating strategy for flow cytometry acquisitions. The scheme characterizes the gating steps upon flow cytometry data acquisition. After cell suspension aspiration, the density plot graph FSC-A vs SSC-A on the left is generated and then the gate (indicated in black) is drawn to cover the cell population within 10 000 events. This gate is used to generate a histogram graph of fluorescence channel vs count where the positive cells peak is counted from the end of the unstained peak. At this point the MFI or % of positive cells is extrapolated.	13
Figure 2.3 Bioluminescence formation reaction by firefly luciferase catalysis. Adapted from Promega, 2006 ⁴⁵	15
Figure 2.4 Colony selection and detection protocol. In the scheme above, the colony picking protocol is described. First, some filter paper was cut in small pieces and autoclaved, to only cover one big colony each. The second part consists in soak them with trypsin-EDTA. In the third, each piece of paper is placed on top of a big isolated colony for it to attach to the paper. In the fourth,	

on a 24 well plate, each selected colony is placed in a well along with the respective paper. The next step was to analyze chemiluminescence for each colony, through the result observed in the sixth phase. 20

Figure 2.5 Cell Growth Graph. Cell growth curve with the respective phases of growth. In the yy axis is the Log number of cells and in the xx axis is the correspondent day. Adapted from Sigma ⁴⁶ 21

Figure 3.1 STn assessment by flow cytometry in MDA STn and LS174T ST6GalNaC I cell lines. On the histograms, the yy axis represents the number of cells and the xx axis represents the fluorescence intensity on the BL1-A channel. A) MDA STn⁺ unstained cells (blue), MDA STn⁺ cells with secondary antibody staining (orange) and MDA STn⁺ cells stained with B72.3 (red); B) LS174T ST6GALNaC I unstained cells (blue), LS174T ST6GalNaC I cells with secondary antibody staining (orange) and LS174T ST6GalNaC I cells stained with B72.3 (red); 24

Figure 3.2 ST6GALNaC I Expression Assessment by RT-PCR. In the presented graph is shown the relative mRNA levels of the ST6GALNaC I gene. The yy axis represents the relative mRNA levels and the xx axis the cell line in study..... 25

Figure 3.3 Fluorescence microscopy images of LS174T. On image A the negative control LS174T WT variant is presented, on image B there is LS174T ST6GALNaC I transduced cells stained with just secondary antibody and on C there is LS174T ST6GALNaC I cells stained with anti-ST6GALNaC I antibody. In the three images, the nuclei were detected by ToPro3 as green and ST6GalNaC 1 by Cy5 in red. 25

Figure 3.4 STn and Tn assessment by flow cytometry in LS174T WT and LS174T ST6GALNaC I cell lines with and without sialidase treatment. On the histograms, the yy axis represents the number of cells and the xx axis represents the fluorescence intensity on the BL1-A channel. A) LS174T WT cells with sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green); B) LS174T WT cells with no sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green); C) LS174T ST6GALNaC I cells with sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green). D) LS174T ST6GALNaC I cells with no sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green). 26

Figure 3.5 STn and Tn assessment by flow cytometry in HEK293 Star and LS174T ST6GALNaC I cell lines with and without sialidase treatment. On the histograms, the yy axis represents the number of cells and the xx axis represents the fluorescence intensity on the BL1-A channel. A) HEK293 ST6GALNaC I cells with sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 3F1 staining (green); B) HEK293 ST6GALNaC I cells with no sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green); C) LS174T ST6GALNaC I cells with sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green). D) LS174T

ST6GALNACI cells with no sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green)	27
Figure 3.6 Luciferase expression in SW948 MOCK LUC transduced cells in a 96 well plate analyzed by Chemidoc. 96 well plate bioluminescence from SW948 MOCK LUC cell lysates from transduced cells with LUC gene. From well A1 to well B12, there were 24 different isolated colonies lysates from SW948 MOCK LUC cell line with luciferin reagent. The C1 well had SW948 MOCK LUC lysate without luciferin reagent, C2 had SW948 MOCK lysate with luciferin reagent and C3 had only luciferin reagent. Luciferase expression was higher in B1, B3 and B8 wells. .	28
Figure 3.7 Glycan expression assessment by flow cytometry in SW948 MOCK and SW948 MOCK LUC cell lines. On the histograms, the yy axis represents the number of cells and the xx axis represents the fluorescence intensity on the BL1-A channel on the right and on RL1-A channel on the left. A) SW948 MOCK cells: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange), with CA19.9 staining (light green), with CD15s staining (dark green), with CD15 staining (light purple), with anti-CEA staining (purple), with anti-Lewis Y staining (dark blue) and with 5F4 staining (yellow); B) SW948 MOCK cells: unstained cells (red), with secondary antibody staining (blue), with HECA-452 staining (orange) and with anti-CD44 staining (green); C) SW948 MOCK LUC cells: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange), with CA19.9 staining (light green), with CD15s staining (dark green), with CD15 staining (light purple), with anti-CEA staining (purple), with anti-Lewis Y staining (dark blue) and with 5F4 staining (yellow); D) SW948 MOCK LUC cells: unstained cells (red), with secondary antibody staining (blue), with HECA-452 staining (orange) and with anti-CD44 staining (green);	29
Figure 3.8 Cell Growth Curve Graphs from SW948 MOCK and SW948 MOCK LUC cell lines. The yy axis corresponds to the Log10 of cell/mL and the xx axis corresponds to days. The first graph describes the cell growth curve of SW948 MOCK converted. The second graph describes the cell growth curve of SW948 MOCK LUC.	30
Figure 3.9 Luciferase expression in HT29 MOCK LUC, HT29 FUT VI LUC and SW948 ST6GAL I LUC transduced cells in a 96 well plate analyzed by Chemidoc. 96 well plate bioluminescence from HT29 MOCK LUC, HT29 FUT VI LUC and SW948 ST6GAL I LUC cell lysates from transduced cells with LUC gene. Wells A1 and A2 had HT29 MOCK LUC cell lysates duplicates with luciferin reagent and A3 had HT29 MOCK LUC lysates without luciferin reagent. B1 and B2 wells had HT20 FUT VI LUC cell lysates duplicates with luciferin reagent and B3 had HT29 FUT VI LUC cell lysates without luciferin reagent. The C1 well had contained only luciferin reagent. Well F1 had SW948 ST6GAL I LUC cell lysates without luciferin reagent, well F2 had only luciferin reagente and from G1 to H10 wells there were present SW948 ST6GAL I LUC cell lysates with luciferin reagent.	33
Figure 7.1 Symbol nomenclature adopted for glycan stuctures.	45
Figure 7.2 Text nomenclature adopted for glycan stuctures, both linear and 2D. Adapted from Nomenclature Committee Consortium of Functional Glycomics ⁵⁴	46
Figure 7.pLenti6/V5-D-TOPO plasmid used for transduction	46

List of Tables:

Table 2.1 Antibody description table. Primary antibodies list, correspondent antigens and secondary antibodies, quantities used and fluorescence, used in cancer expressing glycans screening. 16

Table 2.2 Conversion mix content and respective amounts (μL)..... 18

Table 2.3 Settings of the used program to cDNA conversion 18

Table 2.4 RT-PCR mix content and respective amounts 19

Table 2.5 Conditions for RT-PCR assay 19

Table 3.1 Summary table for the used cell lines 23

Table 3.2 Cell Morphology During Prolonged Culture. Table shows microscope images comparing the confluency and cell morphology in two cell lines the SW948 MOCK and SW948 MOCK LUC. The left columns correspond to the 40x magnification and the ones on the right the 100x magnification. 31

Abbreviations:

APC - Allophycocyanin

ATP – Adenosine Triphosphate

Asn – Asparagine

cDNA – complementary DNA

CFG - Consortium for Functional Glycomics

CMP – Cytidine Monophosphate

CRC – Colorectal Cancer

DMSO – Dimethyl Sulfoxide

DNA – Deoxyribonucleic Acid

ECM – Extracellular Matrix

ER – Endoplasmic Reticulum

EV – Empty Vector

FBS – Fetal Bovine Serum

FITC – Fluorescein isothiocyanate

FSC – Forward-scattered Light

Fuc – Fucose

FUT – Fucosyltransferase gene

FUT VI – Fucosyltransferase VI

Gal – Galactose

GalNAc – *N*-acetylgalactosamine

GDP – Guanosine Diphosphate

Glc - Glucose

GlcNAc - *N*-acetylglucosamine

Le^x – Lewis X

Luc – Luciferase gene

ME – Mercaptoethanol

MFI – Median Fluorescence Intensity

MOI – Multiplicity Of Infection

mRNA – messenger Ribonucleic Acid

NC – Negative Control

Neu5Ac – N-acetyl neuraminic Acid or Sialic Acid

PBS – Phosphate Buffered Saline

PCR – Polymerase Chain Reaction

Pen/Strep – Penicillin and Streptomycin

PFA – Paraformaldehyde

RNA – Ribonucleic Acid

RT-PCR – Reverse Transcriptase Polymerase Chain Reaction

Sia – Sialic Acid

Ser - Serine

SSC – Side-scattered Light

STn – Sialyl-Tn Antigen

ST3Gal I - β -Galactoside α 2,3-Sialyltransferase I

ST6Gal I – β -Galactoside α 2,6-Sialyltransferase I

ST6GalNAc I – *N*-Acetylgalactosaminide α 2,6-Sialyltransferase I

TACAs - Tumor-Associated Carbohydrate Antigens

TE- Trypsin-EDTA

Thr – Threonine

VSV - Vesicular Stomatitis Virus

WT – Wild Type

1. Introduction

1.1 Cancer

Cancer is one of the most prevalent diseases and the second leading cause of death, worldwide. Such high incidence and its related mortality leads to the importance of an early and correct diagnosis and to the discovery of effective treatments, in order to cure or to improve the life quality and expectancy of cancer patients. Lung, liver, colorectal and breast cancers are the most deadly types of cancers, since their etiology is mostly related to dietary or behavioral risk factors¹.

This process occurs by normal cells transformation into malignant cells due to gene mutation that can lead to oncogene activation and/or tumor suppressor genes loss of function. These genetic alterations are the result of the combination of genetics and carcinogens exposition¹. Cancer cells differ from normal cells by the acquisition of certain common capabilities during their development that allows them to survive, proliferate and differentiate in uncontrolled ways. These are called the Hallmarks of Cancer (Figure 1.1). Nowadays, ten common alterations have been described: self-sufficiency in growth signals, insensitivity to anti-growth signals, evasion from cell death signals, unlimited replication capacity, sustained angiogenesis and tissue invasion and metastasis, genomic instability, inflammation, reprogramming of energy metabolism and immune response escape^{2,3}.



Figure 1.1 Hallmarks of Cancer. Adapted from Hanahan, 2011³.

According to the cell type from which the tumor arises, tumors are classified into *carcinomas*, which includes 90% of all cancer, if the primary cell is epithelial; *sarcomas* if the origin is in connective tissue; *lymphomas* for immune system cells; and *leukemias* for blood-forming cells. They are also classified according to their tissue of origin, for example, *adenocarcinoma* have its origin in glandular tissue⁴.

Cancer cells can undergo an invasion process called metastasization. It originates with the loss of cell-cell adhesion ability and it is defined as the tumor capacity to spread and invade from the primary site to adjacent healthy tissue and distant organs. This ability is the main cause of the increasing mortality ratio within cancer patients. While many early stage tumors can be treated with surgery or local irradiation, metastasized ones are more difficult to treat. Cancer cells have genetic instabilities and such heterogeneity that allows them to survive in unfavorable environments⁵. The molecular mechanism behind this process is still unknown, but already some genes and molecules known to be involved with it have been identified. The metastatic process consists of many steps where cells detach from the primary tumor, invade, circulate in the blood stream or in lymph, extravasate, are able to survive, proliferate, and promote angiogenesis⁶. Although most of the cells that enter the circulation do not survive, others develop the ability to adhere the vascular walls in a host organ and extravasate. The adhesion to the extra-cellular matrix (ECM) is mediated by selectins and their respective ligands. Cancer cells can produce collagenase IV which degrades collagen IV, the main component of membranes. When this happens, an inflammatory response is created, and the recruitment of leukocytes and macrophages is initiated, allowing the extravasation and the permeabilization of the membranes/walls. Adhesion molecules such as laminin, fibronectin and collagen are essential for cancer cells to adhere to the endothelium membranes and to attach to basal lamina. It has been described that cell malignancy is associated with overexpression of adhesion molecules receptors, regulating adhesion and motility of these cells⁷. Metastasis have been shown to have preferred spreading sites as happens in bone, in breast and prostate cancers⁸.

1.1.1 Colorectal Cancer

Colorectal Cancer (CRC) is the second most common cause of cancer-related death in women and the third in men⁹. Its causes are various but mostly related with diet habits, prior family cases, age and untreated polyps. What makes this kind of cancer particularly dangerous is the presence of high genetic and molecular heterogeneity among its cells. This high incidence combined with an asymptomatic behaviour translates into an urgent need to develop prognostic tools and therapy methods in order to reduce and avoid disease progression. In fact, most patients with metastatic CRC have higher chances of recurrence, even after surgery and chemotherapeutic treatments.

The majority of CRC are *adenocarcinomas*. They usually originate with a polyp in the colon or rectum that can, sometimes, develop into a tumor: in this scenario, preliminary detection through colonoscopy becomes an essential prevention tool¹⁰. However, CRC can derive in metastatic CRC within the liver and without proper biomarkers, it is very difficult to detect by endoscopy.

1.2 Glycosylation

All mammalian cells have on their surface a layer of sugars, the glycocalyx. Glycans have an important role in biological activities such as cell-cell interactions, cell-ECM interactions, differentiation, adhesion

together with motility, cell growth, immune response, amongst others¹¹, since they are the first face of interaction regarding to cells.

Glycosylation is the most common post-translational modification process in eukaryotic cells, where carbohydrate donors and carbohydrate acceptors react through enzyme mediation, to form chains that are linked to specific protein and lipid glycosites. Glycosyltransferases are the enzymes responsible for glycan biosynthesis and glycoconjugates formation using a single nucleotide-sugar as a donor substrate. This mainly occurs in the lumen of the Endoplasmic Reticulum (ER) and Golgi Apparatus, but also in the Cytoplasm and Plasma Membrane. On other hand, glycosidases are the enzymes that are responsible for hydrolyzing some specific glycan linkages ¹².

There are a wide variety of glycans due to monosaccharide composition, branches, and different linkages. Therefore, a nomenclature was created, the Nomenclature Committee of the Consortium for Functional Glycomics (CFG), where glycans structures are represented as symbols and text (see Appendix I).

Glycosylation can produce several types of glycoconjugates or glycans. From the five classes known, this thesis will only focus on two: O-glycans or N-glycans. N-glycosylation is the binding of a N-acetylglucosamine (GlcNAc) residue to the amine group of an asparagine (Asn) residue of a protein side chain, especially to the motif Asn-X-Ser/Thr and O-glycosylation is the linkage of a N-acetylgalactosamine (GalNAc) residue to the oxygen atom of a subset of serines (Ser) and threonines (Thr)¹² as illustrated in Figure 1.2.

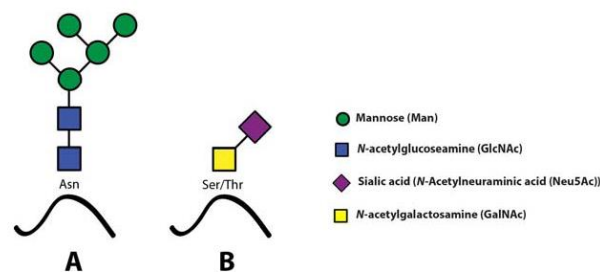


Figure 1.2 N-linked (A) and O-linked (B) glycans representation. From Varki, 2009¹¹.

O-glycosylation process steps occur in the *trans*-Golgi network by the linkage of monosaccharide units by the sequential action of glycosyltransferases. The position in the Golgi lumen, the level of expression and the specificity of each of these enzymes determines the O-glycans pattern¹³, being Tn is the simplest existing O-glycan.

1.2.1 Sialic Acids

Sialic acids (Sia or N-acetylneuraminic acid (Neu5Ac)) are the simplest monosaccharide family of the glycan chain of glycoproteins and glycolipids. Many cell surface glycoproteins and glycolipids can be

modified by sialylation through the linkage of Sia. Due to their several α -linkage possibilities between its nine-carbon backbone, their diversity is vast. Usually they are situated at the terminal branches of *N*-glycans, *O*-glycans and glycosphingolipids (Figure 1.3) and their negative charge center can be H-bonded to the donor groups of cytidine-5'-monophosphate-sialic acid (CMP-sialic acid) by sialyltransferases.

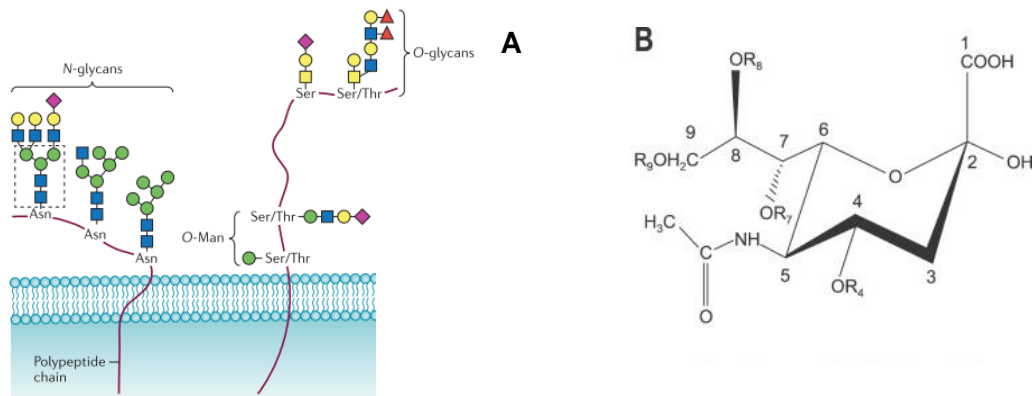


Figure 1.3 Sialic acids structure and position in the terminal branches of *N*-glycans, *O*-glycans and glycosphingolipids at cell surface. A) Usual position of sialic acids. Adapted from Pinho, 2015⁴⁹ B) Structure of *N*-acetyl-neuraminic-acid. Adapted from Park, 2013⁵⁰

Given their non-reducing terminal position in cell membranes, they have several important biological roles. They are related with various biological processes such as cell/cell interactions, B and T cell activation in the immune system and cell differentiation, maturation and they constitute the attachment site for pathogens. An altered sialylation is also implicated in tumorigenicity such as metastasis of malignant cells¹⁴.

1.2.2 Glycosyltransferases and Glycosidases

The enzymes in charge for glycan biosynthesis are called glycosyltransferases and act by the sequential addition of monosaccharides from a high-energy nucleotide sugar donor to specific positions on determined precursors, originating linear or branched sugar chains. Most are found in the ER/Golgi pathway¹⁵.

Glycosidases are responsible for the degradation of sugar chains. These may be endo or exoglycosidases and their action is correlated with low pH levels. Exoglycosidases act by cleaving the linkage of terminal carbohydrates from the nonreducing end. On other hand, endoglycosidases cleave internal glycosidic linkages of larger chains that can later be degraded by exoglycosidases¹⁵.

1.2.2.1 Sialyltransferases and Sialidases

Sialyltransferases are the glycosyltransferase family responsible for adding CMP-activated Sia residues to specific terminal non-reducing positions on oligosaccharides of proteins or lipids. The substrates can

be galactose residues (Gal), GalNAc or Sia. These enzymes are type II membrane proteins from the Golgi apparatus, therefore, Sia are added during glycoproteins biosynthesis.

The sialyltransferase enzyme family vary according to their tissue distribution and to the Sia linkage that they form. Sia can be added by some of them, such as: ST3Gal 1 sialyltransferase in α 2,3 linkage to galactose; ST6Gal 1 sialyltransferase in α 2,6 linkage to galactose or by ST6GalNAc to *N*-acetylgalactosamine; or by the α 2,8 linkage by a family of polysialyltransferases to other Sia. The aberrant expression of these enzymes is highly correlated with cancer. For the aims of this thesis only ST6Gal 1 and ST6GalNAc 1 are going to be considered.

ST6GAL 1 gene is present in chromosome 3 in humans and in chromosome 4 in mice and codes for ST6Gal 1 enzyme. This enzyme is a type II membrane protein, but it can also be found in the soluble form. It is highly expressed in lymphocytes and hepatocytes where it is the mainly responsible for catalyzing α 2,6 sialylation of membrane glycoproteins and of serum glycoproteins, respectively, on *N*-glycan cases ¹¹.

ST6GalNAc 1 is the human enzyme responsible for adding Sia in α 2,6 linkage to *O*-linked GalNAc residues linked to a Ser or Thr, leading to Sialyl-Tn antigen (STn) a cancer related antigen that is one of the main incidences on this work.

Sialidases or neuraminidases are the enzymes responsible for catalyzing the cleavage of the Sia from glycans: they have been described as having an important role in the pathogenesis of diseases¹⁶, since they can be found in lysosomes and endosomes that degrade glycoproteins.

1.2.2.2 Fucosyltransferases

Fucosyltransferases (FUT) are a glycosyltransferase family that transfer a fucose residue from a guanosine diphosphate fucose (GDP-fucose) donor to glycoconjugate acceptors. Human *FUT* genes can be divided into α 1,2-FUT, α 1,3/4-FUT and α 1,6-FUT. The first family is the responsible for blood group antigens, such as ABO, and Lewis γ (Le^y). In the second family, we have From FUT 5 to FUT 7 and from FUT 9 to FUT 11. One of them, FUT6 is responsible for Lewis x (Le^x) and Sialyl-Lewis x (SLe x) biosynthesis¹⁷, another important cancer related antigens that this thesis is going to refer to.

1.3 Glycosylation and Cancer

Due to their importance in several biological aspects, the glycans are also expressed during crucial steps of tumor progression and metastasis development in aberrant forms, increasing tumor malignancy¹⁸. Cancer cells exhibit a high variety of altered glycosylation cell surface patterns such as loss or overexpression of structures such as complex *N*-glycans and incomplete or truncate *O*-glycans, accumulation of precursors and novel structures, all modifications that are absent or diminished in normal healthy cells¹⁵. These altered carbohydrate structures are named tumor-associated carbohydrate antigens (TACAs). TACAs can be divided in two categories: glycoprotein antigens and glycolipid antigens¹⁹. This can be the result of an abnormal expression and activity of

glycosyltransferases or glycosidases or also due to their deallocated position in the Golgi Apparatus as well as due to fluctuations on pH levels²⁰.

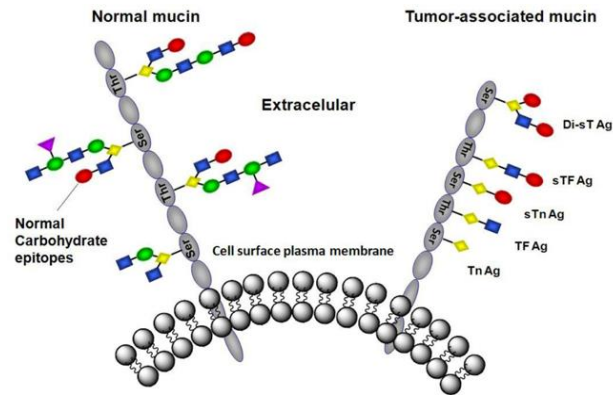


Figure 1.4 Tumor associated carbohydrate antigens structure in normal mammalian cell surface mucins and in tumor cells surface. In normal healthy cells, the glycosylation pattern is more elaborated that in cancer cells. Adapted from Roy, 2013⁵¹.

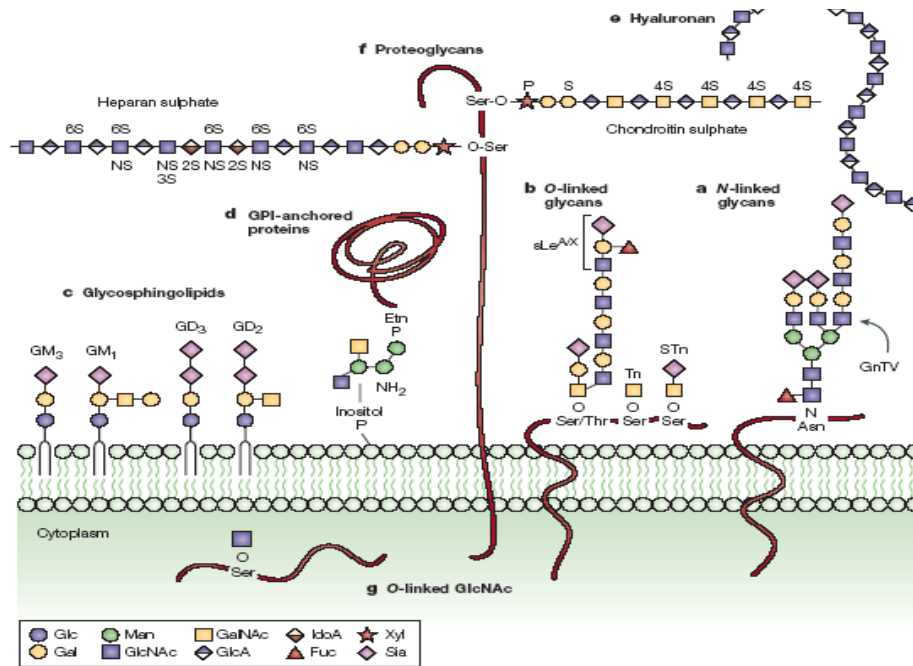


Figure 1.5 Cancer associated glycans. A) N-glycans are covalently bound do Asn residues of protein membranes, with two or more branches, associated with tumor invasion; B) O-linked glycans are covalently bounded to Ser/Thr residues of glycoproteins or mucins, as SLe^xa that are overexpressed in cancer-cell mucins; STn and Tn are tumor antigens correlated with invasion, that are truncated forms of O-glycans; C) Glycosphingolipds are sugar linked to the lipidic layer of cell membranes that includes some sialilated structures associated with tumor growth: GM₃, GM₁, GD₂ and GD₃; D) Glycosylphosphatidylinositol (GPI)-linked proteins are anchored to the membrane by a glycan covalently linked to a phosphatidylinositol. Glycosaminoglycans can also appear as free chains (E); F) Proteoglycans can be related with cell-growth activation and cell adhesion; G) O-linked N-acetylglucosamine (O-GlcNAc) can be associated with cytoplasmic or nuclear proteins and can even be released to the extracellular matrix. (Glc – glucose, Gal – galactose, Man – mannose, GlcNAc - N-acetylglucosamine, GaINAc - N-acetylgalactosamine, GlcA - glucuronic acid, Fuc – fucose, Xyl – xylose, Sia - sialic acid). Adapted from Fuster, 2005²¹.

Aberrant glycosylation leads to, consequentially, altered cell/cell and cell/matrix interactions which can explain some abnormality of tumor cells such as unlimited growth, tumor invasion and metastatic potential²² that combined with absence in normal cells, make glycans good biomarkers for diagnosis, prognosis and therapy. Some studies already correlate the high level of some truncated O-glycans with low rates of survival in cancer patients^{23,24}. Sialylated structures of Thomsen-Friedenreich (TF) antigens are also examples of truncated glycans. Many of these glycans are already used in clinical practice and are related to immune system escape, attachment to capillaries in tumor invasion and new adhesion in metastasis. This aberrant glycosylation can be detected in blood tests since some of these are present in the serum¹⁸.

α 2,6-sialyltransferases have been shown to be upregulated in some tumor cells. ST6Gal I overexpression has already been described in several carcinomas, such as colon, and some studies indicate its upregulation is due to oncogenes activation such as *RAS*. The high activity of this enzyme is correlated with a poor prognosis and metastasis²⁵. The same has been described for ST6GalNAc 1 enzyme, since its overexpression is the mainly responsible for STn formation and upregulation, a cancer related antigen that has been found in several carcinoma types and that is associated with poor prognosis and metastasis development²⁶.

1.3.1 Thomsen-Friedenreich Antigens

TF antigens or disaccharides Gal β 1-3GalNAc α 1-Ser/Thr or T antigens, are the core 1 structure of O-linked mucins, thus being present in membrane glycoproteins. They are truncated in the early stages of glycosylation²⁷ and have been correlated with tumor progression and metastases and also the natural human production of antibodies against them is very low in cancer patients²⁷. Cancer cells can lose the ability to synthesize core 1 and in that case, the Tn antigen, GalNAc α 1-Ser/Thr, is exposed which is also a carcinoma correlated antigen that occurs in approximately 90% of human cancers. Lastly, Tn can be further sialylated originating the STn antigen.

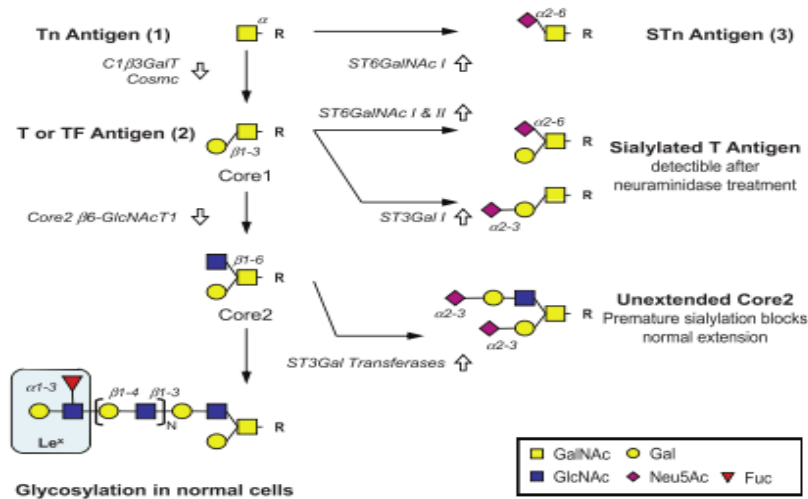


Figure 1.6 Thomsen-Friedenreich Antigens biosynthesis pathway. After the linkage of GalNAc to a Ser/Thr residue, Tn antigen can be converted to core 1 structure, T antigen, and further in core 2, in normal cells. In cancer cells, Sialyl-Tn antigen can be produced due to core 1 synthesis inability. Adapted from Pavão, 2011⁵².

1.3.2 Sialyl-Tn Antigen

STn (Neu5Ac α 2-6GalNAc α -O-Ser/Thr or CD175s) is a truncated O-glycan antigen used as an epithelial cancer biomarker. It has been used not only for diagnosis and prognosis, but also for immunotherapy targeting since it is not found in healthy tissue. It is the simplest sialylated mucin-type O-glycan and it is composed by a GalNAc residue α -O-linked to a Ser or Thr residue in a protein that is sialylated on carbon 6.

The sialylation of a single O-GalNAc has been shown to be performed *in vitro* by two sialyltransferases, ST6GalNAc 1 and ST6GalNAc 2²⁸. However, studies using transfected cells by either one of these enzymes has demonstrated that in a cellular context, only ST6GalNAc 1 overexpression generates STn structures recognized by anti-STn antibodies^{29,26,30}. It can be associated with a variety of membrane proteins such as mucins, integrins and CD44.

It is expressed in almost all carcinomas and in nearly 80% of CRC. STn expression is associated with oncogenesis and the metastatic ability of cancer cells, being therefore correlated with poor prognosis in cancer patients.

1.4 Study of glycans *in vitro* and Luciferase enzyme

To track cancer cells *in vivo* and *in vitro*, the Luciferase enzyme (LUC) can be overexpressed in cells and then used as a reporter. Luciferase is an enzyme that efficiently catalyzes light production through a chemical reaction, present in bioluminescent living organisms. The chemical reaction requires luciferin, Mg²⁺ and O₂ as co-factors, and, in case of insects, also ATP³¹. Through cDNA cloning, the *LUCIFERASE* gene can be introduced into the cancer cells genome by transduction and further expressed. It is widely used as a reporter gene and the detection assay is very sensitive one³². Its detection can be assessed by direct bioluminescence, alternatively, the cells need to be lysate for the enzyme to be approachable.

Firefly luciferase has a wide potential and applicability since it has no endogenous activity in the choose host cells, is able to be detected by a reproducible assay and the protein does not need posttranslational modifications.

In this thesis work, the *LUCIFERASE* gene was cloned in a lentiviral vector and this was transduced into CRC cell lines to be further tracked in living organisms, making sure that this genetic modification in the CRC cell lines did not change their glycan expression either way.

1.5 Introduction to the aims of this thesis

Since metastasis is the major cancer related cause of death, it is extremely important to uncover the molecular mechanism behind metastasis formation and development as a basis to their prevention. Since colorectal metastatic cancer is one of the deadliest, our study is concentrated on this type of cancer.

CRC cells can produce mucin glycoproteins with O-linked oligosaccharides that have been found to express TF antigens. Tn and STn antigens, both examples of these, shown to be highly sensitive and specific tumor-associated antigens in the colon that are absent in normal colon tissue³³, making them potential biomarkers. Also, the SLe^x antigen is another highly expressed glycan in solid tumors, that have potential as biomarker. This aberrant glycosylation is associated with a poor prognosis, cancer progression and metastasis development.

STn is expressed when over or incomplete synthesis of α 2,6-sialylation occurs due to ST6GalNAc 1 overexpression and it has been found in 80% of human *carcinomas* such as colon *adenocarcinomas* and that is associated with a poor prognosis^{35,20}. Moreover, SLe^x provide ligands for selectin adhesion receptors on activated endothelia, platelets or leukocytes what are thought to be involved in metastasis extravasation and its presence is due to the overexpression of fucosyltransferase FUT 6³⁶. ST6Gal 1 expression is upregulated in several *carcinomas* being CRC among them.

In this study, the main goal was to establish experimental genetically modified models based on CRC cell lines expressing both altered glycans and Luciferase reporter. To achieve this, the work consisted in two arms.

In the first part, we aimed to establish a CRC cell line overexpressing the STn antigen. To assess this, the LS174T cell line was transduced with the *ST6GALNAc 1* gene. This transduced cell lines were phenotyped to assess the expression of STn antigen. Since this cell line naturally express high levels of other relevant glycans, namely the SLe^x, but not of STn, it was expected to obtain a CRC cell line that had a dual overexpression of STn and SLe^{x/a} antigens. This cell line could also be used in further transduction experiments to overexpress the LUC gene. However, after several attempts to transduce the cell line with lentivirus containing the *ST6GALNAc 1* gene, it was not possible in this work to obtain the desired LS174T cell line overexpressing STn. Further studies are necessary to understand the mechanism behind these results.

In the second part, we establish CRC cell lines expressing LUC enzyme. We used the SW48, SW948, SW620 and HT29 CRC cell lines that were previously transduced and transfected with sialyl and fucosyltransferases to express relevant altered glycans. Namely the ST6Gal 1 sialyltransferase, responsible for increased α 2,6 sialylation, and the FUT 6 fucosyltransferase, responsible for increased SLe^x expression, respectively. These cell lines were posteriorly genetically modified to express the reporter LUC gene through transduction with pLenti6/V5 directional TOPOvector, containing the LUC cDNA. Luciferase expression was tested by bioluminescence and the cell lines were characterized in terms of glycosylation profiles, comparing them with their negative control (NC) or MOCK cell lines, since previous studies shown that there is no difference between MOCK and WT.

2 Materials and Methods

2.1 Culture of Colorectal Cancer Cell Lines

Several colorectal cancer (CRC) cell lines were used during this work: SW620, LS174T SW48, SW948 and HT29. All CRC cell lines were kindly provided by Professor FABIO DALL'OLIO from the University of Bologna, Italy. As controls, another cell lines were used, MDA and HEK293 Star cells. All of them had the three variants – the wild type (WT); the empty vector (EV) or MOCK or negative control (NC); and *FUT VI* or *ST6GAL I* or STn positive cell lines.

SW620 (ATCC®CCL-227™) cell line derived from a Dukes' type C metastatic *adenocarcinoma* from colon in lymph node from a 51-year-old Caucasian male. LS174T (ATCC®CL-188™) cell line was established from Dukes' type B *adenocarcinoma* of colon from a 58-year-old Caucasian female. SW48 (ATCC®CCL-231™) derived from Dukes' type C, grade IV, *adenocarcinoma* of colon from an 82-year-old Caucasian female. SW948 (ATCC®CCL-237™) was established from Dukes' type C, grade III, *adenocarcinoma* of colon from an 81-year-old Caucasian female. HT29 (ATCC®HTB-38™) derived from an *adenocarcinoma* of colon from a 44-year-old Caucasian female.

The used CRC cell lines were adherent and were kept in T25 and T75 culture flasks (SARSTEDT), in a 37°C incubator (Panasonic) with a humidified atmosphere and 5% CO₂. The culture media used was dependent on the cell line: SW48, SW948 and SW620 cell lines were cultured in Leibowitz's medium (Lonza) supplemented with 10% of Fetal Bovine Serum (FBS) (Gibco), 2mM of L-glutamine (Gibco) and 100 µg/mL Penicillin/Streptomycin (Gibco); LS174T were cultured in MEM (Biowest) supplemented with 10% FBS (Gibco), 2Mm L-glutamine (Gibco), 100 µg/mL Penicillin/Streptomycin (Gibco), Non-Essential Aminoacids (Gibco) and Sodium Pyruvate (Gibco); HT29 and MDA cell lines were cultured with DMEM (Gibco) supplemented with 10% de FBS (Gibco), 2mM of L-glutamine (Gibco) and 100 µg/ml Penicillin/Streptomycin (Gibco). The medium was changed every three days. For passages, at 80/90% of confluency, the culture was washed with 1X phosphate buffered saline (PBS) and 1X Trypsin-EDTA (TE) (Gibco) was used for cell detachment from the flask, followed by a centrifugation (Eppendorf) step at 200xg for 5 minutes and cultured according to the desired dilution. All cell lines were stored at -80°C resuspended in culture medium and 10% (v/v) o dimethyl sulfoxide (DMSO) (Sigma).

2.2 Techniques

2.2.1 Transduction

Transduction is a gene transfer mechanism through viral infection of host cells that allows genome integration and the consequent expression of the gene of interest. In this thesis work, the viral vector used was a lentiviral one.

To construct it, a three-plasmid system was used that contemplates a lentiviral construct with the cDNA sequence of interest, an expression vector with the vesicular stomatitis virus G envelope glycoprotein

gene (VSV-G) and a packing vector. They also have present truncated long-terminal residues (LTR) for self-inactivation³⁷. The coding region of human *ST6GALNAc I* gene was before amplified by PCR and cloned in the pLenti6/V5 Directional TOPO cloning vector containing the CMV promoter, as well for LUCIFERASE. Lentiviral particles were generated by co-transfection of HEK 293 T cells with the three plasmids, using pPACK Lentivector Packing Systems, what was kindly performed by Professor Doutor JOSÉ RAMALHO, CEDOC's. Our plasmid construct is presented on Appendix 3.

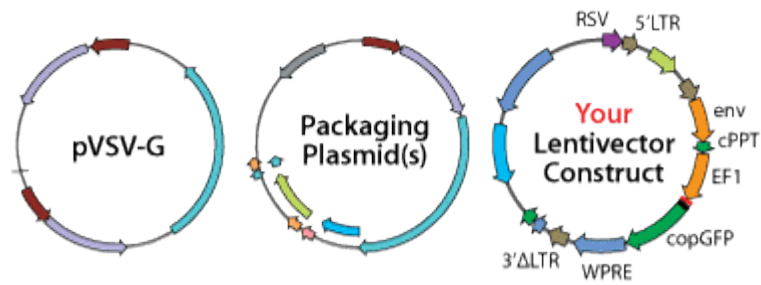


Figure 2.1 Three plasmid packing system. Adapted from System Bio,⁵³

2.2.2 Flow Cytometry

Flow Cytometry analyzes single particles in suspension such as cells, while they flow in a fluid, through a light beam, usually a laser. It is a technique that is able to measure simultaneously particle size, complexity, and fluorescence intensity. These characteristics are determined by how the particle scatters light since its structural complexity scatters light at different angles and quantities. When particles are stained with a fluorophore, the incidence laser excites this molecule making it emit light with a certain wavelength. The fluorescent emitted light is detected, filtered, and amplified to be converted into digital data by a flow cytometer. A flow cytometer allows the determination of the median fluorescence intensity (MFI) or percentage of positive cells that are proportional to the amount of antibody bonded that is proportional to the amount of expressed antigen.

Forward-scattered light (FSC) is proportional to the relative size and area of the particle since scatter at lower angles. Side-scattered light (SSC) is proportional to the cell complexity or granularity, as scatters at larger angles. The correlation of these two measurements leads to a differential population analysis through their physical characteristics.

When a fluorochrome is conjugated with an antibody that recognizes specific receptors, it can be used to detect a cell type according to the antigen at its surface.

The particle characteristics are determined by an optic-electronic system that registers how the particles disperses light and emits fluorescence. Flow Cytometry its formed by three systems: fluidics, that leads the particles to the laser beam; the optic system that its composed by the laser beam and the optic filters that direct the light signals to the proper detectors; and the electronics, that processes the light signals into electronic signals^{38,39}.

The Flow Cytometer used in this work was an Attune® Acoustic Focusing Cytometer (Applied Biosystems) with two lasers, red (638nm) and blue (488nm) that allows up to six different fluorescence's, four for blue laser (BL1: green, BL2: orange, BL3: red and BL4: red) and two for the red laser (RL1: light red and RL2: dark red).

At least 1×10^4 events were acquired in the Attune Cytometric Software (version 2.1) and all data was analyzed in FlowJo (version10) and GraphPad Prism (version7). Upon the flow cytometry acquisition, the strategy adopted to gate cell population was as described on Figure 2.2.

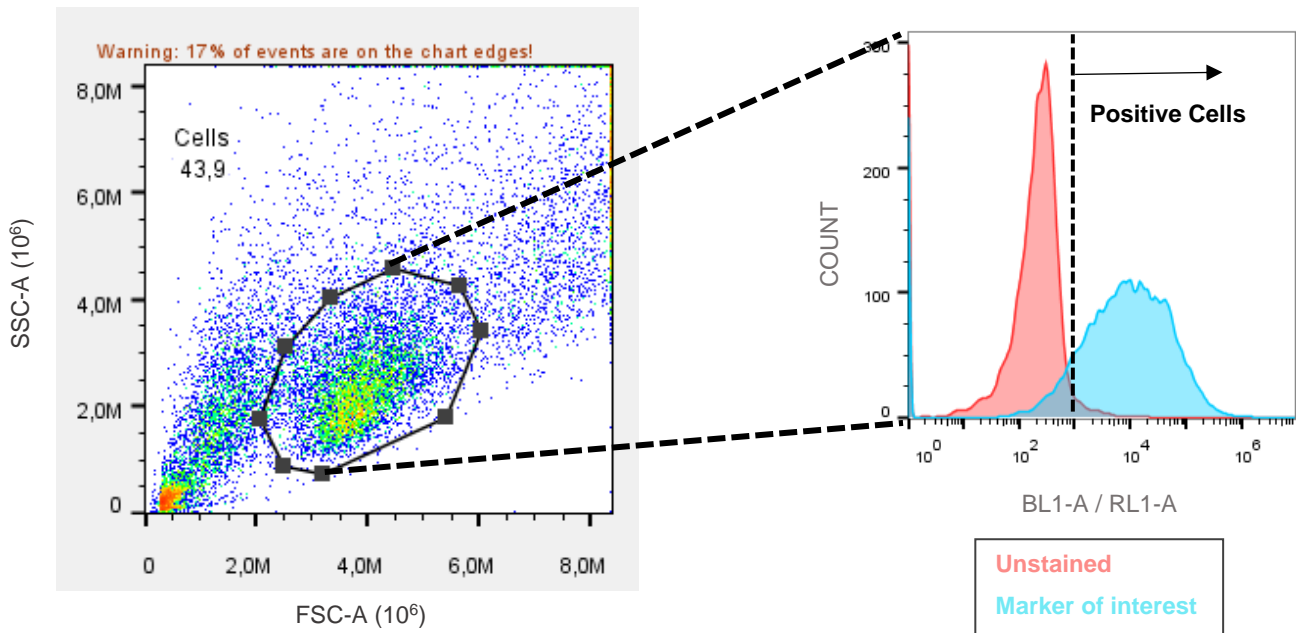


Figure 2.2 Gating strategy for flow cytometry acquisitions. The scheme characterizes the gating steps upon flow cytometry data acquisition. After cell suspension aspiration, the density plot graph FSC-A vs SSC-A on the left is generated and then the gate (indicated in black) is drawn to cover the cell population within 10 000 events. This gate is used to generate a histogram graph of fluorescence channel vs count where the positive cells peak is counted from the end of the unstained peak. At this point the MFI or % of positive cells is extrapolated.

2.2.3 Fluorescence Microscopy

Fluorescence Microscopy is an optical microscopy type that uses fluorescence to generate an image by enlarging the optic resolution and the contrast of the microscopy image. The image is created using an opening, the *pinhole*, that allows a better definition to be obtained in thicker samples.

A laser is focused in the limited place for diffraction in the sample from the objective. The emitted fluorescence is focused by the same lens to a specific point in the detector. The *pinhole* is then placed at this point such that the detector only captures the light prevent from the focal plane. By plotting several planes, it is possible to generate a three-dimensional image of the sample that is processed by a computer⁴⁰.

The microscope used was a Axiolmager D2 – Zeiss, with a 40x objective and the laser filters of 488 nm and 633 nm to excite FITC and Alexa Fluor 633, respectively. The images obtained were processed by Fiji.

2.2.4 Real Time Polymerase Chain Reaction

The Real Time Polymerase Chain Reaction (RT-PCR) is an enzymatic process that converts RNA into its complementary DNA (cDNA) by reverse transcriptase enzyme action, followed by cDNA amplification by PCR⁴¹. It is a very powerful and sensitive tool to analyze and quantify the expression of a gene by mRNA expression. In this thesis, TaqMan technology was used. TaqMan probes are oligonucleotide hydrolysis probes that are dually labelled with a fluorophore linked to the 5' end with a non-fluorescent quencher linked to the other end of the probe. When closer, the quencher reduces the fluorescence emitted by the fluorophore but with the presence of the target sequence, the probe hydrolysis with it and the exonuclease activity of Taq DNA polymerase, degrades the probe annealed to the target sequence, emitting fluorescence that is detected by the device. This fluorescence emitted is proportional to the amplification⁴².

Quantification is performed by relative quantification where the change of expression of the target gene is described by comparison with the expression of an endogenous control gene. The ideal reference gene should have a stable and secure expression regardless of the conditions in which it is used, since they are vital to cell survival. In this thesis two housekeeping genes were used, *GAPDH* (that codes for glyceraldehyde-3-phosphate dehydrogenase) and β -actin. The relative expression was analyzed using the comparative C_T mathematical method, $2^{-\Delta\Delta C_T}$ method, from Livak and Schmittgen⁴³. C_T stands for Threshold Cycle and corresponds to the intersection of the amplification curve with the threshold line. The threshold fluorescence is defined according to the point at which the fluorescence rises significantly above the background fluorescence level. ΔC_T corresponds to the variation between the amplification value of the gene of interest and the endogenous control gene, so $\Delta\Delta C_T$ will correspond to the variation between the amplification values of the gene of interest and the variation between all the controls amplification⁴⁴. Therefore, $2^{-\Delta\Delta C_T}$, will correspond to the amount of target when normalized with the reference.

2.2.5 Luciferase Detection Assay

Luciferase is an oxidoreductase enzyme capable of catalyzing light production in living organisms, bioluminescence, through a chemical reaction of luciferin oxidation that can act as a genetic reporter. Luciferin oxidation and an electron transition, form oxyluciferin. Firefly luciferase cDNA is a 1,6kb fragment that codes for 550 aminoacid residues and a 61kDa protein. The enzymatic reaction requires the cofactors luciferin, ATP, Mg^{2+} and O_2 ⁴⁵. Bioluminescence is a special kind of chemiluminescence present in living organisms that is transformed to visible light. Firefly luciferase emits a yellow-green light corresponding to a wavelength of 560nm.

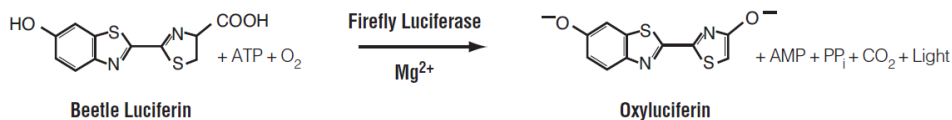


Figure 2.3 Bioluminescence formation reaction by firefly luciferase catalysis. Adapted from Promega, 2006

The luciferase assay offers a nearly instantaneous measure of total luciferase expression in the lysate; therefore, the emitted light is proportional to the luciferase concentration. Although the high efficiency of the assay, the luciferase half-life is less than three hours due to its low intrinsic stability, as a disadvantage.

2.3 Colorectal Cancer Cell Lines Characterization

2.3.1 Transduction Protocol

Before transduction itself, 5×10^5 cells per well in a 12 well plate, were seeded. The next day, the virus were thawed at room temperature in a concentration of 150 μL of virus to a concentration of 8 $\mu\text{g}/\text{ml}$ of polybrene in 2 ml of medium. The cells were washed with 1X PBS and this mixture was placed in the wells. The plate was shaken and incubated overnight, 37°, 5% CO₂. The next day the cells were washed again and the new culture medium was placed with 3,2 $\mu\text{g}/\text{ml}$ of Blastidicin. This selection process was performed during approximately two weeks since the plasmid inside the virus had the resistance to Blastidicin antibiotic and the cells that did not incorporate the plasmid, would die.

2.3.2 Cell Phenotype Analysis by Flow Cytometry

2.3.2.1 Flow Cytometry Protocol

Cancer cells were detached with TE from the culture flasks and washed with 1X PBS by centrifugation at 200xg, for 5 minutes. The resultant pellet was resuspended in 1ml of medium and the cells were counted in a 1:10 dilution in a Newbauer chamber (Marienfeld). 2×10^5 to 3×10^5 cells *per* condition were collected in a centrifuge tube and washed again. At this point, we could proceed to a direct antibody staining or to a sialidase treatment as described forward. The desired concentration of cells was then washed with 500 μl of PBS 1X and centrifuged at 1500xg for 2 minutes. The supernatant was discarded and the previous step was repeated but with 990 μl of 1X PBS. The pellets were resuspended in the desired volume of 1X PBS to be divided as 100 μl *per* condition. The primary antibodies were added and the samples were incubated at 4°C for 30 minutes. Another wash step was performed with 500 μl of 1X PBS and centrifuged at 1500xg for 2 minutes. The supernatant was discarded and the pellets were resuspended in 100 μl of 1X PBS. The secondary antibody was added according to the primary

antibody used and the samples were incubated at room temperature for 15 minutes in the dark. At the end of this step the cells were again washed with 500 μ l of 1X PBS by centrifugation at 1500xg for 2 minutes. The supernatant was discarded, and the pellets were resuspended in 1 ml of 1X PBS to be analyzed by Flow Cytometry. If required, the cells were fixed in 4% Paraformaldehyde (PFA) to be analyzed later.

2.3.2.2 Sialidase Treatment Protocol

The sialidase treatment was performed to remove the sialic acid from structures according to the experimental case, once it had hydrolyses the α 2-6 and α 2-3 sialic acid linkage. After the first wash, the supernatant was discarded, and the cells were resuspended in 400 μ l of sialidase buffer. This 400 μ l of cell suspension was divided between two eppendorf tubes, a sialidase treated one and a non-treated, control. For the sialidase treated one, sialidase (from *Clostridium perfringes*, Roche) was added at a final concentration of 100 mU/ml. Both samples were incubated at 37°C, 5% CO₂ for 90 minutes during which time, the tubes were stirred to avoid cell deposition.

2.3.2.3 Cell Lines Staining with Antibodies

To assess glycans expression at the cell surface, a flow cytometry assay was performed. The primary antibodies used and their correspondent amounts, and the secondary antibody and respective amount and fluorescence method used are detailed in Table 2.1.

Table 2.1 Antibody description table. Primary antibodies list, correspondent antigens and secondary antibodies, quantities used and fluorescence, used in cancer expressing glycans screening.

Primary Antibody	Antigens	Amount	Secondary Antibody	Amount	Fluorochrome
B72.3	STn	1,5 μ L	Goat anti-mouse (Dako)	5 μ L	FITC
5F4	Tn	8 μ L	Goat anti-mouse (Dako)	5 μ L	FITC
A70-C/C8	Lewis ^y	5 μ L	Goat anti-mouse (Dako)	5 μ L	FITC
CD15 (Biolegend)	Lewis ^x	2 μ L	Goat anti-mouse (Dako)	5 μ L	FITC
CD15s (BD Biosciences)	SLe ^x	3 μ L	Goat anti-mouse (Dako)	5 μ L	FITC
antiCD44-APC (Biolegend)	CD44	5 μ L	–	–	APC
HECA-452	SLe ^{x/a}	1,5 μ L	Goat anti-rat	1 μ L	APC

(Biolegend)			(BD Pharmingen)		
CA19.9 (Abcam)	SLe ^a	5 µL	Goat anti-mouse (Dako)	5 µL	FITC
Carcinoembryonic antigen (ImmunoTools)	CEA	5 µL	Goat anti-mouse (Dako)	5 µL	FITC

2.3.3 Cell Phenotype Analysis by Fluorescence Microscopy

2.3.3.1 Fluorescence Microscopy Protocol

0,4 x 10⁶ cells/mL of LS174T WT, LS174T *ST6GALNaC I* and 0,2 x10⁶ cells/mL of MDA STn cells were cultured on top of round coverslips inside 12 well plates (Orange Scientific). After 24h, they were washed three times with 1X PBS for 5 minutes. They were then fixed and permeabilized with 200µl of wash solution (Fixation/Permeabilization Solution Kit from BD Biosciences) for 20 minutes at 4°C. In order to be labeled with the anti-*ST6GALNaC I* antibody (2C3 clone), the cells were incubated with 1 ml of wash solution for 30 minutes and then washed three times with 1X PBS for 5 minutes. The incubation with the primary antibody anti- *ST6GALNaC I* labeled with Cy5, occurred during 30 minutes at a dilution of 1:2 in wash solution. Cells were then washed three times with 1X PBS for 5 minutes. To stain the nucleus, To-Pro3-Alexa Fluor 633 (Invitrogen) was used by covering the cells for 15 minutes and then washed three times with 1X PBS, 5 minutes. Finally, the coverslip was placed face down on the slide, and sealed with montage medium Mowiol (Dabco) and nail polish to be analyzed in the Fluorescence Microscope.

2.3.4 Gene Expression Analysis of *ST6GALNACI* by RT-PCR

2.3.4.1 RT-PCR Protocol

The first step of the RT-PCR protocol was the RNA extraction using GenElute™ Mammalian Total RNA Miniprep Kit (Sigma-Aldrich). A lysis solution/2-ME was previously prepared with 10 µL of 2-Mercaptoethanol (Sigma) to 1 ml of lysis solution and the centrifuge was set at 4°C. Cells were trypsinized and centrifuged at 0,4xg for 5 minutes and washed with 1X PBS. 250 µL of lysis solution/2-ME was added to approximately 5x10⁶ cells and pipetted until no clumps were present. The cell lysate was then collected to a filtration column into a 2ml tube and centrifuged at 12000xg for 2 minutes. The filtration column was discarded and into the resulting lysate was added 250 µL of 70% ethanol and the resulting solution mixed. Up to 700ml of the lysate/ethanol mixture was pipetted into a binding column in a 2ml tube and centrifuged at 12000xg for 2 minutes. The flow-through liquid was discarded but the collection tube was kept to repeat the previous step if the volume was more than 700ml. 250 µL of wash solution 1 was added to the column and centrifuged at 12000xg for 2 minutes. To the mix, 10 µl of

DNase I (Quiagen) were added with 70 μL of digestion buffer (Quiagen) by inversion. From this mixture 80 μL were then added into the filter of the binding column and incubated at room temperature for 15 minutes. 250 μL of wash solution 1 was pipetted into the column, centrifuged at 12000xg for 2 minutes and the flow-through was discarded. The washing step was repeated twice but with 500 μL of wash solution 2. The binding column, now empty, was again centrifuged at 12000xg for 2 minutes because of the remaining liquid and transferred into a new 2ml collection tube. To this, 55 μL of elution solution was pipetted into the binding column and centrifuged at 12000xg for 2 minutes. The flow-through was again passed through the column and centrifuged at 12000xg for 2 minutes since the purified RNA was then in the flow through liquid.

The RNA conversion to cDNA was performed using a High-Capacity cDNA Reverse Transcription Kit as described in the following table.

Table 2.2 Conversion mix content and respective amounts (μL)

Components	Amounts (μL)
Purified RNA	50
Buffer	10
Primers	10
dNTPs	4
Reverse transcriptase enzyme	7,5
RNase free H ₂ O	18,5

The thermocycler machine used was a Programmable Thermal Controller PTC-100™ (MJ Research, Inc.) and it was set according to the settings listed in table 2.3. After the cDNA conversion, the samples could be stored at -20°C.

Table 2.3 Settings of the used program to cDNA conversion

	Step 1	Step 2	Step 3	Step 4
Temperature (°C)	25	37	85	4
Time	10 min	120 min	5 secs	∞

The RT-PCR experiment was performed according to table 2.4 in RT-PCR tubes (Simpport), to a 10 μL final volume.

Table 2.4 RT-PCR mix content and respective amounts

Components	Amounts (μL)
Diluted probe 1:4	2
cDNA	2
RNase free H ₂ O	1
Master Mix Fast (Applied Biosystems)	5

Each experiment was performed in duplicates and using two endogenous controls. The probe of interest was ST6GALNACI (Applied Biosystems) and the endogenous were GAPDH (Applied Biosystems) and β -actin (Applied Biosystems). The samples were run in RotorGene 6000 (Corbett) machine according to table 2.5 and the results were analyzed as previously described followed by GraphPad Prism 7 (GraphPad Software, Inc).

Table 2.5 Conditions for RT-PCR assay

	Step 1 (1 cycle)	Step 2 (40-50 cycles)	
Temperature ($^{\circ}\text{C}$)	95	95	60
Time	20 secs	3 secs	30 secs

2.3.5 Luciferase Expression Assessment

SW48 MOCK and SW48 *ST6GAL I*, SW948 MOCK and *ST6GAL I*, SW620 MOCK and SW620 *FUT VI* and HT29 MOCK and HT29 *FUT VI* were transduced with the lentiviral vector pLenti6/V5-D-TOPO containing luciferase gene from firefly that codes for luciferase enzyme, originating SW48 *MOCK LUC*, SW48 *ST6GAL I LUC*, SW948 *MOCK LUC*, SW948 *ST6GAL I LUC*, SW620 *MOCK LUC*, SW620 *FUT VI LUC*, HT29 *MOCK LUC* and HT29 *FUT VI LUC*.

2.3.5.1 Luciferase Detection Assay Protocol

To assess luciferase expression, a Luciferase Assay System Kit (Promega) was used. This kit can improve the kinetics of the chemical reaction to enhance the light intensity created after reaction by incorporating coenzyme A (CoA).

24 hours before the detection assay, 10^5 cells *per well* were cultured in a 96 well plate. The following day, the medium was removed, and the cells were washed with 1X PBS. 20 μl of 1X lysis buffer RLB from the kit was added to each well and a single freeze-thaw cycle was performed to ensure the complete lysis. At room temperature, 100 μl of luciferase assay reagent were added, composed by luciferase assay buffer and luciferase assay substrate and the samples were read at Chemidoc Touch (Biorad) to check for chemiluminescence, since bioluminescence is a type of chemiluminescence.

2.3.5.2 Colony Picking for SW48 and SW948 Cell Lines

The luciferase enzyme used for transduction in this thesis was isolated from firefly (*Photinus pyralis*). SW948 MOCK, SW948 *ST6GAL I*, SW48 MOCK and SW48 *ST6GAL I* had to be selected by a colony picking method and not by the usual antibiotic one, since they were previously transduced with a plasmid with the same antibiotic resistance.

To performed it, the cell cultures were kept growing until they could form visible and isolated colonies. Twenty-four colonies were then picked by small papers imbibed in trypsin that would detach them. These colonies were cultured in 24 well plates (Greiner Bio-One Cellstar) and then tested to check for chemiluminescence as previously described.

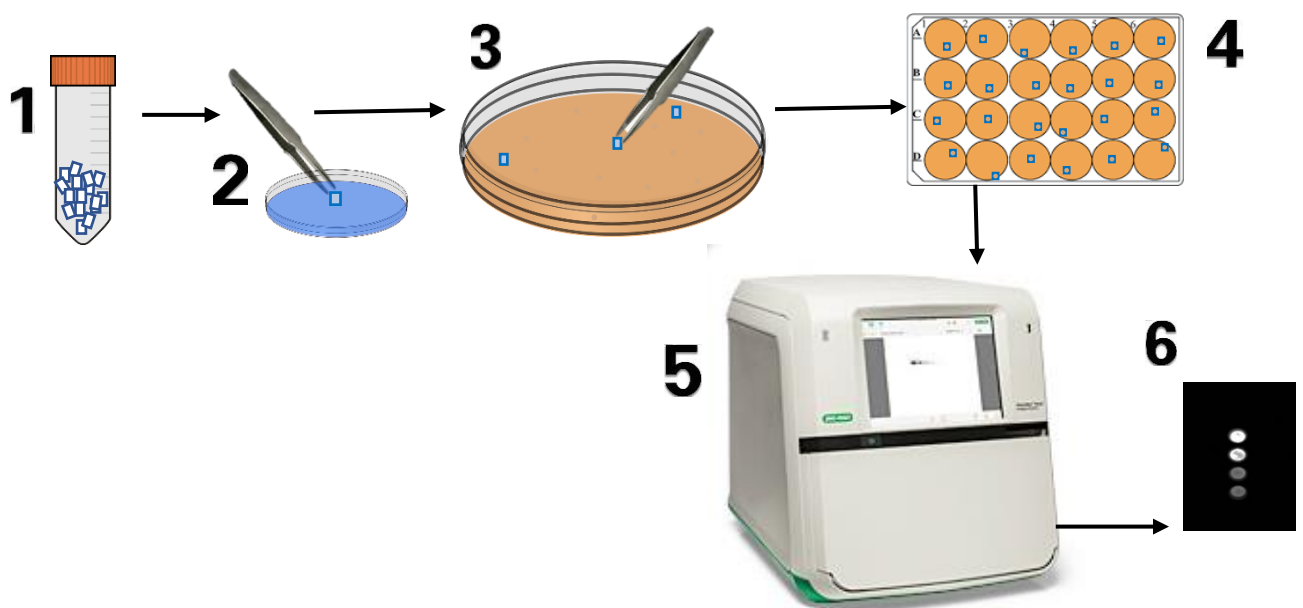


Figure 2.4 Colony selection and detection protocol. In the scheme above, the colony picking protocol is described. First, some filter paper was cut in small pieces and autoclaved, to only cover one big colony each. The second part consists in soak them with trypsin-EDTA. In the third, each piece of paper is placed on top of a big isolated colony for it to attach to the paper. In the fourth, on a 24 well plate, each selected colony is placed in a well along with the respective paper. The next step was to analyze chemiluminescence for each colony, through the result observed in the sixth phase.

2.3.6 Cell Growth Curves

To establish cell growth curves of each cell line, a first culture with 80/90% confluency was counted in a Neubauer Chamber (Marienfeld) at 1:2 dilution with trypan blue and ten T25 flasks were cultured with the initial number of 20000 cells/ml each. The medium was changed every five days and every day for 10 days, at the same hour, one flask was trypsinized and the number of cells were counted as explained above, using a Neubauer Chamber. The curves were then constructed using the Log_{10} value and analyzed in Excel (Microsoft).

A typical growth curve for cultured cells should have a sigmoid pattern of proliferation and the growth phases associated are defined as: Lag Phase, an adaptation period in which cells do not divide; Log Phase, where the cells are more active in proliferation and have an exponential increase of cell density; Plateau Phase, where proliferation slows due to confluency; and Decline Phase, where cells start dying due to the end of cell cycle⁴⁶.

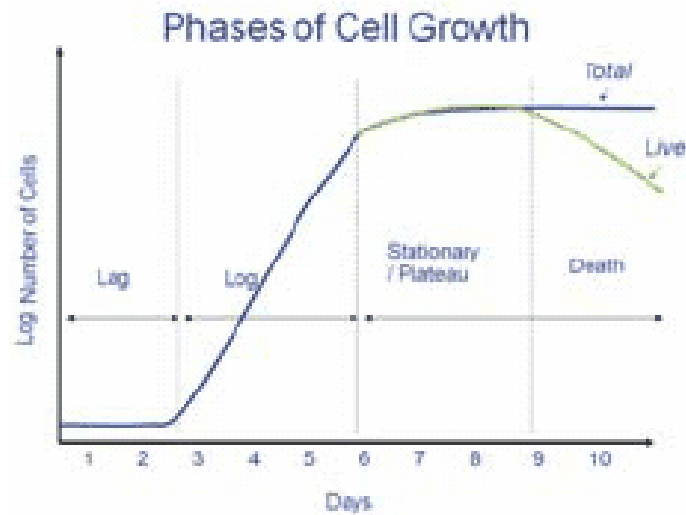


Figure 2.5 Cell Growth Graph. Cell growth curve with the respective phases of growth. In the yy axis is the Log number of cells and in the xx axis is the correspondent day. Adapted from Sigma⁴⁶

2.4 Statistical Analysis

Experimental data was analyzed using GraphPad Prism 7. Statistical differences were determined using Student's t-test, considering a p-value < 0,05 statistically significant, denoted by "**".

3 Results

In this work, we aimed to develop colorectal cancer (CRC) cell line models overexpressing both aberrant glycans and Luciferase to be used in tracking assays. For that purpose, CRC cells expressing aberrant glycans were established through overexpression of glycosyltransferases involved in the glycan biosynthesis and, transduced with *LUCIFERASE (LUC)* gene. These CRC cell line models can be used to assess the impact of glycan overexpression in cell functions and the bioluminescence produced by Luciferase enzyme can be used to track cells *in vitro* and *in vivo*. In this work, we were able to assess cell phenotype of a set of CRC cell lines. In the future, the cell line models here developed can be used to follow and evaluate tumor formation, proliferation, and metastasis development in living organisms.

In particular, the project was developed in two separated parts presented in the following sections. The first part described in section 3.1 had a specific aim to establish a CRC cell line overexpressing the sialyl-Tn (STn) antigen. The work in this section thus consisted on the genetical medication of a CRC cell line that did not expressed STn, the LS174T cell line, by transducing it with the *ST6GALNAc I* gene. This transduced cell line was phenotyped to assess the expression of STn antigen.

The second part, described in section 3.2, had a specific role to establish CRC cell lines expressing the Luciferase enzyme. The work in this section, thus consisted on the genetical modification of several aberrant glycan expressing CRC cell lines, with the *LUC* gene. The resulting transduced cells were then phenotyped to assess their glycosidic profile together with Luciferase activity.

Table 3.1 Summary table for the used cell lines

CELLS	TRANSDUCED	TRANSFECTED	NEEDED
LS174T			<i>ST6GALNAc I</i>
LS174T <i>ST6GALNAc I</i>	<i>ST6GALNAc I</i>		<i>LUCIFERASE</i>
LS174T MOCK			<i>LUCIFERASE</i>
SW948 <i>ST6GAL I</i>	<i>ST6GAL I</i>		<i>LUCIFERASE</i>
SW948 MOCK			<i>LUCIFERASE</i>
HT29 <i>FUT VI</i>		<i>FUT VI</i>	<i>LUCIFERASE</i>
HT29 MOCK			<i>LUCIFERASE</i>
SW48 <i>ST6GAL I</i>	<i>ST6GAL I</i>		<i>LUCIFERASE</i>
SW48 MOCK			<i>LUCIFERASE</i>
SW620 <i>FUT VI</i>		<i>FUT VI</i>	<i>LUCIFERASE</i>
SW620 MOCK			<i>LUCIFERASE</i>

3.1 Establishment of a CRC cell line overexpressing STn Antigen

As mentioned before, in this part of the study, we aimed to establish a CRC cell line overexpressing the STn antigen. To assess this, the LS174T cell line was transduced with the *ST6GALNaC I* gene, a flow cytometry was performed using the mouse anti-STn antibody, B72.3, as described on section 2.3.1.1. The STn antigen was expected to be expressed in the *ST6GALNaC I* transduced LS174T cells and on MDA STn⁺ cells, a previous transduced cell line expressing STn used as positive control. Although the results observed were the expected for MDA STn⁺ cells, with a 67,9% of expression (presented in red in the graph below), LS174T *ST6GALNaC I* did not show any STn expression at cell surface, since only 2,08% of the population were positive for the fluorescence read.

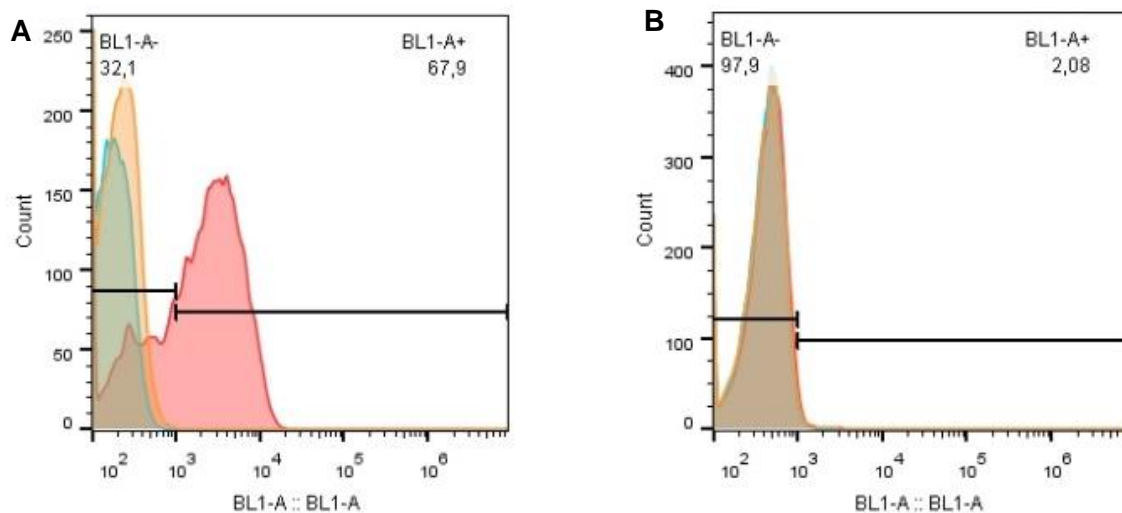


Figure 3.1 STn assessment by flow cytometry in MDA STn and LS174T ST6GalNaC I cell lines. On the histograms, the yy axis represents the number of cells and the xx axis represents the fluorescence intensity on the BL1-A channel. A) MDA STn⁺ unstained cells (blue), MDA STn⁺ cells with secondary antibody staining (orange) and MDA STn⁺ cells stained with B72.3 (red); B) LS174T ST6GalNaC I unstained cells (blue), LS174T ST6GalNaC I cells with secondary antibody staining (orange) and LS174T ST6GalNaC I cells stained with B72.3 (red);

The *ST6GALNaC I* expression in the transduced cell lines was assessed using a Real Time Polymerase Chain Reaction (RT-PCR). The expression in the LS174T cell lines transduced or not with *ST6GALNaC I* gene was compared in between and also with a positive control, the breast cancer cell line, MDA STn⁺. As shown in Figure 3.2, through RT-PCR, the expression of *ST6GALNaC I* in the LS174T *ST6GALNaC I* transduced cell line was very low, 72,42. As expected, the positive control MDA STn⁺ cell line showed a positive expression and, the negative control, LS174T WT cell line showed a negative expression.

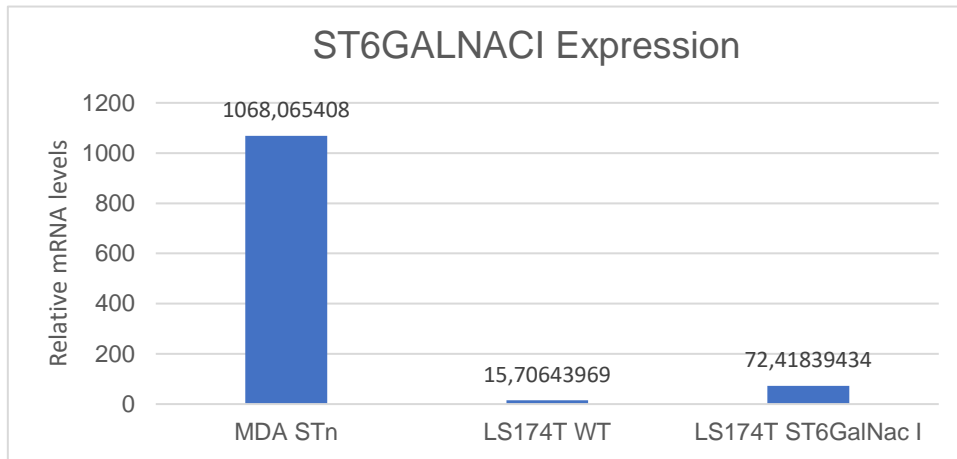


Figure 3.2 ST6GALNaC I Expression Assessment by RT-PCR. In the presented graph is shown the relative mRNA levels of the ST6GALNaC I gene. The yy axis represents the relative mRNA levels and the xx axis the cell line in study.

No STn expression does not necessarily have to be associated with ST6GalNaC 1 absence. Therefore, ST6GalNaC 1 presence was assessed. This assay was performed by Fluorescence Microscopy using an anti- ST6GalNaC 1 antibody that was recognized by a secondary, Cy5. ST6GalNaC 1 expression was observed in the Fluorescence Microscopy images on LS174T ST6GALNaC I transduced cells as presented above, by a red fluorescence around the cell nucleus (on green), being concordant with the Golgi Apparatus localization, where the glycosyltransferases are located. The controls used were LS174T WT variant with no red fluorescence and LS174T ST6GALNaC I stained with the secondary antibody only.

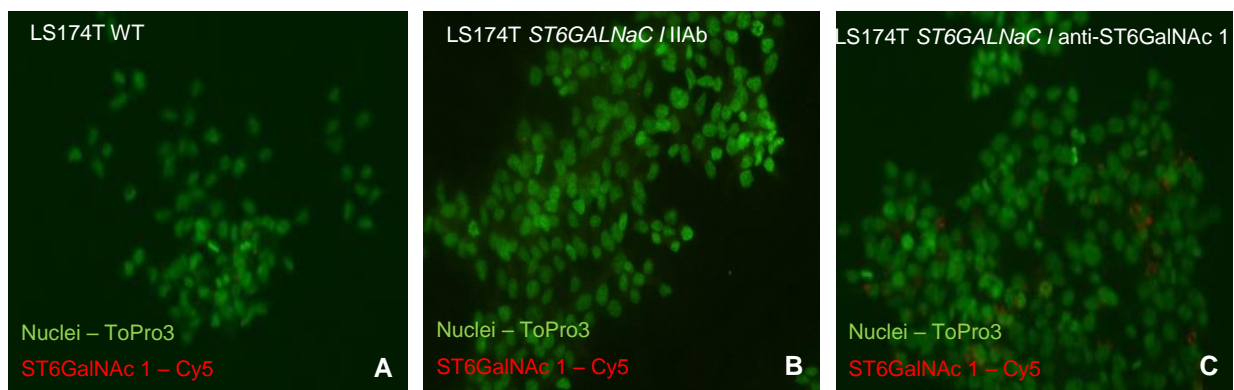


Figure 3.3 Fluorescence microscopy images of LS174T. On image A the negative control LS174T WT variant is presented, on image B there is LS174T ST6GALNaC I transduced cells stained with just secondary antibody and on C there is LS174T ST6GALNaC I cells stained with anti-ST6GALNaC I antibody. In the three images, the nuclei were detected by ToPro3 as green and ST6GalNaC 1 by Cy5 in red.

At this point the presence of the STn precursor had to be controlled and another flow cytometry was performed, using sialidase treatment to remove the sialic acid and uncover the Tn antigen. For this, the anti-Tn antibody used was 5F4. The results showed an absent expression of Tn on LS174T WT and LS174T ST6GALNaC I transduced cells, explaining why STn cannot be present in these cells, although the enzyme responsible for Tn sialylation is expressed.

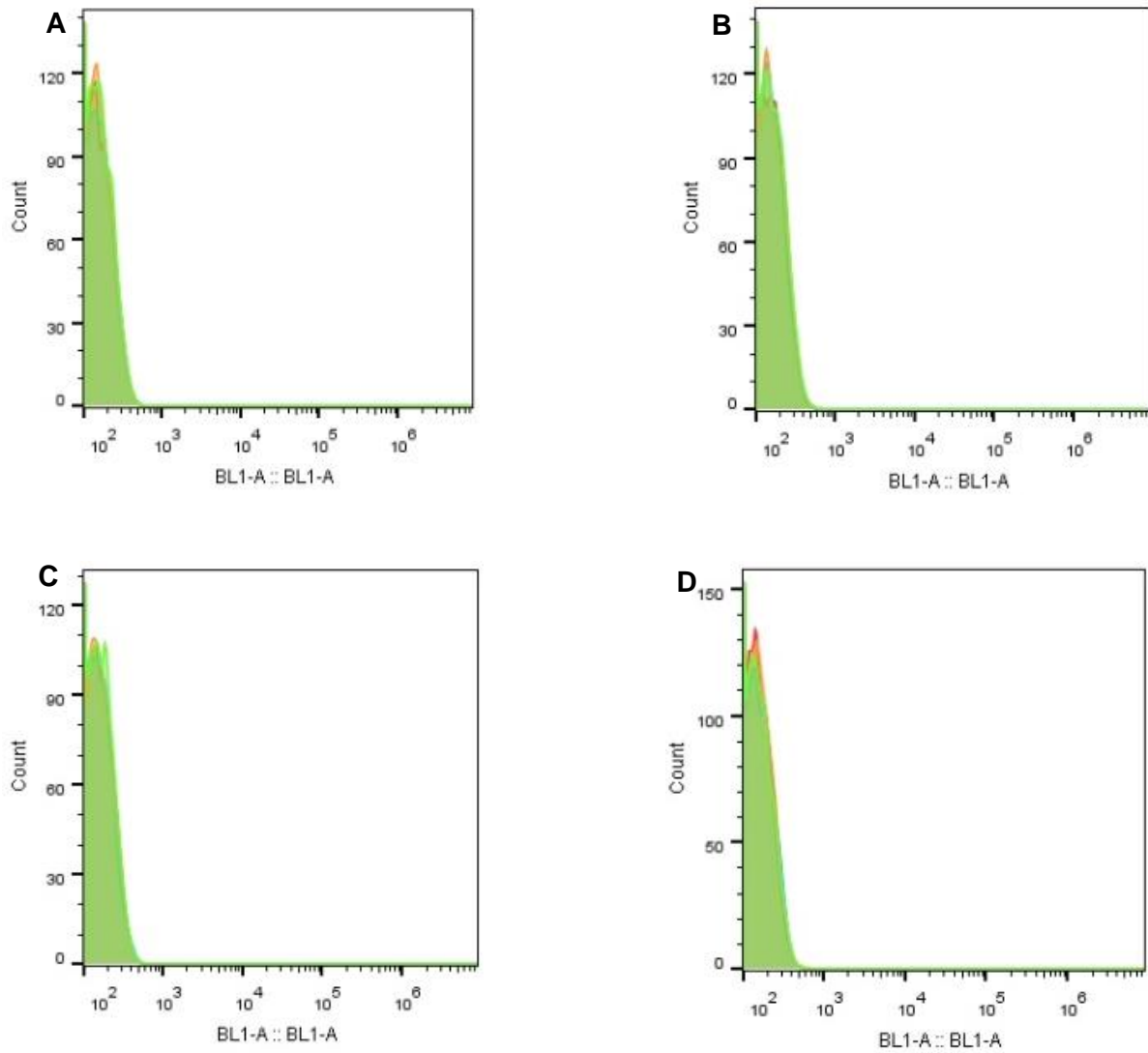


Figure 3.4 STn and Tn assessment by flow cytometry in LS174T WT and LS174T ST6GALNaC I cell lines with and without sialidase treatment. On the histograms, the yy axis represents the number of cells and the xx axis represents the fluorescence intensity on the BL1-A channel. A) LS174T WT cells with sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green); B) LS174T WT cells with no sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green); C) LS174T ST6GALNaC I cells with sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green). D) LS174T ST6GALNaC I cells with no sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green).

To check the cell line transduction efficiency, the transduction was repeated in HEK293 Star, a cell line well known to be efficiently transducible and the results observed were slightly positive for STn expression with the same anti-STn antibody used before, by flow cytometry.

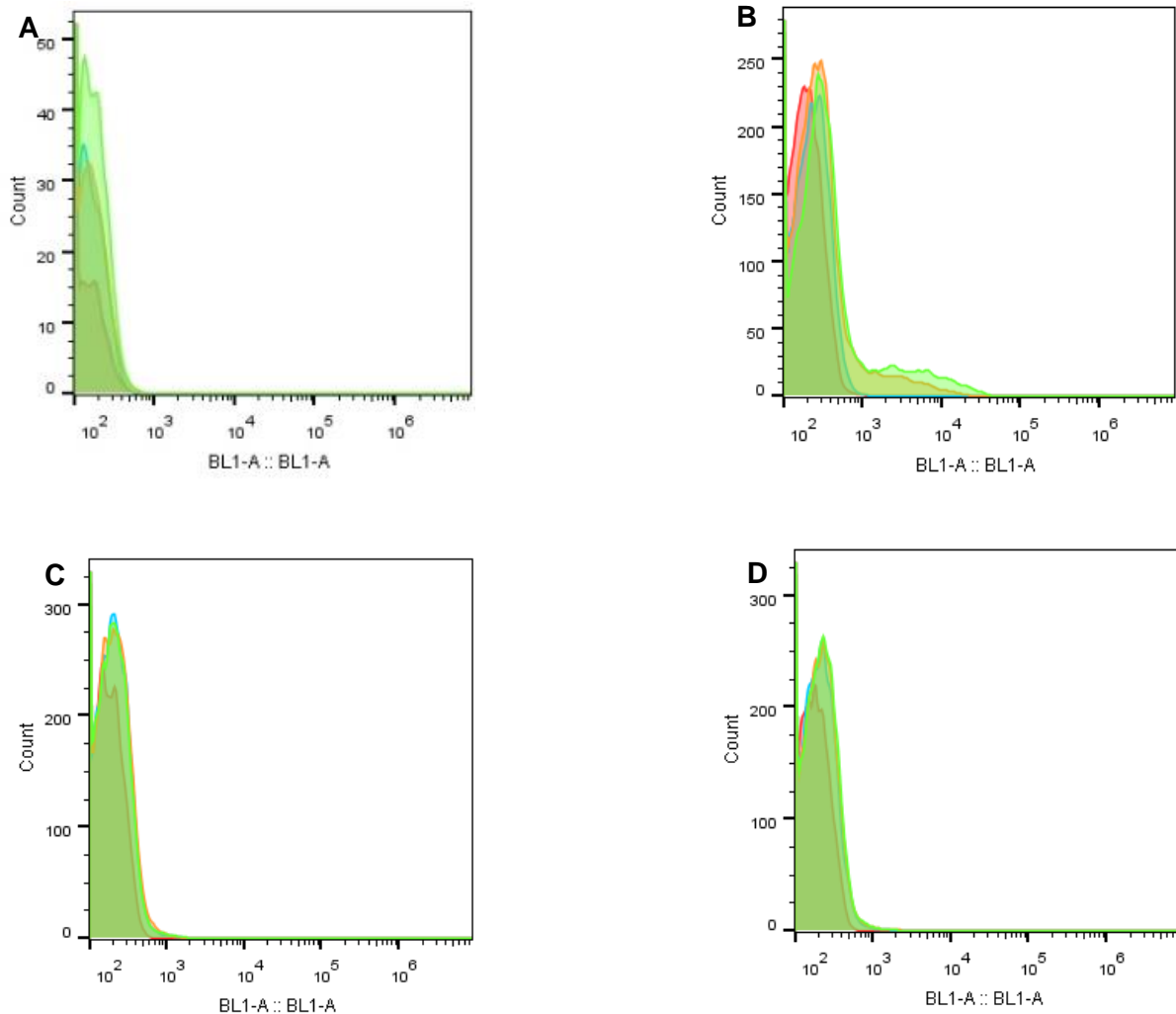


Figure 3.5 STn and Tn assessment by flow cytometry in HEK293 Star and LS174T ST6GALNaC I cell lines with and without sialidase treatment. On the histograms, the yy axis represents the number of cells and the xx axis represents the fluorescence intensity on the BL1-A channel. A) HEK293 ST6GALNaC I cells with sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 3F1 staining (green); B) HEK293 ST6GALNaC I cells with no sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green); C) LS174T ST6GALNaC I cells with sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green). D) LS174T ST6GALNaC I cells with no sialidase treatment: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange) and with 5F4 staining (green)

3.2 Luciferase Expression and Luciferase Cell Lines Characterization

After lentiviral transduction with *LUCIFERASE* gene and transduced cells respective selection methods, SW948 MOCK *LUC*, SW948 *ST6GAL I LUC*, SW48 MOCK *LUC*, SW48 *ST6GAL I LUC*, SW620 MOCK *LUC*, SW620 *FUT VI LUC* and HT29 MOCK *LUC* and HT29 *FUT VI LUC* cell lines were tested for Luciferase expression through bioluminescence emission after cell lysis and substrates addition. The

lysates were tested in Chemidoc with a 50 seconds exposition period in a 96 well plate where the cells were cultured the previous day in a concentration of 10^5 cells. As bioluminescence baseline control, a well with only luciferin reagent was added for each assay.

3.2.1 SW948 MOCK LUC Cell Line

In SW948 MOCK LUC cell line, from twenty-four isolated colonies, only four were positive for luciferase what was observed by a higher bioluminescence emission as shown in figure 3.6. After two weeks in culture, from these four colonies, only two of these maintained Luciferase expression.

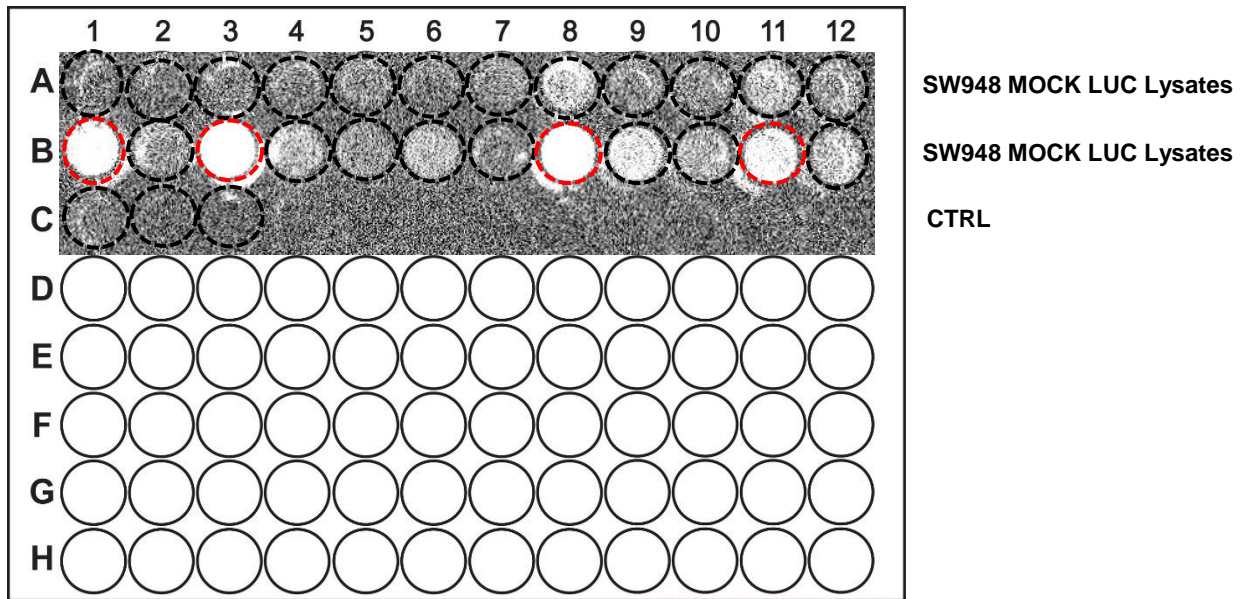


Figure 3.6 Luciferase expression in SW948 MOCK LUC transduced cells in a 96 well plate analyzed by Chemidoc. 96 well plate bioluminescence from SW948 MOCK LUC cell lysates from transduced cells with LUC gene. From well A1 to well B12, there were 24 different isolated colonies lysates from SW948 MOCK LUC cell line with luciferin reagent. The C1 well had SW948 MOCK LUC lysate without luciferin reagent, C2 had SW948 MOCK lysate with luciferin reagent and C3 had only luciferin reagent. Luciferase expression was higher in B1, B3 and B8 wells.

It was also necessary to confirm that Luciferase expression in the cells did not interfere with some tumour associated glycans that our research group is interested in. To assess it, the CRC cell lines SW948 MOCK expressing or not the LUC gene were phenotyped by flow cytometry using several antibodies reactive against glycans. For this purpose, the B72.3 antibody was used to assess STn, the CA19.9 antibody was used to assess SLe^a, the CD15s to assess SLe^x, CD15 to assess Le^x, anti-Lewis Y to assess Lewis Y, the 5F4 antibody to assess Tn, and the HECA-452 to assess SLe^{x/a}. Two proteins known to be potentially carriers of these glycans were also analysed, namely CD44 and the carcinoembryonic antigen, using the appropriate antibodies. No significant differences were observed when comparing SW948 MOCK and SW948 MOCK LUC cells as observed in the histograms above. HECA-452 and anti-CD44 stainings maintained their expression as expected.

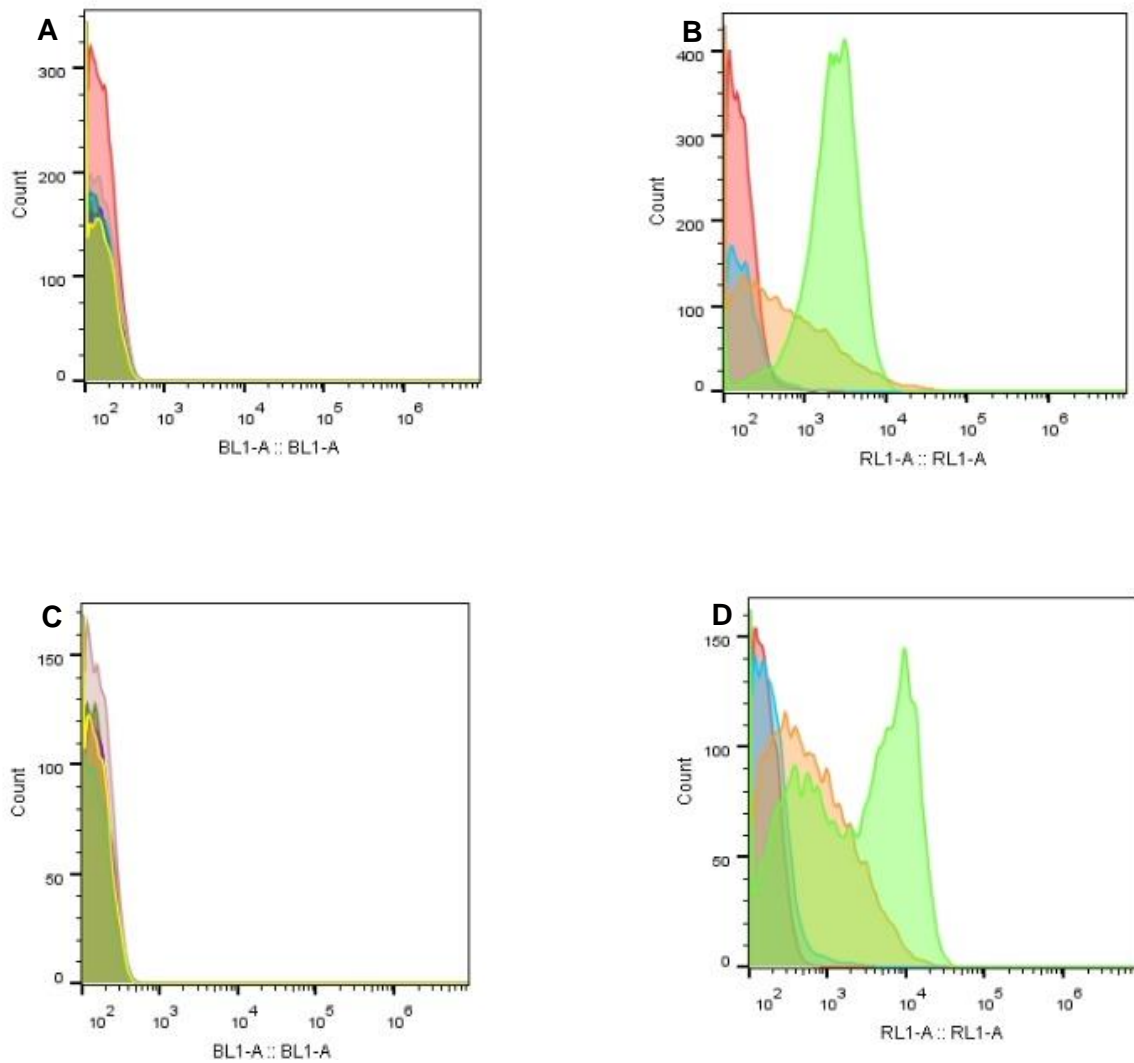


Figure 3.7 Glycan expression assessment by flow cytometry in SW948 MOCK and SW948 MOCK LUC cell lines. On the histograms, the yy axis represents the number of cells and the xx axis represents the fluorescence intensity on the BL1-A channel on the right and on RL1-A channel on the left. A) SW948 MOCK cells: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange), with CA19.9 staining (light green), with CD15s staining (dark green), with CD15 staining (light purple), with anti-CEA staining (purple), with anti-Lewis Y staining (dark blue) and with 5F4 staining (yellow); B) SW948 MOCK cells: unstained cells (red), with secondary antibody staining (blue), with HECA-452 staining (orange) and with anti-CD44 staining (green); C) SW948 MOCK LUC cells: unstained cells (red), with secondary antibody staining (blue), with B72.3 staining (orange), with CA19.9 staining (light green), with CD15s staining (dark green), with CD15 staining (light purple), with anti-CEA staining (purple), with anti-Lewis Y staining (dark blue) and with 5F4 staining (yellow); D) SW948 MOCK LUC cells: unstained cells (red), with secondary antibody staining (blue), with HECA-452 staining (orange) and with anti-CD44 staining (green);

Two ten-day cell growth curves, along with the correspondent photo report was also carried out to compare both cell lines in terms of proliferation rates and confluency. The point correspondent to Day 2 show a higher number of cells/mL in SW948 MOCK LUC cell growth curve, in comparison to SW948 MOCK. Nevertheless, no remaining differences were observed in the cell growth curves and in the photo report.

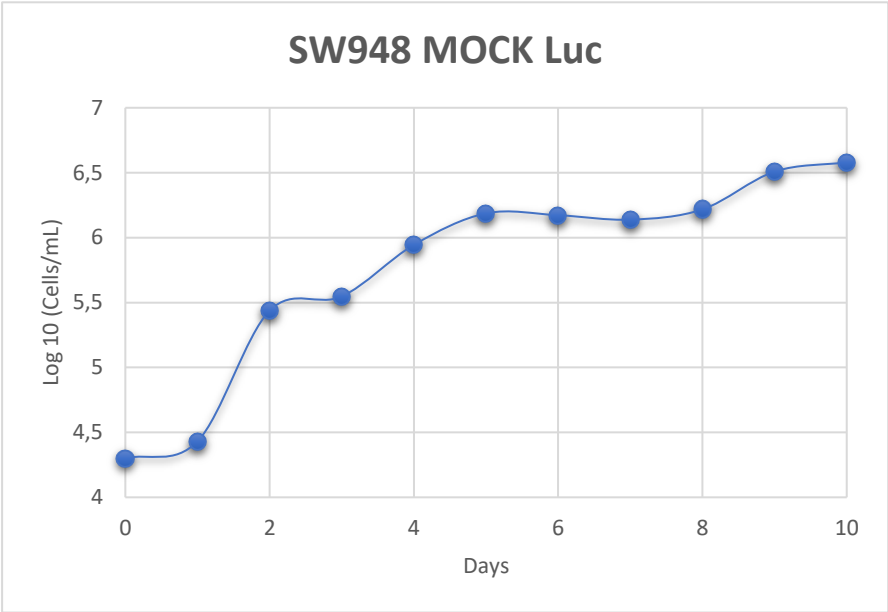
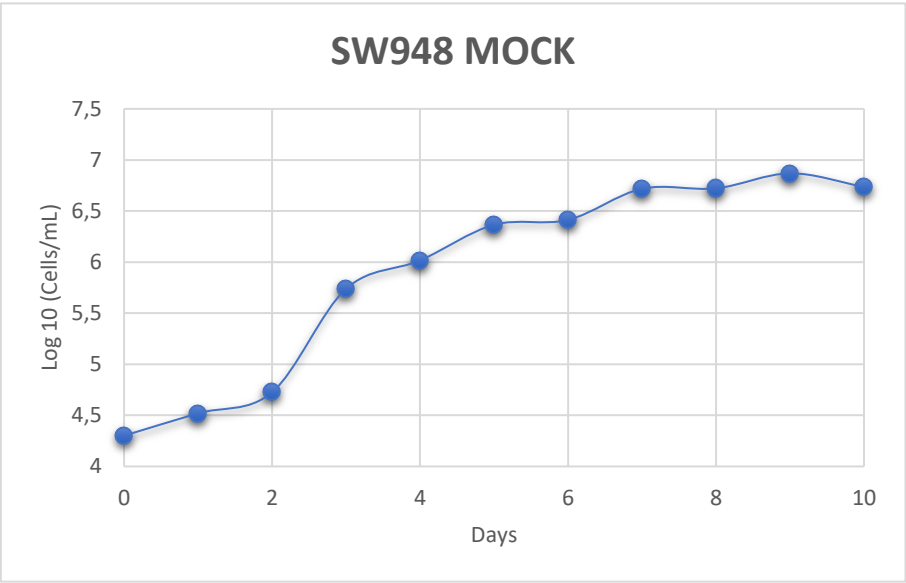
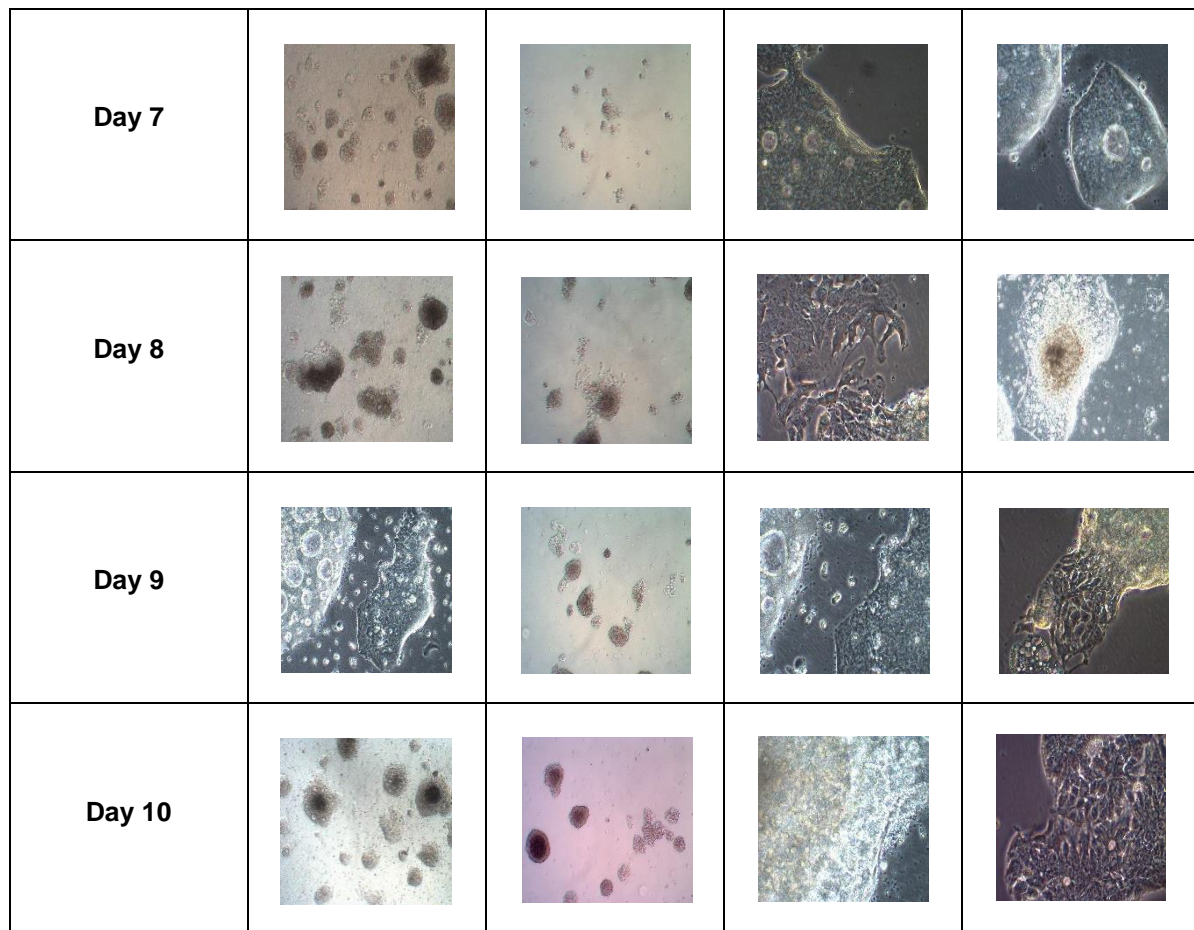


Figure 3.8 Cell Growth Curve Graphs from SW948 MOCK and SW948 MOCK LUC cell lines. The yy axis corresponds to the Log₁₀ of cell/mL and the xx axis corresponds to days. The first graph describes the cell growth curve of SW948 MOCK converted. The second graph describes the cell growth curve of SW948 MOCK LUC.

Table 3.2 Cell Morphology During Prolonged Culture. Table shows microscope images comparing the confluency and cell morphology in two cell lines the SW948 MOCK and SW948 MOCK LUC. The left columns correspond to the 40x magnification and the ones on the right the 100x magnification.

	SW948 MOCK 40x	SW948 MOCK LUC 40X	SW948 MOCK 100X	SW948 MOCK LUC 100X
Day 1				
Day 2				
Day 3				
Day 4				
Day 5				
Day 6				



3.2.2 SW620 MOCK, SW620 *FUT VI*, HT29 MOCK, HT29 *FUT VI*, SW948 *ST6GAL I*, SW48 MOCK and SW48 *ST6GAL I* Cell lines

For all the remaining cell lines, the same luciferase expression analysis was performed. For SW620 MOCK and SW620 *FUT VI LUCIFERASE* transduced, no expression was observed. To SW48 MOCK and SW48 *ST6GAL I LUCIFERASE* transduced, the same absence of Luciferase expression was observed. Despite these negative results, eight isolated colonies from SW948 *ST6GAL I LUC*, HT29 MOCK *LUC* and HT29 *FUT VI LUC* were positive for luciferase expression as presented on image 3.9. From SW948 *ST6GAL I LUC* the chosen colonies were those correspondent to wells G8, G9, H1, H8, H9 and H10.

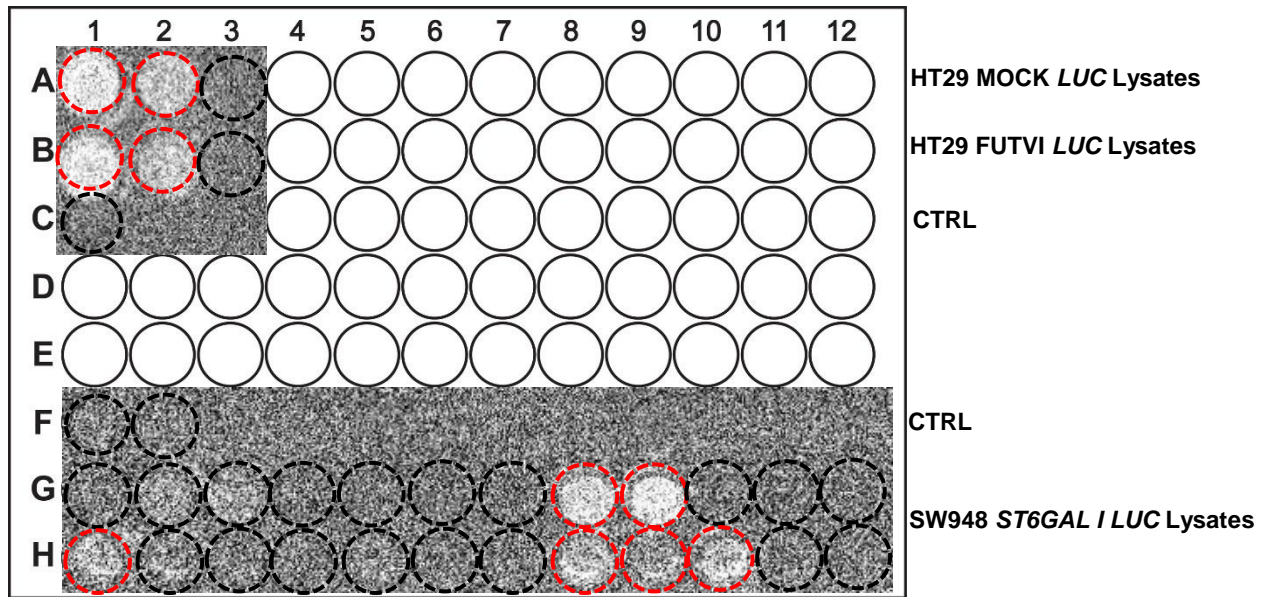


Figure 3.9 Luciferase expression in HT29 MOCK LUC, HT29 FUT VI LUC and SW948 ST6GAL I LUC transduced cells in a 96 well plate analyzed by Chemidoc. 96 well plate bioluminescence from HT29 MOCK LUC, HT29 FUT VI LUC and SW948 ST6GAL I LUC cell lysates from transduced cells with LUC gene. Wells A1 and A2 had HT29 MOCK LUC cell lysates duplicates with luciferin reagent and A3 had HT29 MOCK LUC lysates without luciferin reagent. B1 and B2 wells had HT29 FUT VI LUC cell lysates duplicates with luciferin reagent and B3 had HT29 FUT VI LUC cell lysates without luciferin reagent. The C1 well had contained only luciferin reagent. Well F1 had SW948 ST6GAL I LUC cell lysates without luciferin reagent, well F2 had only luciferin reagent and from G1 to H10 wells there were present SW948 ST6GAL I LUC cell lysates with luciferin reagent.

4 Discussion of Results

Abnormal glycosylation is a hallmark of cancer and is often due to the overexpression of glycosyltransferases involved in the glycan biosynthesis. Several glycans have already been described in the literature as overexpressed in colorectal cancer (CRC) and are associated with growth, cell-adhesion, migration, and proliferation of tumor cells. SLe^x and STn are examples of TACAs, that are overexpressed in carcinoma cells and that are highly related with poor prognosis and low survival of cancer patients.

The CRC cells are an example of cancer cells that can be decorated with this cell surface antigens thus improving its malignancy. SLe^x and SLe^a expression at the cell surface of CRC cells, have been found to be associated with disease progression. In fact, higher expression levels of SLe^x appears to occur in more metastatic colorectal carcinomas than in the less metastatic. Likewise, STn have been described to induce oncogenic features in cancer cells such as proliferation, migration and invasion, and, therefore increasing metastasis formation²⁰.

For the reasons exposed above, in this thesis, we tried to develop a group of colorectal cancer cell models expressing some of the more relevant cancer associated glycans and study their development *in vivo* and *in vitro* to better comprehend how tumors migrate and metastasize in the presence of these aberrant glycans.

To have a consistent starting point of colorectal cancer cell lines expressing the chosen TACAs, we were missing a STn expressing cell line. For this reason, this thesis project was divided into two parts both based on aberrant glycans expression by colorectal cancer cell lines. The first part consisted on genetically modifying a non-expressing STn colorectal cancer cell line, LS174T, to express STn by transduction with the gene of the responsible enzyme for STn synthesis, *ST6GALNAc I*. The second part was based on luciferase expression, to be used as a tracing mechanism by SW48, SW948, SW620 and HT29 CRC cell lines, already expressing tumor associated glycans by previous transduction and transfection with *ST6GAL I* and *FUT VI* genes, respectively, to evaluate tumor progression and metastasis formation through bioluminescence emission *in vitro* and *in vivo*.

Since overexpression of STn has been reported in several epithelial cancers, with colorectal cancer among them⁴⁷, and its expression is also associated with a poor prognosis and low patient survival, one of this work main goals was to lead to this antigen expression, for the study of metastasis formation and tumor proliferation due to its presence. The colorectal cancer cell line used was LS174T, since it has no STn expression. On the contrary of the expected, LS174T cell line transduction with *ST6GALNAc I* gene did not lead to STn antigen expression. Although the transduction worked in LS174T and in HEK293 cell lines, and a small expression of *ST6GALNAc I* was observed on LS174T *ST6GALNAc I* at an early stage, by Fluorescence Microscopy and by RT-PCR, the remaining assays to detect STn expression were not as successful (Chapter 3.1). This may have been due to several reasons, but, when the STn precursor, Tn, was assessed, Tn was absent as well, justifying why STn was not being produced, although ST6GalNAc 1 expression was observed. Even though the cause of this specific result is explained, the reason why Tn is not being expressed by this cell line is not well understood and its comprehension requires the study of all steps of the Tn synthesis pathway. Nevertheless, some

studies have determined that Tn formation is performed by T-synthase an enzyme controlled by a chaperone, Cosmc. Acquired mutations on the gene responsible for this chaperone codification, leads to a decrease on T-synthase activity and consequently lack of Tn and STn expression⁴⁸.

Regarding the second part of the project, as expected, SW948 MOCK transduced with the *LUCIFERASE* gene, expressed Luciferase by bioluminescence detection and maintained the same glycosidic profiles before, meaning that Luciferase expression by these cells, did not interfere with the cell line basal glycan expression levels. Concerning the growth rates between both cell lines, a difference was observed in the second day of the cell growth curve of the Luciferase expressing cell line, since it had a higher cell count. This may be due to counting errors or due to a more abrupt growth during log phase in the presence of Luciferase.

Relative to the other described CRC cell lines, only HT29 MOCK *LUC*, HT29 *FUT VI LUC* and SW948 *ST6GAL I LUC* were positives for Luciferase expression after transduction so, SW620 MOCK *LUC*, SW620 *FUT VI LUC*, SW48 MOCK *LUC* and SW48 *ST6GAL I LUC* were negative for bioluminescence production. This may be due to an unsuccessful transduction either due to the need of more selection time with antibiotic or a higher antibiotic concentration on SW620 cells and due to selection of the non-expressing colonies in case of SW48 cells. Another reason this may have happened, lays on the need to use a higher Multiplicity of Infection (MOI) for this cell lines or to concentrate more the virus solution. MOI indicates the number of vector particles needed, for cell, in the transduction process for the optimal transduction efficiency and must be determined by the number of viral particles divided by the number of cells seeded, that may have been too low to the cells in study. Concerning the glycan profile and proliferation study assays for the Luciferase expressing cell lines HT29 MOCK *LUC*, HT29 *FUT VI LUC* and SW948 *ST6GAL I LUC*, these experiments deferred, due to time restrictions, likewise the repetition need of *LUCIFERASE* transduction, on the negative cell lines.

5 Conclusion and Future perspectives

Cancer is one of the most prevalent diseases worldwide and metastasis are the major cause of cancer related deaths. Some carbohydrates are aberrantly expressed in cancer cells and this altered glycosylation pattern is highly associated with cancer malignancy, tumor progression and migration making tumor associated carbohydrates (TACA), promising candidates as novel cancer biomarkers and therapeutic targets. The upregulation of the enzymes responsible for their biosynthesis, namely the glycosyltransferases, is the main mechanism behind their overexpression. The Sialyl-Tn (STn) has gained particular interest as a cancer specific glycan, that has absent or minimal expression in normal healthy cells. Colorectal cancer (CRC) also shows an altered glycosylation pattern, associated with STn, but also with other TACAs such as the Sialyl Lewis X (SLe^x). Although this correlation is known, none of the proposed glycan biomarkers have already reached clinical application. This can be ascribed to the limited understanding of the molecular mechanisms and the lack of relevant *in vitro* and *in vivo* models to address them.

Therefore, the main goal of this study was to establish experimental genetic models based on CRC cell lines expressing altered glycans and to evaluate their behavior *in vitro* and *in vivo*. To assess it, several CRC cell lines were analyzed: SW48, SW948, SW620 and HT29, expressing relevant glycans by previous transduction and transfection with sialyl and fucosyltransferases, *ST6GAL I* and *FUT VI*. These cell lines were then transduced with the *LUCIFERASE* gene to enable cell tracking by bioluminescence emission. Additionally, CRC cell line LS174T was also used to conduct to STn antigen expression by transduction with *ST6GALNaC I* gene. In this last experiment, no expression of STn was obtained due to Tn precursor absence in this cell line. Regarding the Luciferase expression assays, half of the predicted cell lines expressed Luciferase after transduction, but the other half did not, which may be due to transduction ineffectiveness, what need to be repeated. Between SW948 MOCK and SW948 MOCK *LUC*, no significant differences were observed concerning glycans expression and growth rates.

In the future, the remaining assays to conclude the luciferase cell lines characterization should be addressed and the initial project of using these cell lines and evaluate their migration and proliferation *in vitro* and *in vivo* should proceed, to better understand the metastasis development process in colorectal cancer as a starting point for metastatic colorectal cancer treatment.

6 Bibliography

1. WHO | Cancer. *WHO* (2017).
2. Hanahan, D. *et al.* The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
3. Hanahan, D. & Weinberg, R. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).
4. Cooper, G. M. *The Development and Causes of Cancer*. (2000).
5. Nguyen, D. X. & Massagué, J. Genetic determinants of cancer metastasis. *Nat. Rev. Genet.* **8**, 341–352 (2007).
6. Poste, G. & Fidler, I. J. The pathogenesis of cancer metastasis. *Nature* **283**, 139–146 (1980).
7. Weiss, L., Orr, F. W. & Honn, K. V. Interactions of cancer cells with the microvasculature during metastasis^{1,2}.
8. Chambers, A. F., Groom, A. C. & MacDonald, I. C. Metastasis: Dissemination and growth of cancer cells in metastatic sites. *Nat. Rev. Cancer* **2**, 563–572 (2002).
9. Ferlay, J. *et al.* Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *Eur. J. Cancer* **49**, 1374–1403 (2013).
10. What Is Colorectal Cancer? Available at: <https://www.cancer.org/cancer/colon-rectal-cancer/about/what-is-colorectal-cancer.html>. (Accessed: 11th August 2017)
11. Varki, A. & Lowe, J. B. *Biological Roles of Glycans. Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press, 2009).
12. Ohtsubo, K. & Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **126**, 855–867 (2006).
13. Julien, S., Videira, P. A. & Delannoy, P. Sialyl-Tn in Cancer: (How) Did We Miss the Target? *Biomolecules* **2**, 435–466 (2012).
14. Keppler, O. T. UDP-GlcNAc 2-Epimerase: A Regulator of Cell Surface Sialylation. *Science* (80-.). **284**, 1372–1376 (1999).
15. Varki, A. in Degradation and turnover of glycans. *Essentials of Glycobiology*, 267–268 (Cold Springs Harbor Laboratory Press, 1999).
16. Taylor, G. Sialidases: structures, biological significance and therapeutic potential. *Curr. Opin. Struct. Biol.* **6**, 830–837 (1996).
17. Carvalho, A. S. *et al.* Differential expression of α -2,3-sialyltransferases and α -1,3/4-fucosyltransferases regulates the levels of sialyl Lewis x and sialyl Lewis y in gastrointestinal carcinoma cells. *Int. J. Biochem. Cell Biol.* **42**, 80–89 (2010).

18. Peracaula, R., Barrabés, S., Sarrats, A., Rudd, P. M. & De Llorens, R. Altered glycosylation in tumours focused to cancer diagnosis. *Dis. Markers* **25**, 207–218 (2008).
19. Feng, D., Shaikh, A. S. & Wang, F. Recent Advance in Tumor-associated Carbohydrate Antigens (TACAs)-based Antitumor Vaccines. *ACS Chem. Biol.* **11**, 850–863 (2016).
20. Munkley, J. & Jennifer. The Role of Sialyl-Tn in Cancer. *Int. J. Mol. Sci.* **17**, 275 (2016).
21. Fuster, M. M. & Esko, J. D. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat. Rev. Cancer* **5**, 526–542 (2005).
22. Hakomori, S. in 257–331 (1989). doi:10.1016/S0065-230X(08)60215-8
23. Hakomori, S. Glycosylation defining cancer malignancy: New wine in an old bottle. *Proc. Natl. Acad. Sci.* **99**, 10231–10233 (2002).
24. Kobata, A. & Amano, J. Altered glycosylation of proteins produced by malignant cells, and application for the diagnosis and immunotherapy of tumours. *Immunol. Cell Biol.* **83**, 429–439 (2005).
25. Bellis, S. L. Variant glycosylation: an underappreciated regulatory mechanism for β 1 integrins. *Biochim. Biophys. Acta - Biomembr.* **1663**, 52–60 (2004).
26. Sewell, R. *et al.* The ST6GalNAc-I Sialyltransferase Localizes throughout the Golgi and Is Responsible for the Synthesis of the Tumor-associated Sialyl-Tn O -Glycan in Human Breast Cancer. *J. Biol. Chem.* **281**, 3586–3594 (2006).
27. Yu, L.-G. The oncofetal Thomsen–Friedenreich carbohydrate antigen in cancer progression. *Glycoconj. J.* **24**, 411–420 (2007).
28. Ikehara, Y. *et al.* Cloning and expression of a human gene encoding an N-acetylgalactosamine-2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens. *Glycobiology* **9**, 1213–1224 (1999).
29. Marcos, N. T. Role of the Human ST6GalNAc-I and ST6GalNAc-II in the Synthesis of the Cancer-Associated Sialyl-Tn Antigen. *Cancer Res.* **64**, 7050–7057 (2004).
30. Marcos, N. T. ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues. *Front. Biosci.* **E3**, 1443 (2009).
31. de Wet, J. R., Wood, K. V., Helinski, D. R. & DeLuca, M. Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **82**, 7870–7873 (1985).
32. Gould, S. J. & Subramani, S. Firefly luciferase as a tool in molecular and cell biology. *Anal. Biochem.* **175**, 5–13 (1988).
33. Itzkowitz, S. H. *et al.* Expression of Tn, Sialosyl-Tn, and T Antigens in Human Colon Cancer1. *CANCER Res.* **49**, 197–204 (1989).

34. Kannagi, R. Molecular mechanism for cancer-associated induction of sialyl Lewis X and sialyl Lewis A expression—The Warburg effect revisited. *Glycoconj. J.* **20**, 353–364 (2003).
35. Kjeldsen Thomas Hirohashi Setsuo, Ogawa Tomoya, Iijima Hiroyuki and Hakomori Sen-itiroh, C. H. Preparation and Characterization of Monoclonal Antibodies Directed to the Tumor associated O Linked Sialosyl-2 →6 alpha-N-Acetylgalactosaminyl (Sialosyl-Tn) Epitope. *Cancer Res.* **Vol. 48**, 2214–2220 (1988).
36. Koszdin, K. L. & Bowen, B. R. The cloning and expression of a human α -1,3 fucosyltransferase capable of forming the E-selectin ligand. *Biochem. Biophys. Res. Commun.* **187**, 152–157 (1992).
37. Perletti, G., Osti, D., Marras, E., Tettamanti, G. & Eguileor, M. Generation of VSV-G pseudotyped lentiviral particles in 293T cells. *J. Cell. Mol. Med.* **8**, 142–143 (2004).
38. Robinson, J. P. Flow Cytometry. doi:10.1081/E-EBBE
39. Introduction to Flow Cytometry: A Learning Guide. (2002).
40. Cox, G. Biological confocal microscopy. *Mater. Today* **5**, 34–41 (2002).
41. Farrell, R. E. in *RNA Methodologies* 385–448 (Elsevier, 2010). doi:10.1016/B978-0-12-374727-3.00018-8
42. Wong, M. & Medrano, J. Real Time PCR for mRNA quantification. *Rev. Biotech.* **39**, (2005).
43. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**, 402–408 (2001).
44. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, 45e–45 (2001).
45. Promega, For, I., Of, U. S. E. & Products, E. Luciferase Assay System. *Solutions* 1–13 (2006).
46. Cell Types & Culture Characteristics | Sigma-Aldrich. Available at: <http://www.sigmaaldrich.com/technical-documents/protocols/biology/cell-types-culture.html>. (Accessed: 25th September 2017)
47. Ogata, S. *et al.* Different modes of sialyl-Tn expression during malignant transformation of human colonic mucosa. *Glycoconj. J.* **15**, 29–35 (1998).
48. Ju, T. & Cummings, R. D. A unique molecular chaperone Cosmc required for activity of the mammalian core 1 β -galactosyltransferase. *Proc. Natl. Acad. Sci.* **99**, 16613–16618 (2002).
49. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* **15**, 540–555 (2015).
50. Park, J.-J. & Lee, M. Increasing the α 2, 6 Sialylation of Glycoproteins May Contribute to Metastatic Spread and Therapeutic Resistance in Colorectal Cancer. *Gut Liver* **7**, 629–641 (2013).
51. Roy, R., Shiao, T. C. & Rittenhouse-Olson, K. Glycodendrimers: versatile tools for nanotechnology.

Brazilian J. Pharm. Sci. **49**, 85–108 (2013).

52. Biology of Extracellular Matrix.
53. Lentiviral Technology Delivery Systems Ppack Lentistarter Kit. Available at: <https://www.systembio.com/lentiviral-technology/delivery-systems/ppack/lentistarter-kit>. (Accessed: 2nd August 2017)
54. CFG Nomenclature. Available at: <http://www.functionalglycomics.org/static/consortium/Nomenclature.shtml>. (Accessed: 25th September 2017)

7 Appendixes

Appendix 1: Symbol and Text Nomenclature for Representation of Glycan Structures

(as selected by the Nomenclature Committee Consortium of Functional Glycomics, based on a version originally put forth by Stuart Kornfeld and later adapted by the editors of the textbook *Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press))

Symbol nomenclature must be convenient for the annotation of mass spectra. To this end, it was decided that:

- Each sugar type (i.e. sugars of the same mass: hexose, hexosamine and N-acetylhexosamine), should have the same symbol shape.
- Isomers of each sugar type (e.g. mannose/galactose/glucose) should be differentiated by color or by black/white/shading.
- Where possible, the same color or shading should be used for derivatives of hexose (e.g. the corresponding N-acetylhexosamine and hexosamine).
- Representing different sugars using the same shape but different orientation should be avoided so that structures can be represented either horizontally or vertically.

Colored symbols should still appear distinguishable when copied or printed in black and white.

Considering 10% of the population is color blind, the use of both red and green for the same shaped symbols should be avoided.

When desired, linkage information can be represented in text next to a line connecting the symbols (e.g. alpha4, beta4).



















	Print in color		
Hexoses: Circles			
N-Acetylhexosamines: Squares			
Hexosamines: Squares divided diagonally			
• Galactose stereochemistry: Yellow (255,255,0) with Black outline			
• Glucose stereochemistry: BLUE (0,0,250) with Black outline			
• Mannose stereochemistry: GREEN (0,200,50) with Black outline			
• Fucose: RED (250,0,0) with Black outline			
• Xylose: (5-pointed star) ORANGE (250,234,213) with Black outline			
Acidic Sugars (Diamonds)			
• NeuAc: PURPLE (200,0,200) with Black outline			
• NeuGc: LIGHT BLUE (233,255,255) with Black outline			
• KDN: GREEN (0,200,50) with Pattern & Black outline			
• GlcA: BLUE (0,0,250) /Upper segment with Black outline			
• IdoA: TAN (150,100,50) /Lower segment with Black outline			
• GalA: Yellow (255,255,0) /Left segment with Black outline			
• ManA: GREEN (0,200,50) /Right segment with Black outline			

Figure 7.1 Symbol nomenclature adopted for glycan structures.

Text nomenclature:

The Nomenclature Committee recommends a 'modified IUPAC condensed' text nomenclature which includes the anomeric carbon but not the parentheses, and which can be written in either a linear or 2D version:

- Including the anomeric carbon is important, and is likely to become increasingly more so in the future as more complicated structures are discovered.
- The presence of parentheses (which then necessitates the use of brackets to indicate branching structures) is unnecessarily cumbersome, particularly when representing the structure in 2D form.

Modified IUPAC condensed

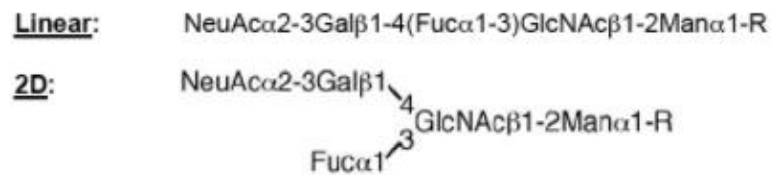


Figure 7.2 Text nomenclature adopted for glycan structures, both linear and 2D. Adapted from Nomenclature Committee Consortium of Functional Glycomics ⁵⁴

Appendix 2: Composition of the solutions and reagents used in the work developed

Complete DMEM culture medium: Simple DMEM supplemented with 10% (v/v) FBS, 2 mM L-glutamine, 100 U/ml of penicillin and 100 µg/ml of streptomycin

PBS 1x: Solution containing 1.47 mM of KH_2PO_4 , 4.29 mM of $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 137 mM of NaCl and 2.68 mM of KCl, in distilled water (pH=7.4)

Sialidase buffer (1x): 10 mM Na_2HPO_4 in MiliQ water (pH=6.0)

Appendix 3: pLenti6/V5-D-TOPO plasmid used for transducing in this work



Figure 7. *Error! Argumento desconhecido de parâmetro.* pLenti6/V5-D-TOPO plasmid used for transduction

Appendix 4: Abstract submitted to XLI Jornada da Genética

Genetic cell models to assess the role of colorectal cancer-associated glycans

C. Amaral^{1†}, R. Zoppi^{2†}, F. Deschepper³, J. Ramalho⁴, P. Videira⁵

1,2,3,5 UCIBIO, Department of Life Sciences, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, Portugal

4CEDOC, Nova Medical School from Universidade Nova de Lisboa, Portugal

†Both authors contributed equally to this work

Cancer is one of the most prevalent diseases worldwide and its metastasization is a major cause of cancer related deaths. Some carbohydrates attached to proteins or lipids, i.e. glycans, are aberrantly expressed in cancer cells and this altered glycosylation pattern is highly associated with cancer malignancy and tumor progression what makes them promising candidates as novel cancer biomarkers and therapeutic targets. Overexpression of certain glycosyltransferases (e.g. sialyltransferases and fucosyltransferases) is the main mechanism behind tumor associated glycans¹. Our group is especially interested in the role of sTn, Tn, sLex, sLea and Ley antigens as cancer specific glycans, with absent or minimal expression in normal cells².

Colorectal cancer (CRC) is the second most common cancer type and it is also correlated with an altered glycosylation pattern. Although this correlation is well known, none of the proposed glycan biomarkers has already reached clinical application. This can be ascribed to the scarce understanding of the molecular mechanisms and the lack of relevant *in vitro* and *in vivo* models to address them.

Therefore, the main goal of this study is to establish experimental models based on CRC cell lines expressing altered glycans and to evaluate their behavior *in vitro* and *in vivo*.

To assess it, we analyzed CRC cell lines, SW48, SW948, SW620 and HT29 that express relevant glycans and that were previously transfected and transduced with sialyl and fucosyltransferases, the ST6GalI responsible for increased alpha 2-6 sialylation, and FUT VI responsible for increased sialyl Lewis X/A/Y expression. These cell lines were genetically modified to express luciferase (LUC) gene as a tracking mechanism by bioluminescence.

After a successful transduction with pLenti6/V5 directional TOPOvector containing the luc gene and blasticidin selection, the luciferase expression was tested by bioluminescence and the cell lines were characterized in terms of glycosylation profiles, comparing them with their negative controls, MOCK cell lines, once previous studies showed that there is no difference between MOCK and WT.

At present, we are analyzing tumor growth of the newly transduced cells lines. In the future progression and metastasis formation in animal models, will be also assessed.

1. Trinchera M, Malagolini N, Chiricolo M, Santini D, Minni F, Caretti A, et al. The biosynthesis of the selectin-ligand sialyl Lewis x in colorectal cancer tissues is regulated by fucosyltransferase VI and can be inhibited by an RNA interference-based approach. *Int J Biochem Cell Biol* [Internet]. 2011 Jan [cited 2017 Apr 30];43(1):130–9.

2. Carrascal MA, Severino PF, Guadalupe Cabral M, Silva M, Ferreira JA, Calais F, et al. Sialyl Tn-expressing bladder cancer cells induce a tolerogenic phenotype in innate and adaptive immune cells. *Mol Oncol* [Internet]. 2014 May [cited 2017 Apr 30];8(3):753–65.