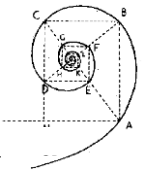
UNIVERSITÀ DEGLI STUDI DI MILANO



DOTTORATO IN MEDICINA MOLECOLARE E TRASLAZIONALE

CICLO XXIX

Anno Accademico 2015/2016

TESI DI DOTTORATO DI RICERCA

BIO10

CA19.9 and type 1 chain Lewis antigens: unraveling the molecular basis of expression in gastrointestinal tissues to improve the clinical effectiveness as tumor markers

Dottorando : Adele ARONICA

Matricola: R10709

TUTORE : prof. Riccardo GHIDONI

CO-TUTORE: dott. Marco TRINCHERA

COORDINATORE DEL DOTTORATO: Chia.mo Prof. Mario CLERICI

SOMMARIO

Introduzione

Il CA19.9 è uno dei marcatori tumorali più utilizzati sia nella ricerca che nella pratica clinica. La sigla CA19.9 identifica un gruppo eterogeneo di molecole aventi come struttura terminale il tetrasaccaride sialil Lewis a (sLea), uno degli antigeni del gruppo Lewis con catena saccaridica cosiddetta di tipo 1. Questo gruppo comprende anche il Lewis a (Lea), il Lewis b (Leb) e il disialil Lewis a che sono comunemente considerati dei prodotti fisiologici dell'epitelio del tratto gastrointestinale. Viceversa, il CA19.9 viene annoverato tra i prodotti delle cellule tumorali. Questa convinzione è dovuta principalmente a studi immunoistochimici che hanno evidenziato intense colorazioni dei tessuti tumorali utilizzando l'anticorpo anti-CA19.9 (che riconosce specificamente il sLea).

La glicosiltrasferasi B3GALT5 è uno degli enzimi necessari alla sintesi degli antigeni Lewis con catena di tipo 1. La sua trascrizione genica è sotto il controllo di vari promotori che presiedono alla sintesi di trascritti differenti nella regione 5'UTR ma identici nella sequenza codificante. Paradossalmente per la sintesi di un marcatore tumorale, questi trascritti sono silenziati nel tumore del colon. Inoltre il CA19.9 è scarsamente riconoscibile nei lisati di tumore del colon analizzati mediante dot-blot e studi condotti nel nostro laboratorio hanno dimostrato che l'anticorpo specifico per il CA19.9 reagisce per immunoistochimica colorando anche i tessuti murini che, però, mancano costitutivamente della fucosiltransferasi necessaria alla sintesi dell'epitopo e che quindi sono sicuramente privi dell'antigene.

SCOPO

Data la rilevanza che ha assunto il CA19.9 come marcatore tumorale, il nostro studio mira a rivisitare da un punto di vista istologico, biochimico e molecolare l'espressione in vivo degli antigeni Lewis al fine di appurare la loro effettiva importanza ed utilità nel cancro gastrointestinale.

MATERIALI E METODI

I livelli di espressione e la sintesi degli antigeni Lewis con catena di tipo 1 sono stati valutati nel colon e nel pancreas mediante immunofluorescenza, immunoistochimica e dot-blot; i trascritti delle glicosiltrasferasi sono stati quantificati mediante real time PCR; gli studi epigenetici sono stati realizzati mediante sequenziamento con bisolfito e saggi di immunoprecipitazione della cromatina.

RISULTATI

Il CA19.9 si è dimostrato poco visibile nella mucosa normale di colon e ancor meno nel tumore, viceversa era ben individuabile (così come il Leb) nei dotti pancreatici, sia nella parte sana che nella parte tumorale dei campioni analizzati. I trascritti della B3GALT5 sono risultati regolati negativamente nel tumore del colon, a differenza dal tumore del pancreas dove si è invece mantenuta la loro espressione. Allo stesso modo l'ST3GAL3 è risultata espressa nel pancreas in modo indipendente dalla trasformazione tumorale mentre i suoi livelli di espressione nel colon sono risultati comunque bassi. Nel tumore del pancreas il livello di metilazione delle isole di CpG fiancheggianti il promotore nativo della B3GALT5 è risultato decisamente inferiore

rispetto a quello individuato nel tumore del colon. Inoltre abbiamo individuato una regione distante circa 1 Kb dal promotore retrovirale della B3GALT5 costituita da una significativa densità di dinucleotidi CpG (stretch di CG) il cui tipo di metilazione era potenzialmente associato alla trascrizione. Questa ipotesi è stata corroborata dai risultati ottenuti mediante immunoprecipitazione della cromatina. Abbiamo infatti trovato un arricchimento degli istoni analizzati di tipo sovrapponibile a quello ottenuto con la regione di legame del fattore di trascrizione.

CONCLUSIONI

I nostri risultati dimostrano che il CA19.9 è un prodotto fisiologico la cui sintesi è fortemente regolata dalla B3GALT5 e dall'ST3GAL3. Nel pancreas sembra assumere le caratteristiche di marcatore tumorale in seguito ad ostruzione dei dotti o ad una inversione della polarità delle cellule duttali. Ciò comporta il riassorbimento dell'antigene nei vasi e un conseguente aumento dei suoi livelli sierici. Lo stesso meccanismo potrebbe anche aumentare i livelli serici degli altri antigeni Lewis, che potrebbero essere utilmente impiegati clinicamente per monitorare i pazienti di cancro pancreatico CA19.9 negativi.

ABSTRACT

Background

CA19.9 is one of the more diffusely used tumor markers in both research and clinical practice. It consists of a heterogeneous group of molecules that share the terminal sialyl Lewis a (sLea) structure, one of the so-called type 1 carbohydrate-chain Lewis antigens, which also include Lewis a (Lea), Lewis b (Leb) and disialyl-Lewis a. Lea, Leb and disialyl-Lea have been considered products of the epithelia of the normal gastrointestinal tract, while CA19.9 has been assumed as an abundant product of cancer cells, due to the reactivity found by immunohistochemical staining of cancer tissues with the anti CA19.9 antibody (recognising sLea).

B3GALT5 is one of the glycosyltransferase enzymes required for synthesizing such antigens, whose gene transcription is under the control of multiple promoters able to drive synthesis of transcripts with different 5' UTRs but identical coding sequence. Paradoxically for the biosynthesis of a tumor marker, all transcripts appear silenced in colon cancer. Moreover, CA19.9 was scarcely detectable by dot-blot staining of colon cancer lysates and we recently found a false reactivity with anti-CA19.9 antibody of mouse tissues stained by immunohistochemistry, since sLea is not synthesized in the rodents due to the lack of any α1,4 fucosyltransferase activity.

Aim

Regarding the current concept of CA19.9 as a tumor marker, our study aims at revisiting the histological, biochemical and molecular aspects of Lewis antigen expression in vivo and to assess their actual relevance in different cancers.

Methods
Expression and biosynthesis of type 1 chain Lewis antigens in the colon and the pancreas were studied by immunodetection in tissue sections and lysates, quantification of glycosyltransferase transcripts, bisulfite sequencing, and chromatin immunoprecipitation assays.

Results
CA19.9 was poorly detectable in normal colon mucosa and almost undetectable in colon cancer, while it was easily detected in the pancreatic ducts, together with Lewis b antigen, under both normal and cancer conditions. B3GALT5 transcripts were down-regulated in colon cancer, while they remained expressed in pancreatic cancer. Even ST3GAL3 transcript appeared well expressed in the pancreas but poorly in the colon, irrespective of normal or cancer conditions. CpG islands flanking B3GALT5 native promoter presented an extremely low degree of methylation in pancreatic cancer with respect to colon cancer. In a DNA region about 1 kb away from the B3GALT5 retroviral promoter, a stretch of CG dinucleotides presented a methylation pattern potentially associated with transcription. Such a DNA region and the transcription factor-binding site provided overlapping results by chromatin immunoprecipitation assays, corroborating the hypothesis.

Conclusions
Our results suggest that CA19.9 is a physiological product whose synthesis strongly depends on the tissue specific and epigenetically regulated expression of B3GALT5 and ST3GAL3. It acquires tumor marker properties in the pancreas due to duct obstruction and/or to inverted polarity of transformed ductal cells, which in turn give rise to reabsorption into vessels and elevation of circulating levels. The data also suggest that other Lewis antigens may share the same properties.

INDEX

INDEX

SIMBOLS LIST

5AZA	5’-aza-2'-deoxycytidine
B3GALT5	β1,3 galactosyltransferase enzyme 5
CEA	carcinoembryonic antigen
ChIP	chromatin immunoprecipitation
CIN	chromosomal instability
CRC	colorectal cancer
DAPI	4,6-diamidino-2- phenylindole
FAP	familial adenomatous polyposis
Fuc	fucose
FUT2	α1,2 fucosyltransferase enzyme 2
FUT3	α1,3/4 fucosyltransferase enzyme 3
Gal	Galactose
GalNAc	N-acetylgalactosamine
GlcNAc	N-acetylglucosamine
H&E	hematoxylin/eosin
HNF1	Hepatocyte nuclear factor 1
HNPCC	hereditary nonpolyposis colorectal cancer
IF	immunofluorescence
IH	immunohistochemistry
LacNAc	N-acetyllactosamine
Lacto-N-biose	Galβ1-3GlcNAc
Lea	Lewis a
Leb	Lewis b
LOH	loss of heterozygosity
LTR	long terminal repeat
MBD	Methyl-CpG binding domain
MMR	mismatch repair genes
MSI	microsatellite instability
PanIN	pancreatic intraepithelial neoplasia

PDAC	pancreatic ductal adenocarcinoma
qPCR	quantitative real-time polymerase chain reaction
RT	reverse transcription
sLea	sialyl-Lewis a
sLex	sialyl-Lewis x
ST3GAL3	galactose α2,3 sialyltransferase enzyme 3
ST6GALNAC6	N-acetylgalactosamine α2,6 sialyltransferase enzyme 6
UTR	untranslated region

FIGURES AND TABLES LIST

2. Material and methods

3. Results

1. INTRODUCTION

1.1 Gastrointestinal tumors: statistical data and most important aspects

1.1.1 Brief statistical report

Cancer represents a severe health challenge worldwide that is expected to increase due to the growth and aging of the population, as well as to the adoption of behaviors and lifestyles known to favor cancer. According to The World Health Organization, in 2012 8.2 million people died for cancer and approximately 14 million new cases appeared [1]. In addition, over the next two decades it is expected that the annual new cancer cases will rise to 22 million [1].

Gastrointestinal cancer is a collective term for cancers affecting the digestive system. This includes the cancers of the esophagus, gallbladder, bile ducts, liver, pancreas, stomach, small intestine, colon, rectum, and anus.

Despite advances over the last decade in understanding the etiology of these malignancies, gastrointestinal cancers collectively remain the most lethal malignancies worldwide [2].

In this context, colorectal and prostate cancers are the most common after lung and breast cancer (in women), with1.4 million and 1.1 million new cases diagnosed in 2012 [3], respectively (Figure 1).

Although the diagnostic and therapeutic approaches have improved in recent years, they are not yet sufficient to overcome the emerging problem of cancer, which requires new and more effective tools.

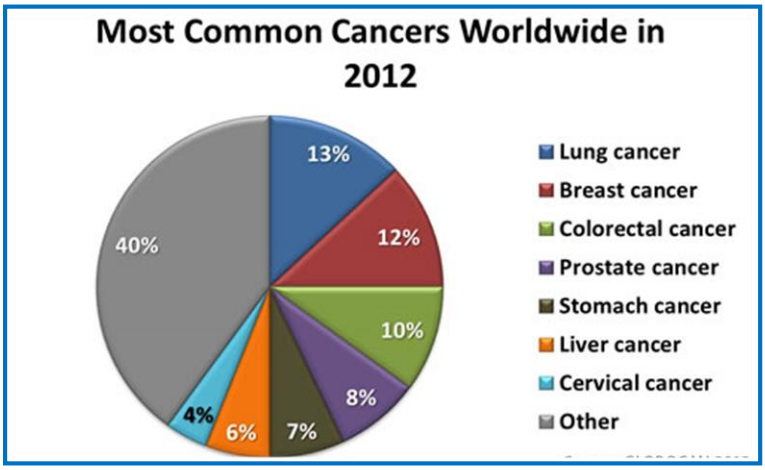


FIGURE 1: **Estimated most common cancer cases worldwide.** Colorectal cancer: 10% of all cancers diagnosed (1.4 million people). Prostate cancer: 8% of all cancers diagnosed (1.1 million people). Source GLOBOCAN 2012.

1.1.2 Colorectal cancer

Colorectal cancer (CRC) is the third most commonly diagnosed cancer in the world accounting for 10% of the estimated 14.1 million new cancer cases registered in 2012. Moreover, it is the third leading cause of cancer-related death in women and the fourth in men, with 693,600 deaths occurring worldwide in 2012 [3].

CRC is a multifactorial disease process, with etiology encompassing genetic factors. The vast majority of CRC cases have been linked to environmental causes, while approximately 30% of all CRC cases are an inherited form of the disease [4].

This tumor develops as the result of the progressive accumulation of genetic and epigenetic alterations that lead to the transformation of normal colonic epithelium into colon adenocarcinoma.

Two main genetic mechanisms play a prominent role in the occurrence of CRCs:

- **Chromosomal instability (CIN) pathway,** characterized by accumulation of mutations in tumor suppressor genes and oncogenes, as well as by other epigenetic changes such as hyper- or hypomethylation of DNA, which drive the cells to become malignant.

  The concept of the multi-step 'adenoma–carcinoma' sequence involves the inactivation of *APC* (a tumor suppressor gene) and the subsequent stepwise mutation of several other genes such as *KRAS* and *TP53* [5]. Germline *APC* mutations initiate the neoplastic process in patients with familial adenomatous polyposis (FAP) and endow all colonic crypt stem cells with a high risk of clonal proliferation [5]. Moreover, approximately 70% of colorectal carcinoma develops sporadically through the loss of heterozygosity (LOH) following inactivation of the *APC* gene [6] (Figure 2).

- **Suppressor pathway, also known as MSI (Microsatellite Instability)**, in which inactivation of the mismatch repair genes (MMR) leads to the accumulation of mutations, such as insertions or deletions, in microsatellites (repetitive DNA sequences). Microsatellite mutations may lead to genomic instability, which in turn may accelerate further accumulation of mutations in other cancer genes during tumorigenesis [7]. Microsatellite instability is involved in the genesis of about 15% of sporadic CRCs and most of HNPCC (hereditary nonpolyposis colorectal cancer) [6, 8], and may results from mutations affecting different mismatch repair genes [9]. In particular, heterozygous germ-line mutations in genes *MSH2* and *MLH1* are responsible for most HNPCC families, while *MSH6* and *PMS2* are less frequently involved [9, 10] (see Figure 3 for details).

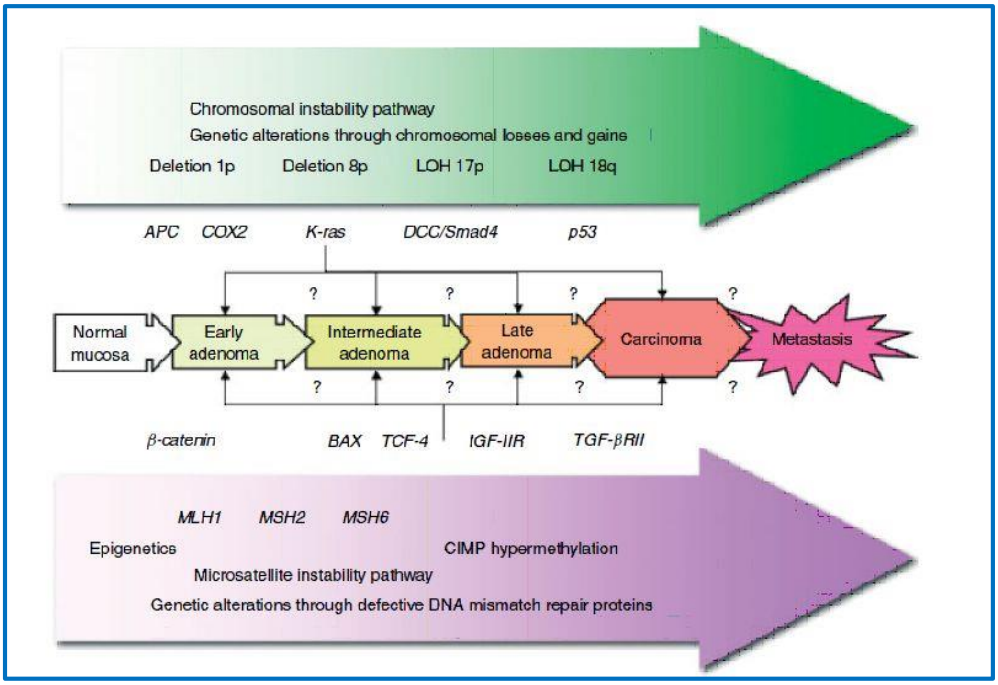


FIGURE 2: Conceptual model for colon cancer carcinogenesis. From normal epithelium to carcinoma with two molecular pathways of genomic instability: microsatellite instability (MSI) and chromosomal instability (CIN) [11].

In addition to genetic mutations, aberrant methylation of DNA is frequently found in CRCs. It includes global hypomethylation of promoters, which leads to overexpression of oncogenes and can results in chromosomal instability [12], and *MLH1* silencing due to hypermethylation of its promoter [13] (Figure 2).

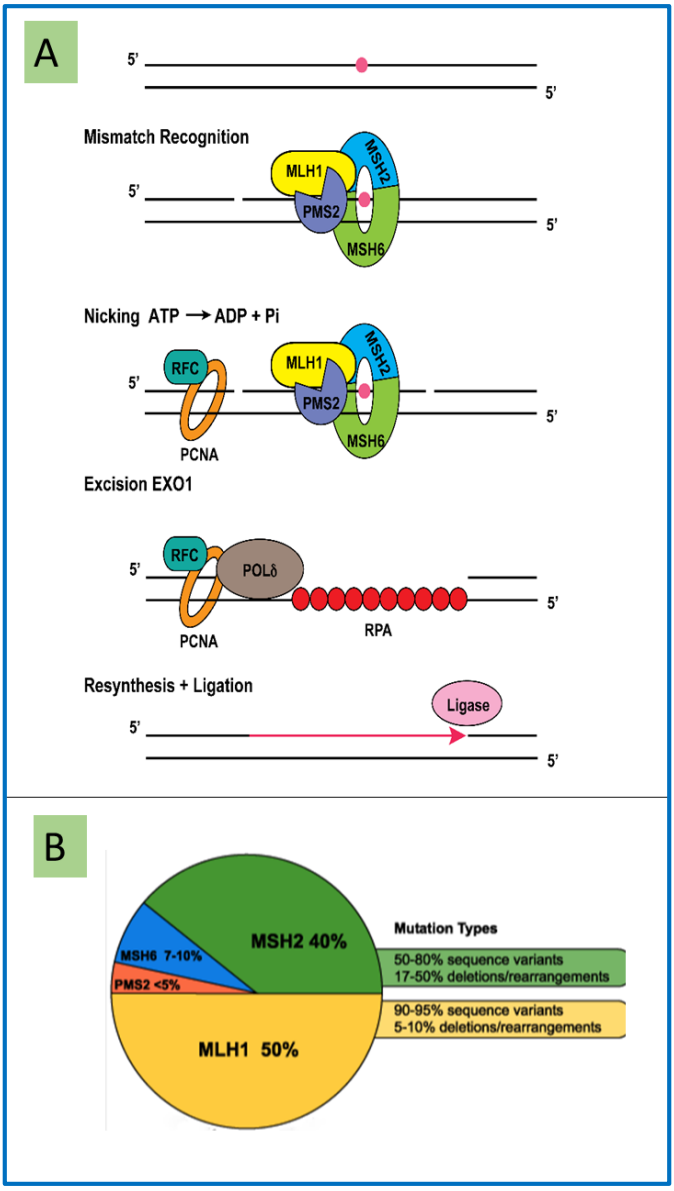


FIGURE 3: A) **The DNA mismatch repair (MMR) mechanism**. MMR proteins lead to excision of replicated DNA containing an error and to restore DNA integrity. An aberration in one of the MMR genes prevents accurate repair of base mismatches produced during DNA replication, resulting in production of a DNA chain of altered length. This phenomenon is called microsatellite instability (MSI) and can lead to an increased frequency of errors in target genes involved in carcinogenesis, resulting in cancerization of the cell. B) **HNPCC mutations**. Mutations in MLH1 and MSH2 account for about 90% of all HNPCC mutations (source: Affiliated Pathologists Medical Group, APMG).

A number of different staging criteria have been used to estimate the depth of cancer penetration in the colon as well as the extent of extra-colonic involvement. Currently, a commonly used staging method for colon cancer is based on the TNM (tumor/node/metastases) system as delineated by the American Joint Committee on Cancer (AJCC), now in its 7th edition [14]. According to this system, early stage colon cancer can be defined as a lesion that has been completely resected with no subsequent evidence of involvement of adjacent organs, lymph nodes or distant sites. While these cases are usually curable, in the advanced stage of the disease the prognosis becomes poor, due to metastatic involvement of lymph nodes or other organs [14].

Although current clinical practice in colorectal cancer screening (fecal occult blood test and colonoscopy) has contributed to a substantial rise in survival over the last two decades, the 5-year disease-specific overall survival rate is still < 60% [9]. Moreover, 70% of newly discovered CRC are detected at an advanced stage, presenting poor patient prognosis [15]. Thus, in order to eradicate CRC death, early detection is a crucial point, and identification of diagnostic markers should be urgently developed.

Currently, available serum tumor markers are carcinoembryonic antigen (CEA) and carbohydrate antigen CA19.9, but their clinical usefulness is controversial for clinical assessments in diagnostic, prognostic, and surveillance protocols [15].

1.1.2 Pancreatic cancer

Pancreatic ductal adenocarcinoma (PDAC) is the most common neoplasia of the exocrine pancreas [16]. As one of the most aggressive human malignancies, it presents an exceptionally rapid progression [9-11].

Early stage pancreatic cancer is generally clinically silent. First symptoms of pancreatic cancer become apparent only after tumor invasion of surrounding

tissues or even metastasis to distant organs, leading to a later diagnosis [17, 18]. As a result, PDAC patients have an extremely poor prognosis and the five-year survival rate is only 7% [19].

Although PDAC is the fourth leading cause of cancer death in Europe and the United States, it is not found among the most common cancers; new cases of PDAC represent, in fact, only 3% of all cancers diagnosed [3, 19]. At the same time, this tumor has been projected to become the second leading cause of cancer-related deaths by 2030 [20].

PDAC is a heterogeneous disease at the molecular, pathological, and clinical levels. According to the American Cancer Society, 10% of PDAC may be caused by an inherited DNA alteration, meaning a genetic predisposition to cancer due to family lineage. There are different identified PDAC precursor lesions that affect the pancreatic ducts [21, 22] :

- intraductal papillary mucinous neoplasm (IPMN);
- mucinous cystic neoplasm (MCN);
- pancreatic intraepithelial neoplasia (PanIN), which is the most common PDAC precursor lesion [5, 20]. These non-invasive precursor lesions reflect different grades of dysplasia in the epithelium of the pancreatic ducts, while the high-grade PanIN (PanIN-3) may be converted into invasive pancreatic ductal adenocarcinoma [23] (Figure 4).

Some of the major molecular alterations involved in PDAC carcinogenesis are known. Pancreatic cancer, like other cancers such as breast and lung cancer, results from the accumulation of gene mutations.

They include the activation of *KRAS*, *HER2/neu*, *BRCA2, AKT2*, and *BRAF* oncogenes, and the inactivation of *p16/CDKN2A*, *TP53*, and *SMAD4/DPC4* tumor suppressor genes [24, 25].

- *KRAS* (Kirsten Rat Sarcoma Viral Oncogene Homolog) oncogene mutation occurs in 95% of patients with PDAC [26]. Mutations of *KRAS*, impair intrinsic GTPase activity of the KRAS protein. This leads to activation of

KRAS downstream pathway that drives tumor progression. Mutations in *KRAS* are typically present in the PanIN lesions, being the earliest genetic alteration that promotes formation of PDAC and leading to cell proliferation, differentiation or survival.

- CDKN2A (Cyclin-Dependent Kinase Inhibitor 2A), also known as p16, is a tumor suppressor gene that inhibits cell cycle progression through the G1-S checkpoint. CDKN2A is inactivated in 90% of patients with PDAC and is found in PanIN-2 precursor lesions. Loss of function occurs by either homozygous deletion, intragenic mutation or epigenetic silencing by promoter methylation [27].
- The *TP53* protein is a fundamental regulator of genome integrity and plays an important role in regulation of cell cycle, such as G1-S checkpoint, G2-M arrest and induction of apoptosis. Loss of *TP53* has been found in approximately 50%–75% of patients with PDAC [28], whereas gene mutations are found in PanIN-3 lesions. Morton and coworkers have shown that p53 loss or mutation permits growth of *KRAS*-mutated cells and promotes progression of premalignant lesions to PDAC through failed growth arrest and senescence [29].
- SMAD4 (Mothers Against Decapentaplegic Homolog 4) deficiency can inhibit TGF-β-induced cell cycle arrest and migration, contributing to tumor formation. SMAD4 inactivation are late events and can be detected in PanIN-3 and intraductal papillary mucinous neoplasm and are mainly evident in carcinomas [27].
- *HER2/neu* is overexpressed in approximately 70% of infiltrating ductal carcinomas of the pancreas [30].
- Inactivation of *BRCA2 is* a rare event in pancreatic carcinogenesis, occurring in 7-10% of PDACs [31] only.
- The mutations in *BRAF* [32] and *AKT2* [33, 34] genes occur in 5% and 10-20% of PDAC tumors, respectively.

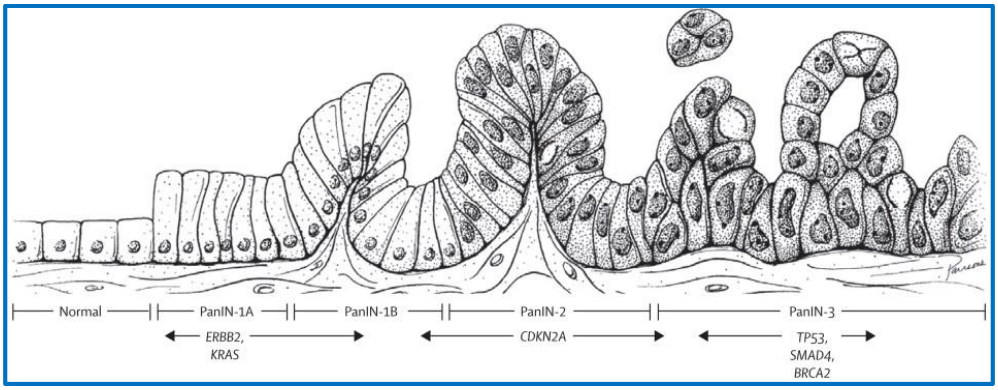


FIGURE 4: Model for the histological and genetic progression of pancreatic cancer: from normal cells through a series of histologically defined precursor lesions (PanIN) to invasive pancreatic cancer (*left* to *right)* [35].

PDAC patients are classified into different groups: resectable, borderline resectable, locally advanced, and metastatic disease. Unfortunately, the majority of PDAC patients (80-85%), are diagnosed with advanced unresectable disease [5] and only 15–20% of PDAC cases can be expected to be resectable [35].

Surgery is the only potentially curative treatment, but due to the frequently late diagnosis of PDAC (approximately 40% of patients show distant metastases and between 30–40% have locally advanced neoplasms at the time of diagnosis), surgery is not always feasible. Under these circumstances, chemotherapy and radiotherapy are the most suitable options. Despite advances in surgical approaches, post-operative survival rates have not improved significantly over the past decades.

Unfortunately, not a single tumor marker has been approved for early PDAC diagnosis. However, CA19.9 and CEA are widely used in clinical settings, to

distinguish PDAC from other cancers, monitor recurrences, and evaluate responses to therapy.

1.1.4 Tumor markers

The term “tumor markers” is commonly used to define biological molecules that indicate the presence of cancer or provide information about progression or response to therapy [36].

There are a great variety of markers, such as proteins, nucleic acids, antibodies, or peptides [37]; alternatively, a biomarker can also be understood as a collection of alterations involving gene expression, proteomic, and metabolomic signatures. These molecules can be found in the circulatory system, in excretions or in secretions.

The molecular alterations found in patients with malignancies can be due to several factors, including germ-line or somatic mutations, as well as transcriptional and posttranslational modifications [37].

The main characteristics for an ideal tumor marker should be the following:

- high sensitivity (the proportion of cancer patients identified correctly) and specificity (the proportion of non-cancer patients identified correctly).
- High positive and negative predictive value [36]. In particular, the positive predictive value is the probability that subjects with a positive test truly have the disease, while the negative predictive value is the probability that subjects with a negative test truly do not have the disease.
- Accuracy in differentiating between healthy individuals and tumor patients and between neoplastic and non-neoplastic diseases [38].
- Suitable for early detection of the disease, prediction of response or resistance to specific therapies, and for the follow-up of the patients after primary therapy [39].

Unfortunately, a tumor marker that possesses all these characteristics does not exist [36]. Despite this, in the management of patients with cancer a number of markers are still indispensable.

Over the years, many ways to classify biomarkers have been used. According to their origin, they can be classified into two different categories:

- tumor-derived markers, which are produced by neoplastic cells;
- tumor-associated antigens, which include substances produced by normal tissue in response to the presence of the neoplastic cells.

Moreover, cancer biomarker can be distinguished on the basis of their different clinical use and the main categories are listed below.

A **diagnostic tumor marker** can be used in the detection of malignant disease in a patient. If the marker is used for screening purposes, it should exhibit high levels of diagnostic sensitivity more than specificity, and to be of clinical value. A **prognostic marker** is commonly used to assess the risk of disease recurrence and/or cancer-related death for a patient, following the initial surgical removal of the cancer without administration of adjuvant therapy, and independently of the effects of future treatments. **Predictive tumor markers** are associated with either response or resistance to a specific therapy [40, 41]. Finally, **monitoring markers** are useful for the detection of recurrence or remission and are used during the follow-up of patients undergoing therapy [41].

A change in the tumor marker value during treatment is expected to reflect a change in the disease status [41]. Moreover, the early detection of recurrent/metastatic disease, immediately followed by the treatment, should enhance the chances of cure and hence of survival. However, these assumptions are not supported by current evidences and the marker values are not so obvious [36].

A typical example of monitoring marker is the CA19.9 [41] (see next paragraph).

1.1.5 CA19.9 as a tumor marker

CA19.9 is one of the more diffusely used tumor markers in both research and clinical practice. This antigen was first defined by a monoclonal antibody (NS-1116-19-9) obtained upon immunization with a membrane extract prepared form the colon carcinoma cell line SW-1116 [42]. It was later characterized as a heterogeneous group of glycoconjugates, including N- and O- linked glycoproteins, some of mucin type, and glycosphingolipds. They all share the terminal NeuAcα2,3Galβ1,3[Fucα1,4]GlcNAc sequence, named sialyl-Lewis a tetrasaccharide (sLea), representing the actual epitope [43, 44] (Figure 5).

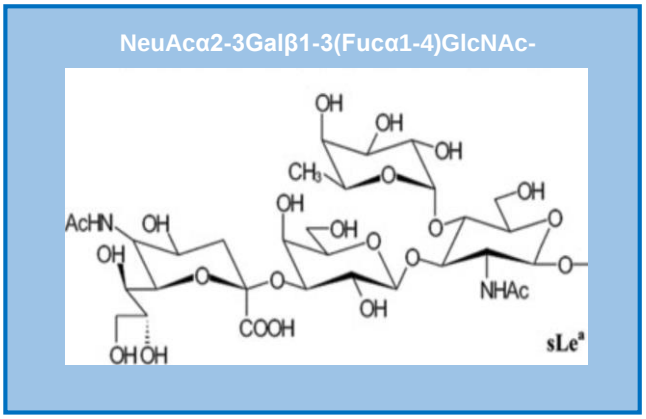


FIGURE 5: Chemical structure of sialyl Lewis a (Extracted and modified from [45]).

CA19.9 is synthesized by biliary and pancreatic ducts as well as by gastric, colonic, endometrial, and salivary normal epithelial cells [46].

The antigen is detectable in the serum and other biological fluids or extracts through RIA or ELISA procedures, and elevated serum values are associated with various pathological conditions.

Currently, CA19.9 has a primary role in the management of gastrointestinal cancers in order to monitor the response to therapy and predict recurrence

of carcinoma during follow-up [46], although its reliability is accepted for pancreatic cancers only [47]. Particularly, increased serum levels of this marker above the standard cut-off of 37 U/ml are associated with tumors of the pancreatobiliary system [48], and considered a sign of recurrence or metastatic lesions [49].

Overall, the assay (albeit not exclusive) has a reported sensitivity of 80% and specificity of 90% for the diagnosis of pancreatic carcinoma [48].

Despite its widespread use in clinical practice, the efficacy of CA19.9 as a tumor marker remain controversial and several limitations undermine is utility as a tumor marker for pancreatic cancer.

First, the diagnostic utility of CA19.9 is limited due to a low or modest sensitivity (79-81%) in symptomatic patients and a low positive predictive value (0.9%). Second, CA19.9 serum levels fail to identify early/small tumors or precancerous lesions even in individuals at high risk of developing pancreatic cancer (such as hereditary pancreatitis, family history of pancreatic cancer, or Peutz-Jeghers syndrome) (10-15% of cases). Third, CA19.9 serum levels alone cannot differentiate between benign or precursor lesions and malignant pancreatic conditions, and are not predictive of tumor location or differentiation.

It also important to remember that the use of CA19.9 is limited by false positivity (10-30%) [50] due to patients suffering benign diseases and false negativity [51, 52].

In particular, regarding the first issue, elevated CA19.9 levels can be observed in patients suffering non-malignant diseases [52, 53], such as chronic pancreatitis [54], liver cirrhosis, cholangitis [55], obstructive jaundice [46], and lung diseases [52] (Table 1).

Organ/system	Pathologic condition	CA 19-9 range (U/mL)
Pancreatic diseases	Acute pancreatitis Chronic pancreatitis Pancreatic abscess Pseudo-pancreatic cyst	**3-22**
Hepato-biliary diseases	Cholangio-carcinoma Cholangitis Choledocholithiasis Cholelithiasis Cirrhosis of liver Hepatitis Hepatocellular carcinoma Liver cyst Liver abscess Polycystic liver disease	**50-99000**
GI malignancies	Colorectal cancer Esophageal cancer Gastric cancer	**37-100**
Miscellaneous	Bronchitis Congestive heart failure Cystic fibrosis Diverticulitis Hashimoto's thyroiditis Lung cancer Ovarian cyst Pleural effusion Renal cyst Rheumatoid arthritis	**112-1338**

TABLE 1: False positive elevations of CA19.9 serum levels detected in various pathological conditions (table extracted and modified from [50]). U/mL: unit/milliliter. GI: gastrointestinal.

The second issue concerns the impossibility to detect CA19.9 in the subjects with the Lewis negative phenotype (representing about 5-10% of the population [56]) that lack the α1,4 fucosyltransferase activity required for the biosynthesis of type 1 chain Lewis antigens, including sLea (see over for details on glycosylation). Therefore, they cannot express sLea, even in presence of advanced pancreatic tumor.

Together, these critical aspects make the effectiveness of CA19.9 doubtful and bring limitations in interpreting the clinical course.

Over the last three decades, there have been many investigations aimed at developing a surrogate biomarker, but unfortunately none of these has been shown to be superior to CA19.9 in terms of sensitivity and specificity [55]. Therefore, improvements to the CA19.9 assay are clearly needed for a better management of gastrointestinal cancers.

The MEDLINE database contains over 2600 articles dealing with CA19.9, and their number per year has constantly increased from the 80s to date, reaching over 250 in 2015. Moreover, recent epidemiological survey reported that the number of serum CA19.9 determinations routinely performed in some Western country is enormous, largely exceeding the number expected on the basis of the prevalence of the diseases for which it is recommended by the scientific literature [57, 58]. In fact, medical associations provided rather restrictive recommendations, suggesting that CA19.9 should be used as a tumor marker only for the management of pancreatic cancer, but not for the early diagnosis or for other gastrointestinal cancers [55].

Notwithstanding such a big number of data, the rationale underlying the clinical use (or perhaps abuse) of CA19.9 as a tumor marker has never been addressed in detail and remains ambiguous. Is the circulating antigen produced by the cancer cells but not by their normal counterparts? Is the ability to produce CA19.9 common to cancer cells originated from different

types of gastrointestinal tissues? Why is CA19.9 a tumor marker while the other type 1 chain Lewis antigens are not?

The work presented in this thesis is aimed at answering these questions, supposed to be obvious for a tumor marker, which have never been given for CA19.9.

1.2 Abnormal glycosylation in gastrointestinal cancers

1.2.1 Glycans and glycoconjugates

Glycans (*N*- and *O*-linked to proteins or linked to sphingolipids) are structural components of the cell surface or part of secreted glycoproteins, which transmit information from the cells to the environment to control fundamental aspects of cell behavior [59], including differentiation, development, fertilization, inflammation, and cell–cell recognition [60].

Glycans can be covalently linked to proteins or lipids in the process called glycosylation, thus forming different types of glycoconjugates (Figure 6).

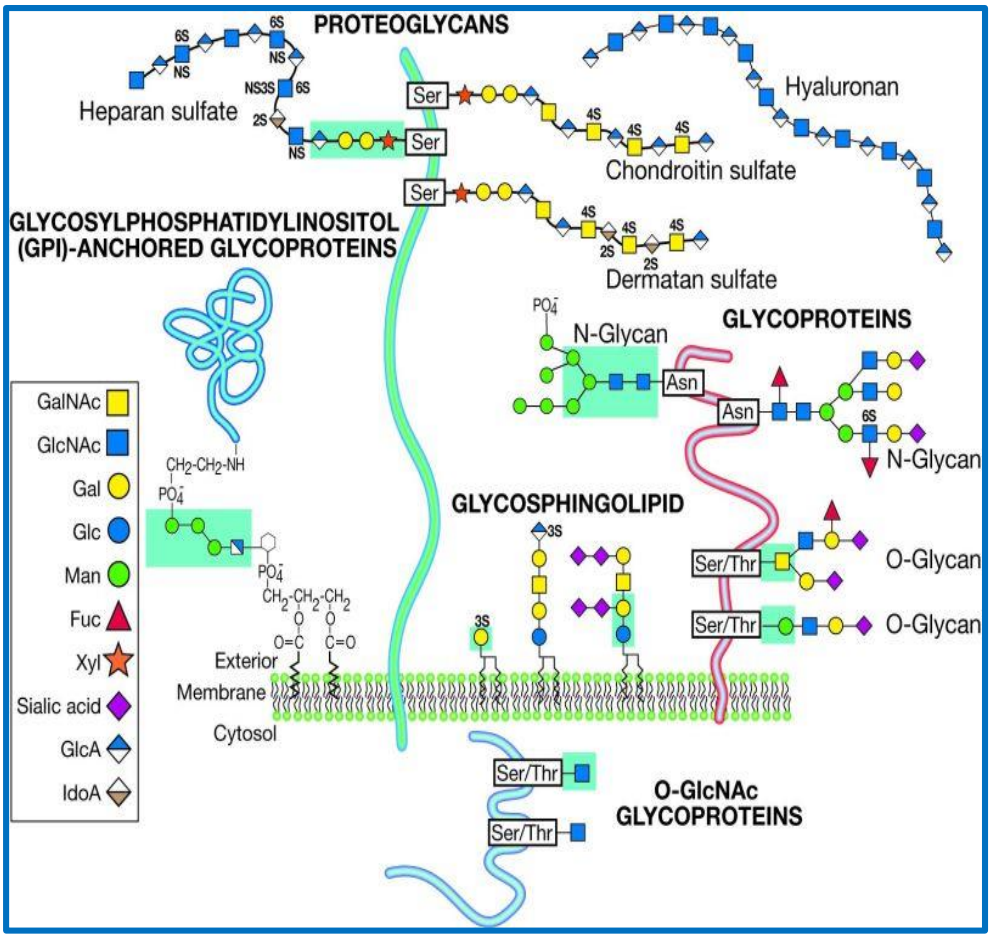


FIGURE 6 : Major classes of glycoconjugates and glycans present on the surface or inside the eukaryotic cells [61]. A glycoprotein is a glycoconjugate in which a protein carries one or more carbohydrate chains (glycans) covalently attached to a polypeptide backbone, usually via asparagine (N-linked) or serine/threonine (O-linked). A proteoglycan is a glycoconjugate that has one or more glycosaminoglycan (GAG) chains attached to a “core protein”. A glycophosphatidylinositol anchor is a glycan bridge between phosphatidylinositol and a phosphoethanolamine that is linked to a protein. A glycosphingolipid (often called a glycolipid) consists of a glycan usually attached to a ceramide, which is a lipid composed of a long chain base (sphingosine) and a fatty acid.

Glycoconjugates, such as glycoproteins, glycolipids, and proteoglycans, play a key role in regulating various cellular events and in the maintenance of homeostasis in the tissues. Their biochemical and biological properties depend strongly on the specific structure of the glycan portion. Consequently, impairment of glycan biosynthesis can have dramatic effects on the biosynthesis, function, stability and turnover of each glycoconjugate molecule [61].

In mammals, glycosylation is one of the most common and complex post-translational modifications of proteins that involves approximately 250 gene products [62]. It is affected by a multitude of factors, including the activity of the enzymes devoted to the sugar chain initiation and elongation (glycosyltransferases and some glycohydrolases), the availability of nucleotide sugars (which act as donors of the monosaccharide units in the glycosyltransferase-catalyzed reactions), and the kinetics of glycoconjugate transport [60]. Altogether, the high-coordinate network of glycosyltranferases and the competition between these enzymes for acceptor intermediates during glycan elongation [59], define the structural variability of glycans [63]. Deficiency of glycosylation enzymes or transporters results in impaired glycosylation, and consequently in the pathological modulation of many physiological processes [62], including the acquisition of malignant features by cancer cells [64]. Regarding that, various studies in the gastrointestinal tissues have identified alterations of glycosyltransferase mRNAs that lead to the biosynthesis of specific glycan epitopes potentially involved in carcinogenesis and cancer progression [64, 65].

1.2.2 Biosynthesis of type 1 chain Lewis antigens in normal and malignant gastrointestinal tissues

As mentioned above, the neoplastic transformation is often accompanied by alteration in protein and lipid glycosylation that implies changes in the expression of cell surface components, including antigens of the so called Lewis histo-blood group.

Lewis antigens are in fact oligosaccharide determinants (3-5 sugar residues) constituting the end of various carbohydrate chains of glycoproteins and glycolipids present on cell surface of epithelia, endothelia and erythrocytes [66].

Expression of these carbohydrate antigens is complex and involves coordinated activity of multiple glycosyltransferases and their regulation at the genetic, epigenetic and transcriptional level. Aberrant expression of Lewis antigens in cancer cells has been widely reported in gastrointestinal carcinogenesis.

Among the terminal or sub-terminal part of many carbohydrate chains, the presence of a galactose residue (Gal) linked to a hexosamine is very common, especially to N-acetyl-glucosamine (GlcNAc). Since the Gal unit is usually the β anomer and can be attached to either the 3– or 4- position of GlcNAc, sugar chains are frequently classified as type 1 chain, presenting the Galβ1-3GlcNAc sequence, or the type 2 chain, having the Galβ1-4GlcNAc sequence instead. Expression of type 1 chain Lewis antigens in the gastrointestinal tract requires the concurrent expression of a set of specific glycosyltransferases that interact and compete with each other to produce the final products [67].

Their biosynthesis requires the activity of a β1,3 galactosyltransferase (usually the enzyme 5, named B3GALT5), which catalyzes the addiction of the Gal residue to GlcNAc through the β1-3 linkage typical of these chains. The biosynthesis of the type 2 chains requires instead the activity of a β1-4

galactosyltransferase [68]. The Galβ1-3GlcNAc end can be further modified by an α1,4 fucosyltransferase (commonly the enzyme 3, FUT3) giving rise to the minimal trisaccharide structure (Galβ1-3[Fucα1-4]GlcNAc-R) of Lewis a antigen (Lea). The Lewis b (Leb) epitope, Fucα1-2Galβ1-3[Fucα1-4]GlcNAc-R, is formed when the Galβ1-3GlcNAc precursor is acted upon by an α1,2 fucosyltransferase (usually the enzyme 2, FUT2) forming a different substrate for the successive action of FUT3, giving rise to the final difucosylated tetrasaccharide structure [69].

The synthesis of sialyl-Lewis a (sLea), the epitope of CA19.9 antigen, requires the activity of another enzyme activity, an α2,3 sialyltransferase acting on galactose (commonly the enzyme 3, ST3GAL3) that transfers sialic acid to the Gal residue of the type 1 chain in an α2,3- linkage, resulting in the synthesis of the sialylated precursor NeuAc-α2-3Galβ1-3GlcNAc (sialyl-lacto-N-biose). Subsequently, FUT3 transfers fucose to such precursor completing the NeuAc-α2-3Galβ1-3[Fucα1-4]GlcNAc structure of sLea. As mentioned above, substitution of GlcNAc by a β1,4-linked galactose leads to N-acetyllactosamine, the basic unit of the type 2 chains. The α2,3-sialylation of the type 2 chain, followed by the addition of an α1,3-linked fucose, leads to the biosynthesis of sialyl Lewis x (sLex) antigen, an isomer of sLea of important biological functions [70] (Figure 7).

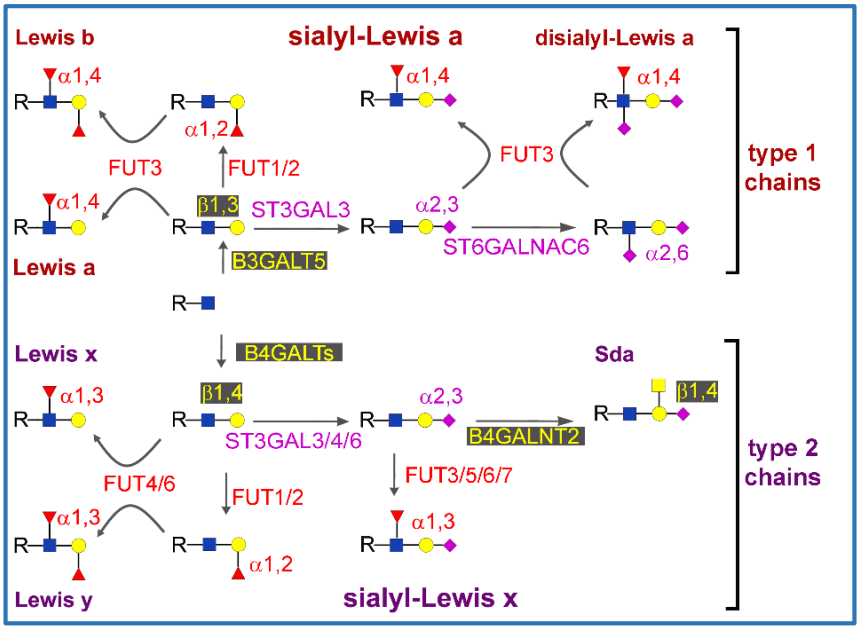


FIGURE 7: Schematic representation of Lewis blood group related antigens. The substitution of a GlcNAc residue by a β1,3-linked Gal residue, catalyzed by B3GALT5, originates Type 1 chains, whereas substitution with a β1,4-linked Gal residue forms type 2 chains. The α2,3-sialylation of type 1 or 2 chains, followed by the addition of α1,4- or α1,3-linked fucose, respectively, leads to the biosynthesis of sialyl Lewis a and sialyl Lewis x antigens, respectively. The α1,2-fucosylation of type 1 or 2 chains, followed by the addition of α1,4- or α1,3-linked fucose, respectively, leads to the biosynthesis of Lewis b and Lewis y antigens, respectively.

1.2.3 Role of glycosyltransferases in regulating expression of sLea

The role of FUT2 and FUT3 is known through in vitro studies [71, 72] and well established in vivo by the relevant effects of their polymorphisms in the human population [73, 74]. It has been amply demonstrated that there are null alleles for FUT3 and FUT2, and their genomic dosage does affect the amount of individual antigens expressed, including circulating CA19.9, as reported [67, 74]. In particular, individuals who genetically lack the FUT3 enzyme ($FUT3^{-/-}$) are known to lack any type 1 chain Lewis antigen, while $FUT3^{+/+}/FUT2^{-/-}$ individuals are candidate for higher expression of CA19.9 than $FUT3^{-/+}/FUT2^{+/+}$ individuals. The major expression of CA19.9 in subject

lacking FUT2 seems to depend on the competition for the precursor substrate between FUT2 and ST3GAL3 that initiate the synthesis of the two antigens, sLea and Leb, along mutually exclusive pathways [67]. In the colon, the expression of ST6GALNAC6 reduces the levels of sLea giving rise to disialyl-Lea [75].

While the fucosyltransferases have been extensively characterized, there are not a lot of information concerning the role of ST3GALT3 in the biosynthesis of sLea in gastrointestinal tract. Currently, few *in vitro* studies reported a preferentially action of the enzyme on type 1 disaccharides (Galβ1-3GlcNAc, the backbone of sLea antigen) [76] and a high correlation between its activity and the expression of the sLea epitope. In particular, overexpression of ST3GAL3 leads to an increase of the sLea levels, confirming its potential role in the biosynthesis of sLea antigens found on glycolipids and glycoproteins. Although it has been shown that ST3GAL3 expression correlates with tumor malignancy in several carcinomas (including those of the extrahepatic bile ducts, breast, cervix, colon, stomach and pancreas [77]), its mechanistic role has been not yet fully evaluated.

Since its cloning in 1999, B3GALT5 was proposed as the most probable candidate participating in the synthesis of the sLea epitope in gastrointestinal and pancreatic cells [78]. Overall, several studies suggest a down-regulation of the biosynthesis of the type 1 chain in colon cancer. This assumption is founded on experimental evidence found in colon cancer [79, 80], including up regulation of lactosaminic chains [81] and of their biosynthetic enzymes β1,4-galactosyltransferase I [82] and -IV [83], and down-regulation of the B3GALT5 [78, 84], indicating a switch towards the synthesis of type 2 chains in the malignant transformation of colonic tissues. Moreover, the impairment of B3GALT5 using anti-sense DNA caused a down-regulation of sLea and up-regulation of sLex and lactosaminic chains in a pancreatic cancer cell line [85]. Recent data from our lab, demonstrated that in cancer cell lines

synthesizing CA19.9, the amount of antigen secreted is proportional to that expressed on the cell surface, and depends on appreciable levels of B3GALT5. In fact, B3GALT5 is so strongly down-regulated in colon cancer that the antigen could not be detected in cancer specimens, or is more difficult to detect than in the normal mucosa. Conversely, it is widely reported that its concentration is elevated in the sera of patients affected by various cancers of the digestive organs, including the colon. Such controversial issues open questions about the biology and the origin of circulating CA19.9.

1.2.4 Implication of sLea and related Lewis antigens in tumorigenesis: past and current evidences

The notion that CA19.9 is present in various normal secretions, including seminal fluid, bile, and pancreatic juice, has been known for several years.
However, after the discovery that sLea (and sLex) act as E-selectin ligands, the interest about the expression of sialyl Lewis antigens on cancer cells increased enormously [69]. In fact, the tumor cells may employ the trafficking mechanisms utilized by leukocytes to target distant sites as a tool favoring metastasis [86]. This process includes their detachment from the primary tumor, passage into the bloodstream, and adhesion to vascular endothelia prior to extravasation [87].
Tethering and rolling require interactions between the cell adhesion molecules E-selectin and their ligands such as sLea and sLex carbohydrates [71] (Figure 8).
Thus, sLea and sLex expressed on cell surfaces could play functional roles in the adhesion of malignant cells to the endothelium facilitating the extravasation of the cells from the bloodstream and promoting cancer metastasis, mimicking neutrophils in inflammation [88-91].

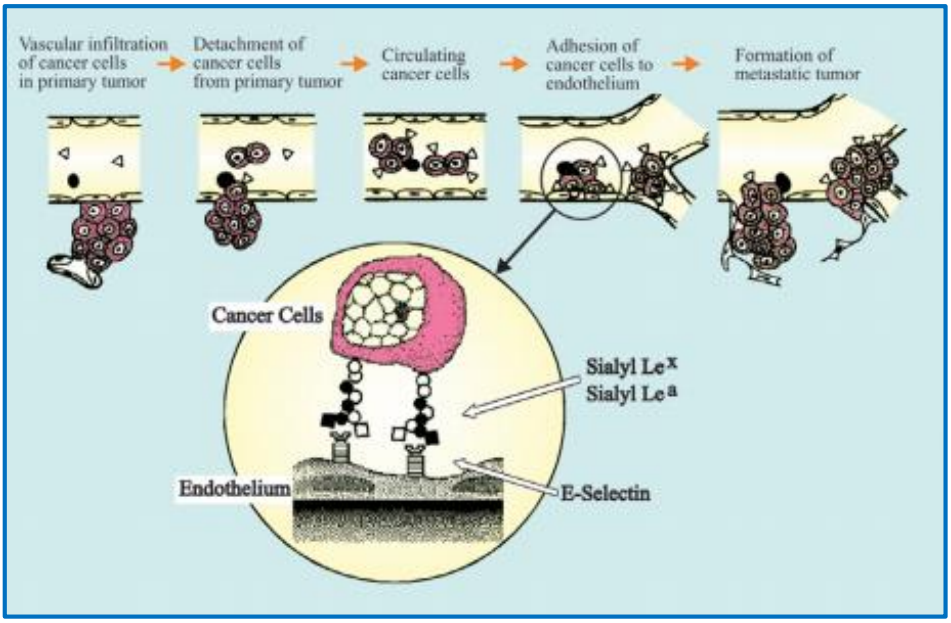


FIGURE 8: Schematic representation of the multi-step process of hematogenous metastasis of cancer [88].

The expression and even up-regulation of sLea on the cell surface of cancer cells was postulated upon the discovery of another glycosyltransferase, the N-acetylgalactosamine α 2,6 sialyltransferase enzyme 6 (ST6GALNAC6), able to form disialyl-Lewis a and silenced in colon cancer [75, 92] (Figure 9). According to studies published about a decade ago, impairment of 2-6 sialylation at the GlcNAc moiety occurs during malignant transformation of colonic epithelial cells and leads to the loss of the disialyl Lewis a determinant, causing a concurrent increase of the sLea determinant in colon cancer cells [75]. As hypothesized, the disialyl Lewis a was expressed preferentially in the normal epithelium of the colon, mediating normal interaction with intramucosal lymphoid cells, while sLea was considered a product of the tumor cells potentially involved in hematogenous metastasis.

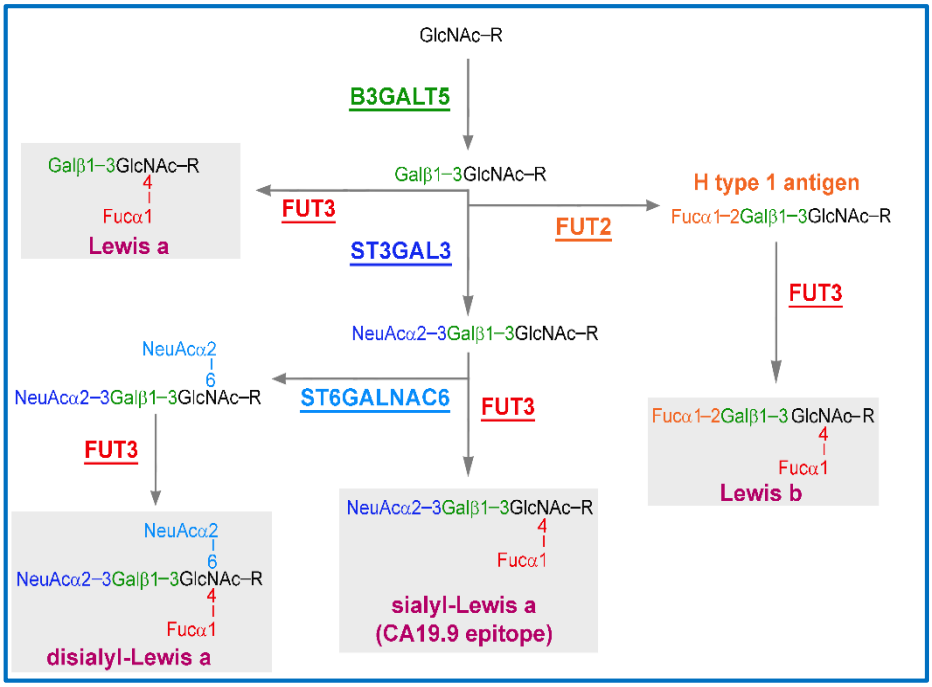


FIGURE 9: Biosynthesis of type 1 chain Lewis antigens in epithelial cells of the gastrointestinal tract. Enzymes (underlined) are named according to the HUGO gene nomenclature committee.

Coherently with previous findings, immunohistochemical studies revealed an abundant reactivity with the anti-CA19.9 antibody in colon and pancreatic cancer tissues [93, 94]. These observations lead to the speculation that sLea and cognate CA19.9 antigen were a rather specific product of cancer cells [93-96]. Otherwise, Lea and Leb have always been considered products of normal epithelia of the gastrointestinal tract [97].

However, a finding contradicting such view was reported in a study aimed at elucidating the malignant properties of colon cancer cells expressing sLea [66]. Analyzing xenografts obtained in nude mice using anti CA19.9 antibody by immunohistochemistry (IH), it was found that mouse tissues were stained although they lack any type 1 chain Lewis antigen because no α1,4 fucosyltransferase is present in mice [98]. These observations suggest that

the1116-NS-19-9 antibody does not recognize the antigen by IH, but other unknown structures that could be expressed in normal mouse tissue as well as in normal and cancerous human tissues. As a very relevant consequence, it appeared possible that the cellular and tissue distribution of the antigen detected so far by IH could be unreliable.

1.3 DNA Methylation and Cancer

1.3.1 Epigenetics and glycosylation

The term 'epigenetics' literally means 'outside conventional genetics' and is commonly used to describe the regulation of gene expression by mechanisms that do not involve changes in the DNA sequence [99].

The epigenetic regulation is involved in various physiological functions including cell division, differentiation, and apoptosis [100], thus it is necessary that the whole epigenetic machinery function properly. The epigenetic mechanisms are regulated in such a sophisticated fashion that disruption of the balance between genetic and epigenetic information, either at the genome scale or even restricted to a certain gene, is usually the cause of many diseases, including cancer [101, 102]. The epigenetic alterations frequently found in cancer include aberrant DNA methylation, abnormal histone modifications, and altered expression levels of various noncoding RNAs [103]. The very first case was documented in 1983 when researchers found an aberrant level of DNA methylation in the tissues of colorectal cancer patients [104]. Since then, a growing number of studies have demonstrated the contribution of abnormal DNA methylation to human diseases.

A variety of glycogenes is regulated by epigenetic mechanisms [105], but at present, there have been few studies addressing the epigenetic mechanisms of glycosyltransferase expression. They include the fucosyltransferase FUT3 [106], the sialyltransferases ST3GAL6 [107] and the galactosyltransferases B4GALT1 [108] and B3GALT5 [109].

1.3.2 Gene expression regulated through DNA methylation

DNA methylation is a fundamental epigenetic modification in mammalian cells and occurs at the 5 position of the cytosine rings in the dinucleotide sequence CG in vertebrates [110].

Until recently, the majority of DNA methylation studies focused on the analysis of CpG islands associated to promoter regions. There are approximately 20,000 CpG islands in the mammalian genome [111] representing 5% of all CGs and 1% of the genome. CpG islands usually consist of stretches of DNA that have very high frequencies of CGs dinucleotide repeats, occupying approximately 60% of human gene promoters [110].

Genome wide DNA methylation profiling studies show that there is a strict correlation between promoter DNA methylation and promoter activity depending on the CpG density of the promoter [112] (Figure 10). According to these recent evidences, high CpG content promoters (HCPs) are globally unmethylated or hypomethylated even when the promoter is inactive [112]. HCPs are often associated with housekeeping genes, and genes with more complex expression patterns such as those expressed during embryonic development [113]. The expression pattern of genes with intermediate CpG content promoters (ICPs) is commonly inversely correlated to the degree of methylation [112]. In contrast, it seems that DNA methylation at low CpG-density promoters (LCP) is not correlated with transcriptional silencing, and indeed most LCPs are methylated irrespective of their expression state [114], suggesting that other mechanisms regulate the activity of LCPs. Highly tissue-specific genes are considered to belong to this class of promoters [113]. Overall, these results indicate that HCPs remain unmethylated regardless of transcriptional activity, ICPs are repressed when methylated and LCPs are usually methylated.

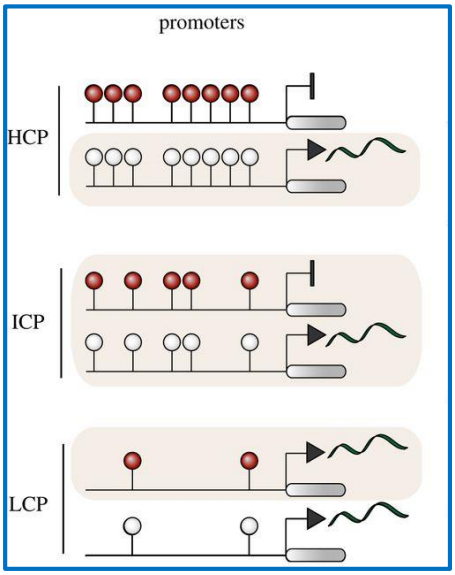


FIGURE 10: Distribution and roles of DNA methylation. The distribution of DNA methylation varies according to genomic landmarks. High-CpG density promoters (HCP) are usually hypomethylated, intermediate CpG-density promoters (ICP) are usually methylated, and low CpG-density promoters (LCP) can be either methylated or unmethylated (the methylation does not correlate with repression).
Figure extracted and modified from [115].

The best-studied epigenetic alteration in cancer is the methylation that occur within CpG islands present in 70% of all mammalian promoters [116]. Experimental evidences have shown that about 5%–10% of unmethylated CpG islands in a promoter becomes abnormally methylated in several cancer genomes [116], and many studies have investigated the mechanisms that underpin this relationship. Two main models have emerged (Figure 11). In one model, methylation causes a physical hindrance of transcription factor binding to their recognition motifs, inhibiting transcription. A second model involves the recruitment of transcriptional corepressor complexes mediated by methyl CpG-binding domain protein (MBD) that recognizes methylated DNA [117]. In particular, binding of MBD protein to the methylated DNA may physically block the binding of transcription factors to the promoter region, repressing transcription initiation. Alternatively, binding of MBD protein to the methyl CpG regions leads to a cascade of downstream events, including

recruiting of repressors and other proteins that mediate histone modifications and induce chromatin remodeling, blocking the access of transcription factors to the promoter region [118].

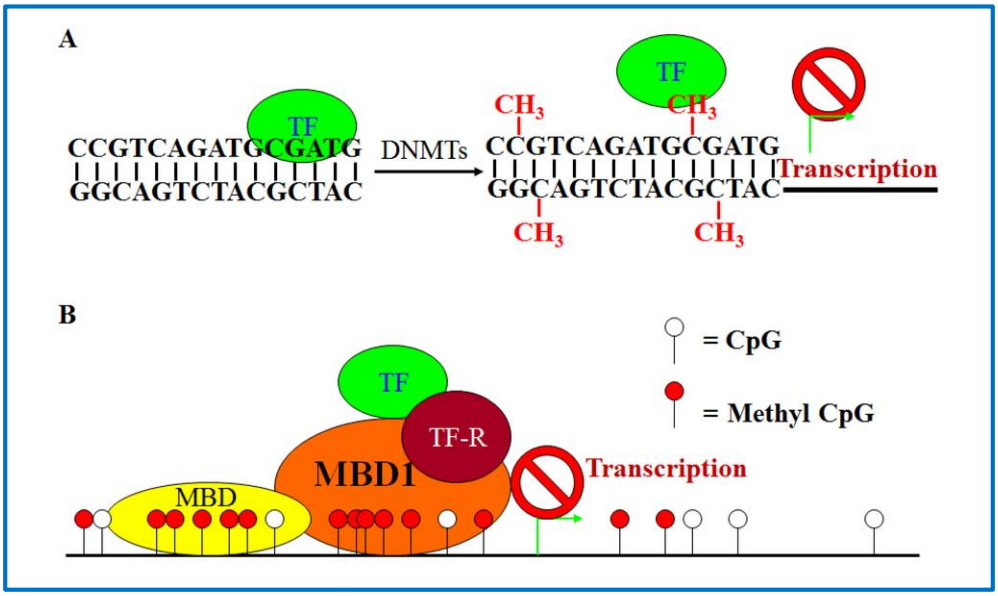


FIGURE 11: Model of DNA methylation-mediated gene silencing [118]. (**A**) Methylation of the cytosine of the CpG dinucleotides in a gene promoter region within a transcription factor's binding consensus sequence motif may be recognized by the transcription factor (TF) as a "mutation" and thus results in loss of TF binding and transcription initiation. (**B**) Methyl-CpGs may recruit MBD protein binding to the methylated DNA region in a gene promoter region. The MBD proteins may physically block TF binding to the promoter region or recruit repressors to inhibit TF activity and thereby silencing gene transcription.

While the strong association between CpG island hypermethylation and gene silencing in cancer has been widely documented and characterized, the role of the hypomethylation of the other CpG dinucleotides in cancer has not yet been established well.

However, the general dogma that DNA methylation equates with transcriptional silencing has recently been questioned by genomewide DNA methylome studies. In fact, these studies have shown that many

transcriptionally active genes present high levels of DNA methylation within the gene body and at the so called CpG island ‘‘shores” [116].

The term CpG island shores refers to regions of lower CpG density that lie in close proximity (~2 kb) of CpG islands [119] and are reported to be methylated in a tissue-specific manner [120, 121].

Overall, although DNA methylation is typically considered to be involved in silencing mechanism, tissue specific low CG content promoters are methylated irrespective of their expression. This aspect can be explained by two mechanisms. One possibility is that the CG methylation of such promoters creates transcription factor binding sites, the other is that low CG content promoters can be associated with some distal enhancer containing high CG regions. Differential methylation of these enhancer regions might be the regulator for the transcription of these genes [114].

1.3.3 Gene body methylation

Gene body methylation is common in ubiquitously expressed genes and is positively correlated with gene expression [122]. It has been proposed that it might be related to elongation efficiency and prevention of spurious initiations of transcription. In general, DNA methylation is thought to block transcription initiation but not elongation. In fact, intragenic nucleosomes with trimethylation of H3K36, which is associated with transcript elongation, seems to recruit DNA methyltransferases, thus facilitating the methylation of intragenic DNA [123]. Despite alternative promoters are classically located upstream of the translation start site, some of them are also present within genes. Therefore, it seems that a major function of the tissue-specific and conserved intragenic methylation may be to regulate the activity of such alternative promoters and to interfere with expression of the main transcripts [124, 125]. An interesting study observed a context-dependent correlation of CpG gene-body methylation, related to whether the CpGs were located

inside or outside intragenic CpG islands. For CpGs outside intragenic CpG islands, the methylation status correlated positively with gene expression, while for CpGs located within CpG islands, the methylation status can be either positively or negatively correlated with gene expression levels [126].

1.3.4 Collaborative model for DNA methylation

Recently, a new model has been proposed for DNA methylation, named the “collaborative” model. This innovative model is based on the collaboration between CpG sites, so that the methylation reactions at one site are affected by the methylation status of the nearby CpGs [127]. This can be achieved by means of systems with positive feedback. Thus, the probability that a CpG becomes or remains methylated increases if other CpGs in the cluster are methylated. Conversely, the probability that a CpG becomes or remains unmethylated increases if the CpGs in the cluster are unmethylated. In the collaborative model, the generated systems are sensitive to CpG density [127] and to the separation between CpGs [128] (Figure 12). The bias toward hyper-methylation for less dense CpG clusters suggests that collaborative methylation reactions generally act more efficiently than collaborative demethylation reactions over longer CpG–CpG distances. Conversely, the bias toward hypo-methylation for more dense CpG clusters suggests that collaborative demethylation reactions are favored at shorter CpG–CpG distances [128].

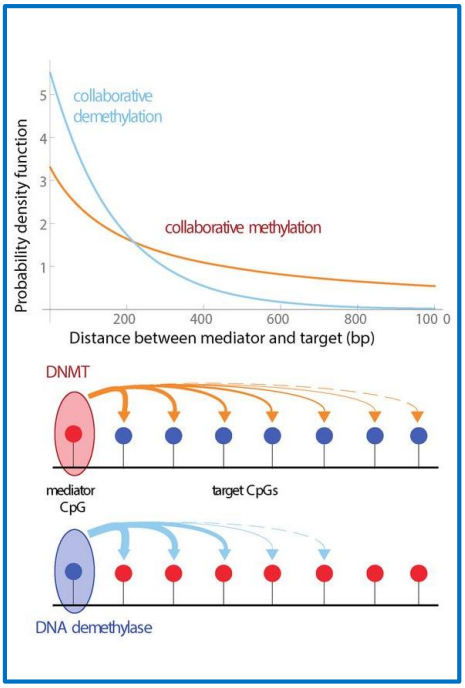


FIGURE 12: Collaborative model design. The probability of interaction between a target CpG and an enzyme (methylase or demethylase) recruited to a mediator CpG is dependent on the DNA distance between the two CpG sites. Moreover, hyper-methylated CpGs (low density of CGs) are associated with a large zone of increased methylation, while hypo-methylated CpGs (high density of CGs) seem to have effects over short distances only. Thus, the collaborative methylation is more efficiently over long distance than collaborative demethylation. The low CpG density outside islands is likely to strongly disfavor the collaborative demethylation reactions that are thus restricted to CpG islands.
Figure extracted and modified from [128].

1.3.5 B3GALT5: expression mechanisms

In 1999, the B3GALT5 cDNA was first cloned from the human colon adenocarcinoma cell line COLO-205 by Isshiki *et al.*

Studies on the B3GALT5 gene showed that at least two main promoters drive expression of alternative transcripts differing in their 5' untranslated region (UTR) [129]. In mammary gland, thymus, and trachea, as well as in some human cancer cell lines, transcription is mainly driven by the **native** promoter, which is conserved through mammalian evolution. It is sensitive to the ubiquitous transcription factor NF-Y [130] and is epigenetically regulated through the methylation of two flanking CpG islands [131] (Figure 13).

In the organs of the gastrointestinal tract (such as colon, stomach and pancreas), is active another promoter, named **LTR**, which represents a long terminal repeat of retroviral origin inserted about 25–30 million years ago

(Figure 13). This alternative promoter is present only in primates following the division of Old World monkeys [132]. It is stronger than the native promoter and is sensitive to HNF1α/β transcription factors [79], which are necessary but not sufficient to drive expression of the transcript [79, 129].

A complex interplay of epigenetic mechanisms modulates the expression of *B3GALT5* from both promoters, leading to a strong downregulation in colon cancer. The methylation status of both CpG islands flanking the native promoter is increased in cancer, leading to the inhibition of B3GALT5 expression [131]. Conversely, Zulueta and coworkers have recently shown that treatment with a demethylating agent inhibits the B3GALT5 LTR promoter, reproducing *in vitro* the down-regulation of the transcript observed in cell lines and colon cancer specimens [133]. Therefore, although HNF1-α/β transcription factors are necessary to activate the B3GALT5 LTR, transcription modulation seems to depend on unknown regulatory elements that are active when methylated [133].

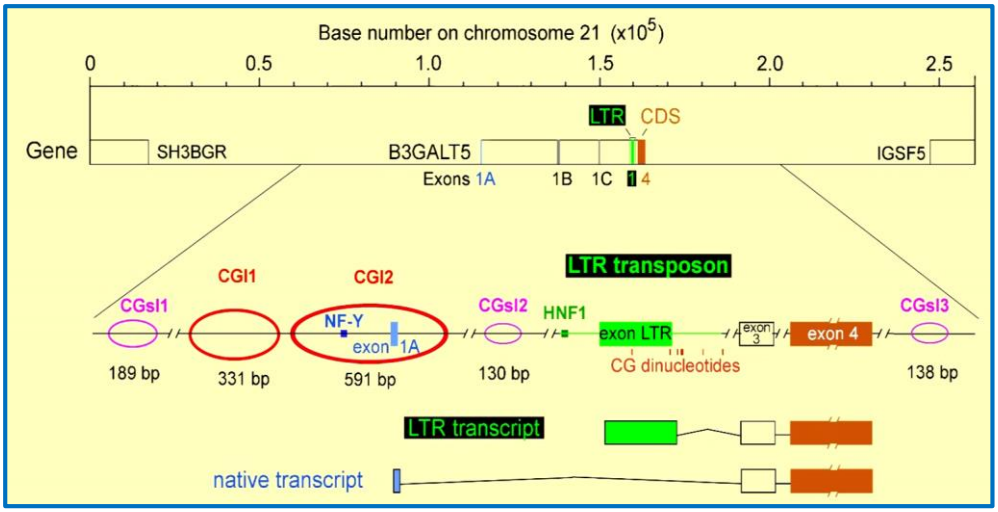


FIGURE 13: Genomic structure and DNA methylation of the human B3GALT5 gene. In the upper part, the genomic sequence of chromosome 21 (0.26 Mbp) encompassing B3GALT5 and intergenic regions is presented together with the positions of the exons and the LTR transposon. In the lower part, the enlarged region show transcription factor binding sites, CpG islands (CGI, in red), CpG short islands (CGsI) of at least 100 bp (pink), CG dinucleotides present in the LTR transposon and the two main transcripts [109].

The *B3GALT5* gene offers an emblematic example of the role of differentially methylated CpG sites in regulating gene transcription. In fact, the concurrent methylation of the CpG islands nearby the native promoter and the demethylation of distant regulatory elements, determine gene silencing in colon cancer.

2. AIMS

Our working hypothesis is that the sLea epitope of CA19.9, as well the other type 1 chain antigens Lea and Leb, are all down-regulated in gastrointestinal adenocarcinomas due to the epigenetic silencing of B3GALT5. According to our model, elevation of circulating sLea in cancer is thus paradoxical. Such a hypothesis relies on recent published data and our preliminary results showing how a pitfall occurred in the past with IH detection, leading to the wrong assumption that CA19.9 was synthesized by tumor cells. In addition, since elevated levels of circulating CA19.9 are reported also in benign conditions, we hypothesize that the circulating antigen could be secreted by the healthy glandular ducts. In cancer, the obstruction of the ducts impairs secretion and determines the reabsorption of the molecules (i.e. mucins) in the bloodstream. In order to confirm our working hypothesis and clarify the doubts regarding the current concept of CA19.9 as a tumor marker, our study is aimed at defining:

- the actual distribution of sLea and related Lewis antigens in pancreas and colon cancer, by immunofluorescence staining of sections prepared from cancer specimens and surrounding normal tissue in order to understand whether or not the circulating antigen is produced by cancer cells.
- The expression levels in the corresponding tissue of cognate glycosyltransferase transcripts involved in the synthesis, determined by real time PCR.
- The status of the chromatin affecting the main B3GALT5 promoters, defined by bisulfite sequencing and ChIP analysis performed on both cell lines and selected cancer specimens.
- What is the distant DNA region recruited by the transcription machinery of the B3GALT5 LTR promoter that boosts transcription, and to prove the inhibitory effect of its demethylation.

3. MATERIAL AND METHODS

3.1 Cell and tissue processing

COLO-205 (colon cancer), MKN-45 (gastric cancer), Huh-7 (hepatocarcinoma), MDA-MB-231 (breast cancer), MDA-MB-231/ HNF1α (a clone expressing HNF1α, [133]), BxPC-3 and Capan-2 (both pancreatic cancer) cell lines were grown and treated with 5AZA (5'-aza-2'-deoxycytidine) as previously reported [85, 133]. Six human colon and two human pancreas samples were collected at surgery as reported [130]. A matched pair of RNAs from cancer and adjacent normal pancreas (Ambion) and two RNAs from normal pancreas (Clontech and Stratagene, respectively) were of commercial origin. Specimens of tumor and adjacent normal tissue (about 80 mm^3 each) were immediately frozen in dry ice and kept at −80 °C. Frozen material was cut in aliquots (about 20 mm^3) that were thawed and used as follows. For RNA extraction, samples were homogenized in 0.14 ml lysis buffer (ReliaPrep RNA miniprep system, Promega) using a rotary homogenizer (UltraTurrax) and processed according to the manufacturer's protocol, including DNAse treatment. Elution was done with 30 µl of nuclease free water heated at 70 °C. For genomic DNA extraction, samples were treated with 0.2 ml of buffer ATL (QIAamp DNA Mini Kit, Qiagen) containing 2.0 mg/ml proteinase K at 60 °C overnight, and then processed according to the manufacturer's protocol. For total lysate preparation, samples were homogenized (Ultra Turrax) in PBS containing Halt protease inhibitor cocktail (Thermofisher) and then brought to RIPA buffer (50 mM Tris-HCl pH 7.4, 1% Nonidet-P40, 0.5% Na-deoxycholate, 0.1% SDS, 150 mM NaCl) containing protease inhibitor cocktail, and kept on ice with frequent vortexing for 1 h. After spinning at 12.000 ×g for 10 min at 4 °C, the clean supernatant was removed and stored at −80 °C. Nuclear protein was extracted as reported [131]. Nucleic acid concentrations were determined using the Qubit fluorometer (Thermofisher) and protein was measured with the Bradford method.

3.2 Reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR)

First strand cDNA was synthesized from 1 to 4 µg of total RNA by Moloney Murine Leukemia virus reverse transcriptase as reported [133]. Control reactions were prepared by omitting the reverse transcriptase. cDNAs (0.2–1.0 µl of first strand reactions) were amplified in a volume of 20 µl using Sybr Premix Ex Taq (Tli RNase H Plus) (Takara), ROX as reference dye and Step One Plus instrument (Applied Biosystem Life Technologies). Primer sequences are reported in Table 2. Annealing temperature was 60 °C. The amounts of amplified target cDNAs were calculated from their respective standard curves and normalized by those for GAPDH.

B3GALT5 NATIVE	
Forward	5'-GGGCTGCCCCGCGCAC
Reverse	5'CACAGTGTAAATGCCAAGAGGA
B3GALT5 LTR	
Forward	5'-CTGGCCTTTGGACCCGAG
Reverse	5'-CACAGTGTAAATGCCAAGAGGA
ST3GAL3	
Forward	5'-CTCTGGGGTCACGAATTGAC
Reverse	5'-TGCTCAGGCCGCTGCATG

TABLE 2: PCR primers for quantification of specific transcripts.

3.3 Dot-blot and western blot analysis

Serial dilutions of total lysates were prepared in water to obtain 80 µl aliquots containing 0.03 to 4 µg of protein, which were applied to a nitrocellulose membrane using a 96-well blotter (BioRad), washed with 0.4 ml water and air dried. Membranes were blocked 30 min at RT with 5% defatted milk in

PBS containing 0.1% Tween-20 (PBS-T), washed 3-times with PBS-T, and incubated at 4 °C overnight with 0.5 µg/ml of each of the following monoclonal antibodies: anti-CA19.9 (NS-1116-19-9, ATCC hybridoma HB 8059), anti-Leb (130–3-F7–5, ATCC hybridoma HB 8325), anti-Lea (151–6-A7–9, ATCC hybridoma HB 8324), which were purified from the respective culture media as reported [68, 84].

Cell pellets and surgical specimens were processed to obtain nuclear extracts using a commercially available kit (NE-PER, Thermofisher) as reported [131]. Aliquots of nuclear extracts (5–20 µg of protein) were separated by 8% SDS-PAGE, transferred to a nitrocellulose membrane (Trans-Blot SD Semi Dry Transfer Cell, Bio\\Rad Laboratories) and blotted with rabbit polyclonal anti-HNF1α/β (sc-8986, 1:200, Santa Cruz Biotechnology) or rabbit polyclonal anti-NF-YA (sc-10779, 1:200, Santa Cruz Biotechnology) according to our published protocol [134].

3.4 Bisulfite sequencing

Genomic DNA (0.1–1.5 µg) was submitted to bisulfite treatment and purification using the Epitect bisulfite Kit (Qiagen). For cloning, the obtained material was amplified by PCR as follows. Amplifications (35 cycles) were performed in 25 µl using a hot-start Taq (Promega) according to the manufacturer's recommendations in the presence of an enhancer (PCRx Enhancer system, Invitrogen) with 2 µl of bisulfite converted DNA as template and first reaction primers (Table 3). Amplification program included a single treatment at 94 °C for 3 min followed by cycles consisting of 1 min at 94 °C, 1 min at 58 °C, and 1 min at 72 °C, and a final extension step at 72 °C for 8 min. Nested PCRs (25 cycles) were performed in a final volume of 50 µl using 1–2.0 µl of the PCR products and inner primers designed to contain restriction sites (Table 3). Reaction mixtures and amplification programs were as above described, but annealing temperature was 62 °C. Amplified fragments were column-cleaned, digested with appropriate restriction

enzymes, column-cleaned again, and cloned into pGl3 or pCDNA3 vectors for sequencing (at Eurofins sequencing service).

	Region 1	
First reaction	forward	5'-GTTTTAGGTTAGGTTTAAGGTAAGG
	reverse	5'-CTACCTCCAAACACATCTAAACC
Second reaction	forward	5'-CGCGAAGCTTTGGTTTGGGGAGAGTTTTAGG
	reverse	5'-CGCGTCTAGATAAACAAACAAATAAATAAAAAAAAATTTC
	Region 2	
First reaction	forward	5'-GGTTAGGGGTAGGTTTAGGTTTG
	reverse	5'-CAATAACCAAACCCCACTATAAAACC
Second reaction	forward	5'-CGCGAAGCTTGGAAAAGAAAATAGTAGAGGTGAG
	reverse	5'-CGCGTCTAGACAAAAAATCTCCAAAAACTATAACCC
	Region 3	
First reaction	forward	5'-GGAAAATAGGGTTTTAATATTATAGTG
	reverse	5'-CAATTCTAACCCCCACATCAATC
Second reaction	forward	5'-GCGCAAGCTTAATTTTTGAGATGGAGTTTTATTTT
	reverse	5'-GCGCTCTAGAATACAAACAACTTCTTATAACTTAC

TABLE 3: Primers for nested PCR after bisulfite treatment. Restriction sites for cloning are underlined.

3.5 ChIP assay

Chromatin was prepared from about 10^7 cells and immunoprecipitated using a commercially available kit (SimpleChIP, Cell Signaling) with the modifications reported [131]. About 1 µg of chromatin DNA, determined by fluorometry upon reversing cross linking (Qubit, Invitrogen), diluted to a final volume of 0.25ml, was used for each immunoprecipitation. Aliquots (10–20 µl) were kept as input DNA. Antibody binding reactions were performed at 4 °C overnight under rotation.

Precipitation with agarose-bound G-protein, washing, elution, reverse of cross-linking, and DNA purification were performed with the kit reagents according to the manufacturer's protocol. The following antibodies were used: anti-rabbit IgG, 1 µl (Cell Signaling); anti-trimethyl-Histone H3 (Lys4), named H3K4me3, 2 µl (GeneSpin, PAb005); anti-trimethyl- Histone H4 (Lys20), named H4K20me3, 2 µl (Millipore #07–749); antiacetyl- Histone H3, named H3KAc, 1 µl (Millipore #06–599). To quantify the immunoprecipitated DNA, qPCR was performed using specific primers (see Table 4 for sequences). Reaction mixture, 20 µl, and amplification conditions were as above described for RT-PCR.

HNF1-binding region	
Forward	5'- TCATGAGTTACAGAGCAAAGCC
Reverse	5'- GCCCACTGACTCACAGCCAATC
Region 2	
Forward	5'- TCAAGGACCAGTGCAGATGC
Reverse	5'- GCTCAAGGG GTCTCCAGG
Region 3	
Forward	5'- TGCCTGGCCAACATCACATG
Reverse	5'- CCA GGA CCT GTC GCT CAC

TABLE 4: Primers for qPCR after ChIP. See Figure 23 for primer location in the genome sequence.

3.6 Immunofluorescence and immunohistochemistry

IF and IH were performed on two primary colon and pancreas adenocarcinoma samples collected from the archive of the Unit of Human Pathology of the San Paolo Hospital Medical School, University of Milan. Archival material was used in compliance with the national guidelines and in respect to the privacy of the patients. IF was performed on formalin-fixed paraffin-embedded 4 μm thick tissue sections as previously reported [135]. Sections were deparaffinised in xylene and rehydrated through a graded series of alcohols. Primary antibodies were those above reported for dot blotting, and were used at the following concentrations: anti-CA19.9, 0.4 μg/ml; anti-Lea 1 μg/ml; anti-Leb 1 μg/ml. Polyclonal Rabbit Anti-Mouse Immunoglobulins/FITC (Dako, dilution 1:80) was used as fluorophore-conjugated secondary antibody. Nuclei were subsequently counterstained with DAPI (Invitrogen). Serial dilutions of primary antibodies were tested on positive and negative controls and on a colon sample to assess the effect of antibody concentrations (Figure 14 and 15). Immunohistochemical studies were carried out on formalin-fixed paraffin-embedded 4 μm thick tissue sections as previously reported [66], with the anti-CA19.9 antibody only, used as for IF. Antigen retrieval was performed at 97.5 °C for 35 min in 9 mM sodium citrate at pH 6.0. Endogenous peroxidase activity was quenched with 3% H_2O_2 for 10 min. Staining was performed with 3,3′ diaminobenzidine as chromogen for 8 min and slides were counterstained with hematoxylin. The reaction was detected with the autostainer (Optimax i6000, Biogenex) by Novolink Max polymer detection system (Novocastra Laboratories). Slides were immunostained in the same batch, to prevent incubation variability and to ensure identical condition for comparison. Slides with absence of the primary antibody were also included as negative controls. Images of hematoxylin/eosin (H&E) staining, immunohistochemical staining and IF

staining were obtained using the digital slide scanner NanoZoomer 2.0 (Hamamatsu Photonics, Japan).

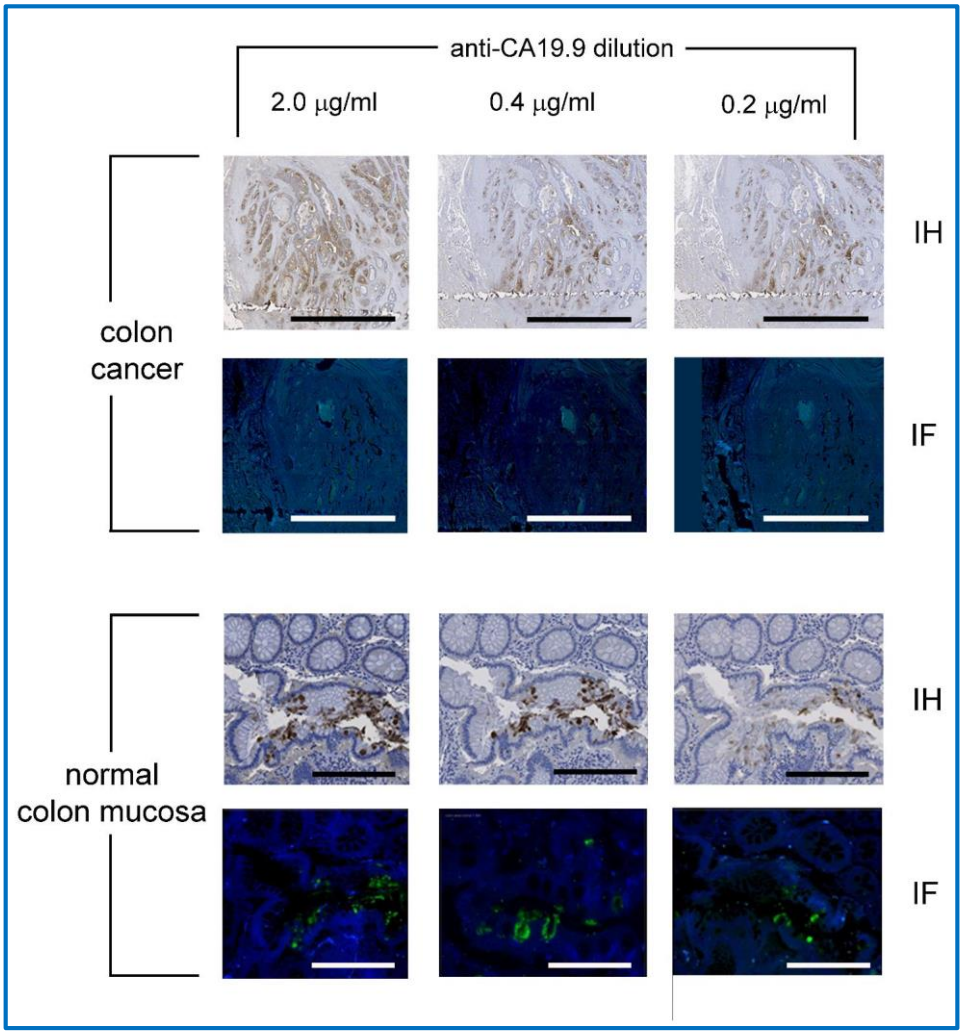


FIGURE 14: Detection of CA19.9 in colon sections using serial dilutions of the NS-1116-19-9 antibody. Sequential sections from a colon cancer sample were studied by IH or IF at the indicated antibody dilutions. For IF detection, nuclei were blue-counterstained with DAPI as a reference. Upper part: sections showing neoplastic transformation are presented at lower magnification (scale bars 2.5 mm). Lower part: sections showing normal mucosa architecture are presented at higher magnification (scale bars 0.25 mm).

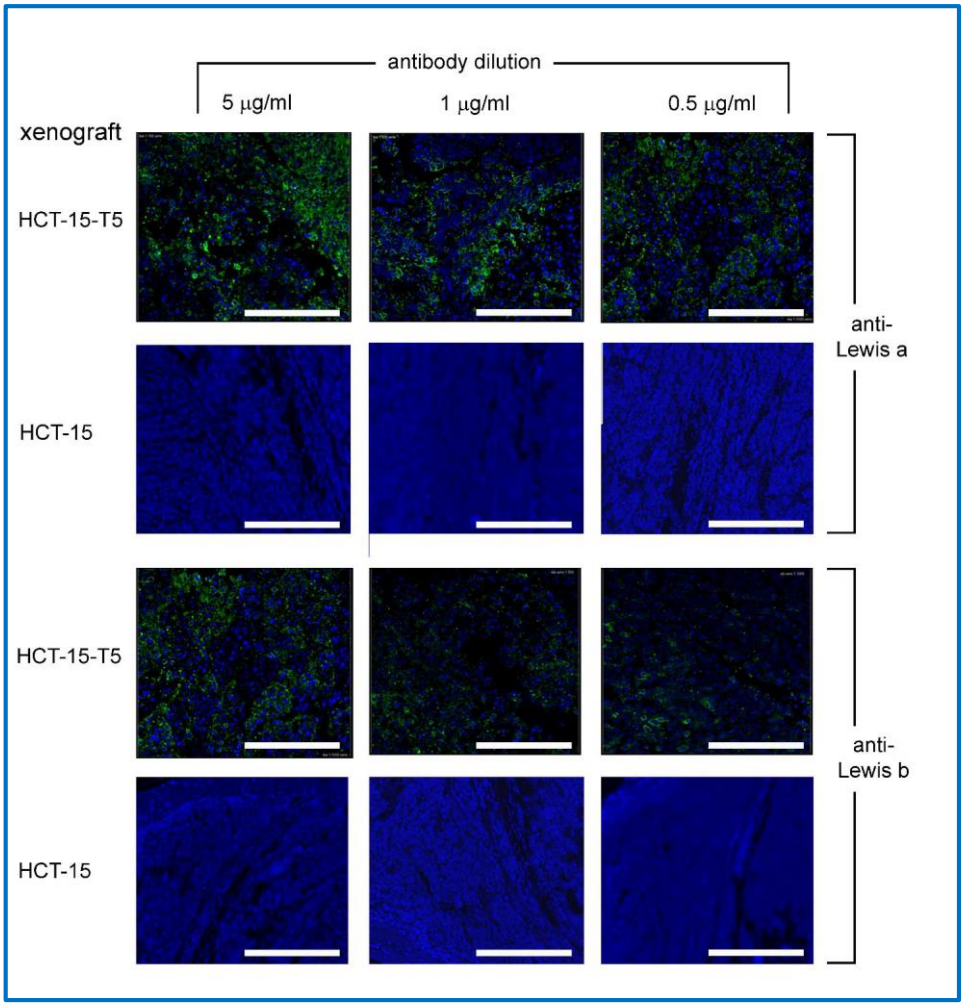


FIGURE 15: Detection of Lewis a and Lewis b antigens by IF using serial dilutions of the antibodies. Sections from two xenografts were studied by IF at the indicated antibody dilutions and nuclei were blue-counterstained with DAPI as a reference. Scale bars = 0.25 mm. HCT-15 denotes the negative control. It is a xenograft obtained inoculating mice with the human colon carcinoma cell line HCT-15 that lacks type 1 chain Lewis antigens due to the absence of B3GALT5 enzyme. HCT-15-T5 denotes the positive control. It is a xenograft obtained inoculating mice with a recombinant clone that expresses type 1 chain Lewis antigens upon transfection of B3GALT5 cDNA. Both xenografts were characterized in a previous report [66].

4. RESULTS

4.1. Detection of Lewis antigens in colon and pancreas samples by IF on tissue sections

To investigate the actual expression of CA19.9 in human gastrointestinal cancers, we analyzed two pairs of colon and pancreas samples by IF with antibodies against type 1 chain Lewis antigens. In both the colon samples there was high expression of Lea in the portion of the section where the normal architecture of the mucosa was well preserved (Figure 16 B and C), which disappeared in the cancerous part (Figure 16 D). CA19.9 and Leb were also lacking in this part, while they presented a spotted reactivity scattered through the preserved normal mucosa of the section (Figure 16 B, C and D). We also tried detection of CA19.9 on the same sections by IH. It is noteworthy that the NS-1116-19-9 antibody did reveal ambiguous CA19.9 reactivity. In fact, the antibody apparently detected the antigen where no Lewis antigen, or Lea but not CA19.9, was present as revealed by IF. (Figure 16). In the case of the pancreas, the features of the two analyzed samples were different from each other. In fact, by IF one reacted mostly with Leb (Figure 17), while the other mostly with CA19.9 (Figure 18). In both cases, reactivity was evident in the part of the sections where the glandular architecture was recognizable, and it was restricted to the pancreatic ducts (Figures. 17 B and 18 B). In the less differentiated part of the sections, the reactivity was not present in the bulk of the tumor cells, but restricted to ducts or ductal-like structures (Figures 17 C and 18 C, D). By IH detection, the obtained patterns were misleading. In fact, false CA19.9 reactivity was more pronounced in the less differentiated parts of the sections, and very evident in those expressing Leb or Lea but not CA19.9 (Figure 17 C). On the other hand, in the sections expressing CA19.9, IH sometimes overlapped IF, making the picture even more ambiguous (Figure 18 E). These results indicated that the data obtained so far by IH are affected by technical

artifacts, while IF suggests that the colon and pancreas may present different patterns of expression of Lewis type 1 antigens.

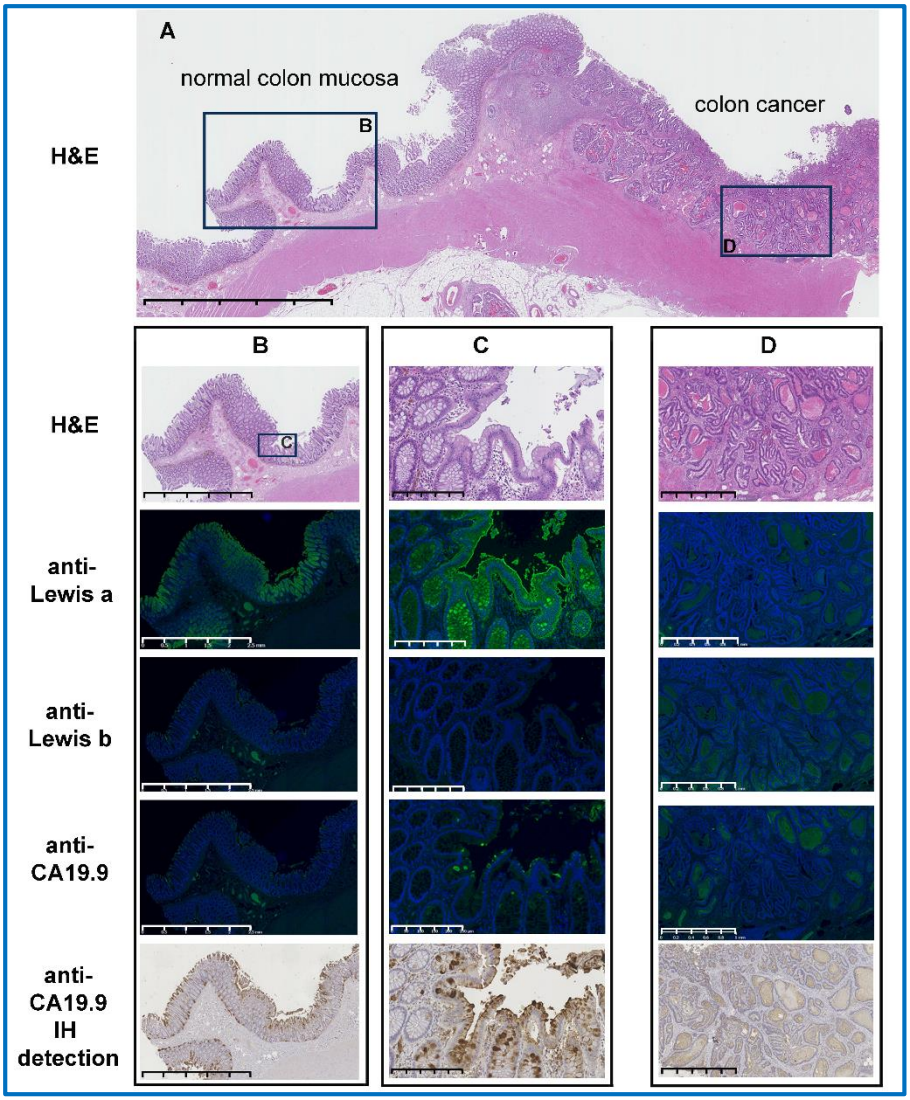


FIGURE 16: Detection of type 1 chain Lewis antigens in colon sections. A section from the colon presenting normal mucosal architecture (left side) or neoplastic transformation (right side) was studied by IF with anti-Lea, −Leb and -CA19.9 antibodies, or by IH with anti-CA19.9 antibody. For IF detection, nuclei were blue-counterstained with DAPI as a reference. (A) Only H&E staining is presented, at the lowest magnification, scale bar=5mm. (B) The area boxed on the left side of panel A (normal colon mucosa) is presented at higher magnification after staining with all antibodies, scale bar=2.5mm. (C) The area boxed in panel B is presented at the highest magnification, scale bar=0.25mm. (D) The area boxed on the right side of panel A (colon cancer) is presented at higher magnification, scale bar = 1 mm. Another section from a second case provided very similar results and is not presented.

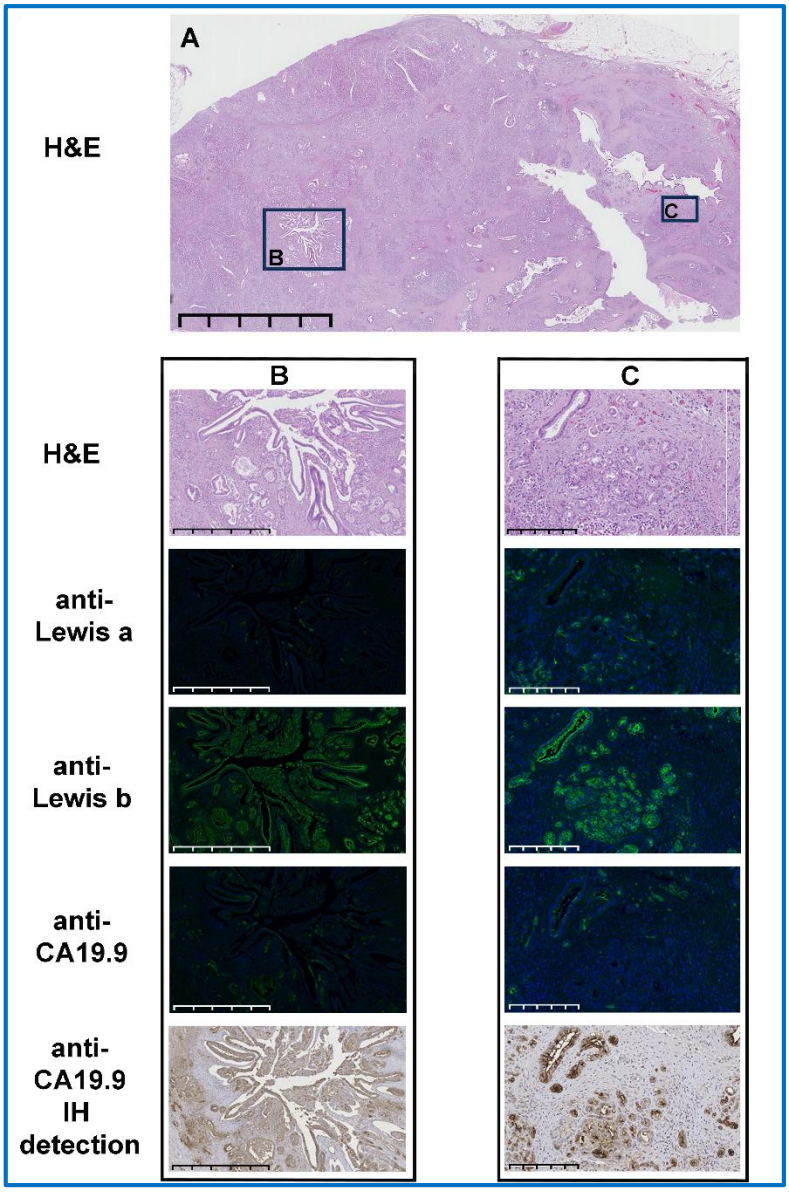


FIGURE 17: Detection of type 1 chain Lewis antigens in pancreatic carcinoma sections (first case). A section from the pancreas presenting well differentiated (left side) or poorly differentiated glandular structure (right side) was studied by IF or by IH as in Figure 16. (A) Only H&E staining is presented, at the lowest magnification, scale bar=5mm. (B) The area boxed on the left side of panel A (well differentiated architecture) is presented at higher magnification after staining with all antibodies, scale bar = 1.0 mm. (C) The area boxed on the right side of panel A (poorly differentiated structure) is presented at higher magnification, scale bar= 0.25 mm.

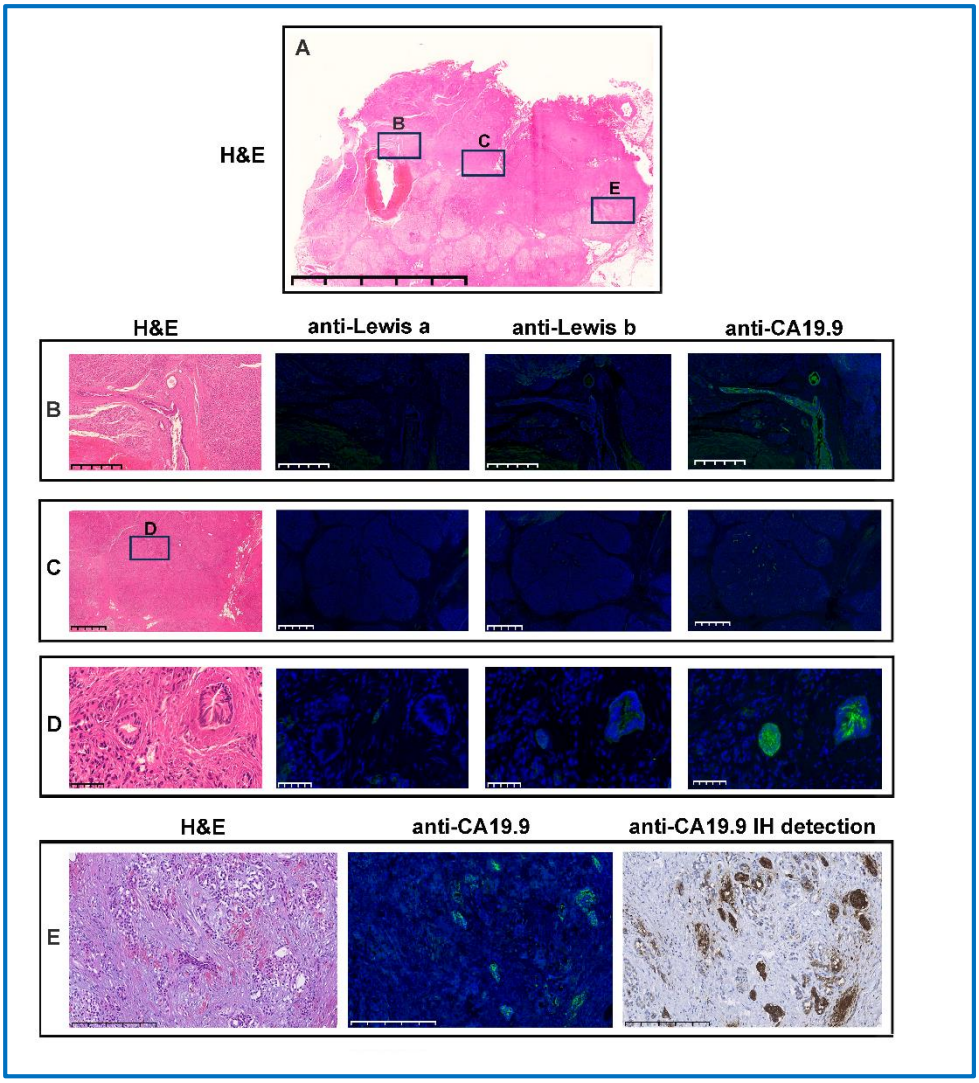


FIGURE 18: Detection of type 1 chain Lewis antigens in pancreatic carcinoma sections from another case. Sections were studied as reported in Figure 17. (A) Only H&E staining is presented, at the lowest magnification, scale bar=10mm. (B) The area boxed on the left side of panel A (well differentiated architecture) is presented at higher magnification, scale bar=0.5mm. (C) The area boxed on the right side of panel A (poorly differentiated structure) is presented at higher magnification, scale bar=1.0mm. (D) The area boxed in panel C is presented at the highest magnification, scale bar=50 μm. (E) A section of poorly differentiated structure is presented to show the misleading effect obtained by IH that provides partially overlapping results with IF; scale bar = 0.25 mm.

4.2 Detection of glycosyltransferase transcripts and cognate Lewis antigens in colon and pancreas tissues

Expression levels of ST3GAL3 and B3GALT5 native and LTR transcripts were determined by RT-qPCR in matched normal and cancer specimens from the colon and the pancreas, and in pancreatic cancer cell lines BxPC-3 and Capan-2, known to express Lewis type 1 antigens [136]. As expected [79, 131, 133, 137], the main features of B3GALT5 transcripts in the colon were the following (Figure 19): first, the high levels of expression in the normal mucosa due to the large amounts of the LTR transcript, which was higher than the native transcript; second, the strong down-regulation of both transcripts in cancer, which made the overall expression of B3GALT5 mRNA negligible. In the pancreas (Figure 19), expression levels of B3GALT5 LTR transcript were similar to, or even lower than, those of the native transcript, without difference between normal and cancer tissues for both transcripts. Moreover, B3GALT5 LTR transcript was undetectable in both BxPC-3 and Capan-2 cell lines. The levels of ST3GAL3 transcript in both normal colon mucosa and colon cancer were low (mean ± SEM, 1.92 ± 0.49 fg/pg of GAPDH), while they were significantly higher (8.59 ± 1.73 fg/pg of GAPDH, $p = 0.006$) in pancreas tissues (Figure 19). The high expression levels of B3GALT5 LTR transcript in the colon and the cancer-associated silencing of B3GALT5 thus appeared to be tissue specific since they were not found in the pancreas.

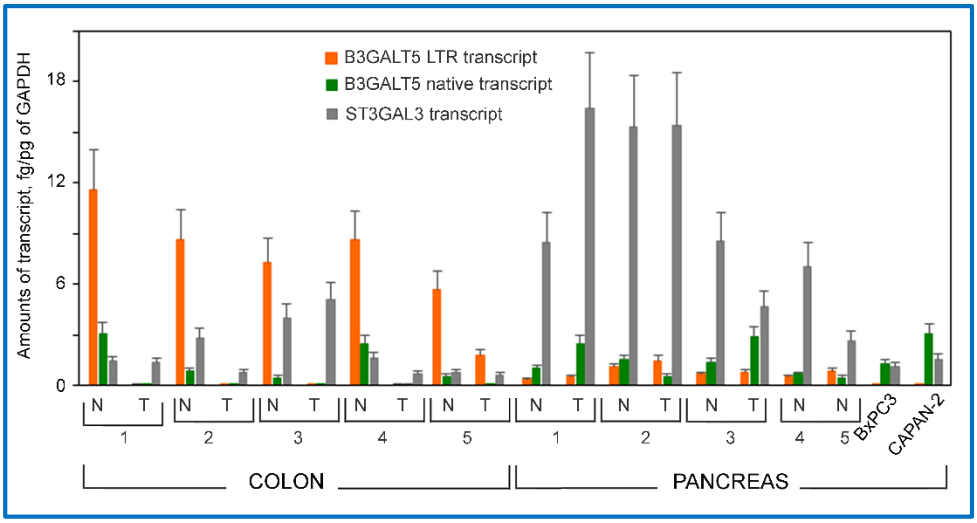


FIGURE 19: Detection of B3GALT5 and ST3GAL3 transcripts in colon and pancreas samples. Matched pairs representing cancer and surrounding normal tissue from the colon and the pancreas were collected at surgery and immediately frozen. The frozen material was cut into aliquots that were processed for analysis. Total RNA was extracted, reverse transcribed, and amplified by qPCR as described in sections 2.1 and 2.2. Pancreas RNA samples 3 (matched pair), 4 and 5 (normal tissue) were of commercial origin.

Results are the mean ± standard deviation for three determinations. Amounts were calculated from a standard curve and normalized to the amounts calculated for GAPDH. N, normal tissue; T, tumor.

The pattern and amount of Lewis type 1 antigens were determined by dot-blot staining of lysates available from the same matched pairs of colon and pancreas tissues. Lea was found to be very abundant in normal colon mucosa, at much higher levels than Leb, and the amount of CA19.9 was the lowest in the analyzed samples. In colon cancer, the levels of all Lewis antigens were dramatically lower, and CA19.9 was almost undetectable. An exception was apparent in colon cancer sample 2, where Leb expression was relevant (Figure 20). The data suggest that Lewis antigen expression depends in part on glycosyltransferase levels, and in particular that B3GALT5 and ST3GAL3 are limiting in the colon for biosynthesis. To verify this aspect, we analyzed a further colon cancer sample known to maintain residual B3GALT5 transcripts (about one third of the adjacent normal mucosa). In this case (Figure 20, colon sample 6), Lewis antigens were detectable in the cancer sample, but still at lower levels than in the normal counterpart, and the amount of CA19.9 remained the lowest. Expression of the antigens in the two pancreas samples was different with respect to the colon. The amounts in cancer were higher than in normal tissue, and those of CA19.9 and Leb higher than Lea. Pancreas sample 2 expressed much more antigen than sample 1. These data suggest that tissue-specific expression of glycosyltransferases is responsible for the different levels of Lewis type 1 antigens, but other mechanisms affect the total amount of antigens present in the tissues, especially in cancer.

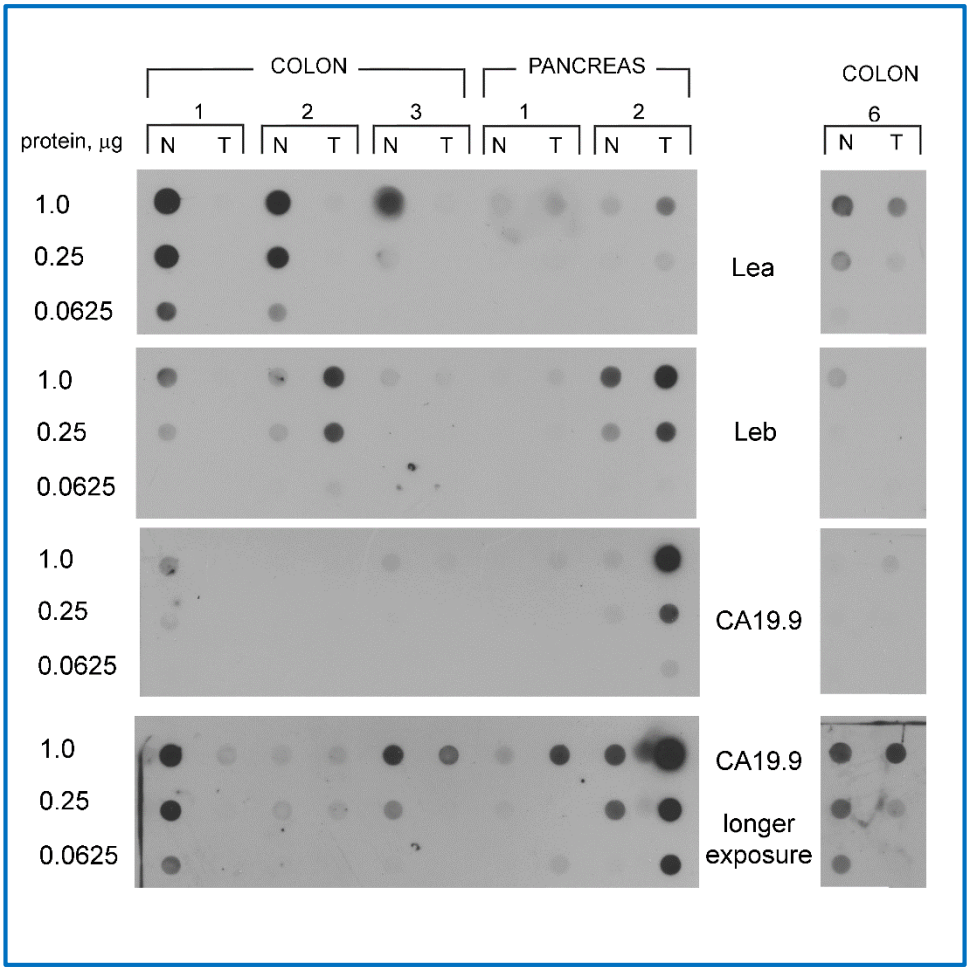


FIGURE 20: Detection of Lewis type 1 antigens in tissue lysates. Tissue samples are those reported in Figure 18. Lysates were prepared in RIPA buffer as described in section 2.1. Antigens were determined by immunostaining of dot-blots prepared with three amounts of sample lysates brought to the same final volume, using monoclonal antibody followed by peroxidase-labeled secondary antibody, as described in section 2.3. Labels show the actual amount of sample protein blotted.

4.3 Regulation of B3GALT5 expression: mechanisms affecting the native promoter

Nuclear proteins were extracted from normal and cancer specimens of pancreas samples 1 and 2, as well as from BxPC-3 and Capan-2 pancreas cell lines, and analyzed by western blotting. NF-Y was easily detected in all specimens, while HNF1α/β was faintly detected in both normal and cancer tissues (Figure 21).

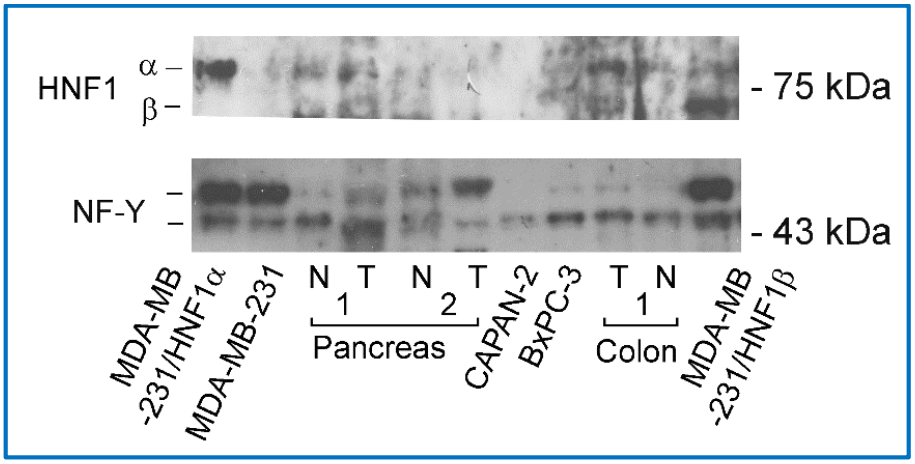


FIGURE 21: Detection of NF-Y and HNF1 transcription factors in pancreas tissues and cell lines. Nuclear extracts (5-20 µg of protein) were separated by 8% SDS-PAGE and transferred to a nitrocellulose membrane that was blotted with anti NF-Y or anti-HNF1 antibodies followed by HRP-labeled secondary antibody and chemiluminescence detection, as detailed under section 2.3. N: normal tissue; T: tumor. Parental MDA-MB-231 lacks HNF1 and expresses the NF-Y doublet corresponding to the long (upper band) and short form (lower band) of the factor [131]. MDA-MB-231/HNF1α and MDA-MB-231/HNF1β are recombinant clones expressing either form of the factor. All such cells and colon sample 1 were used as references.

We then determined the methylation status of CpG islands flanking the native promoter by bisulfite sequencing of genomic DNA extracted from the same pancreas samples (Figure 22). Very low levels of methylation were found in both samples, without differences between normal and cancer tissues. In particular, the methylation levels were even lower than in the BxPC-3 cell line, but comparable with those previously found in Hucc-T1 cells [131], a cell line derived from bile ducts known to express the highest levels of B3GALT5 native transcript. These data indicated that the methylation-dependent mechanism of silencing reported in colon cancer [131] appears not to be active in the pancreas, and may explain the presence of B3GALT5 native transcript and cognate reaction products (including Lewis antigens) found in these cancers only.

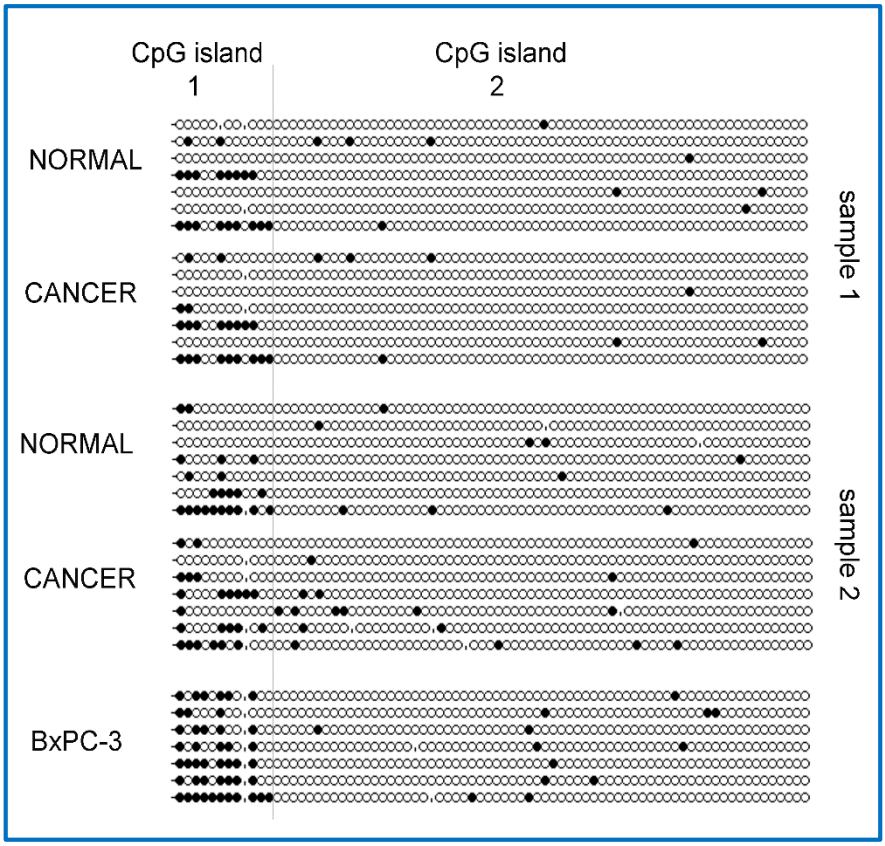


FIGURE 22: **Lollipop representation of methylated CG residues in CpG islands 1 and 2 flanking B3GALT5 native promoter**. Genomic DNA extracted from BxPC-3 cells and pancreas tissue samples reported in Figures 19 and 20 was submitted to bisulfite treatment, amplified through nested PCR, and the obtained fragments cloned as reported in section 2.4. Seven independent clones were submitted to direct DNA sequencing and the obtained sequences analyzed using BiQ Analyzer software (http://biq-analyzer.bioinf.mpi-sb.mpg.de/). Empty circles represent unconserved (unmethylated) cytosines, full circles represent conserved (methylated) cytosines. Dots (no circle) represent potential polymorphisms or sequencing errors. Twelve out of 15 CG pairs were analyzed for CpG 1, and are shown on the left of the thin line; 66 out of 76 pairs were analyzed for CpG 2 and are shown on the right.

4.4. Epigenetic regulation of B3GALT5 LTR promoter: individuation of a candidate DNA sequence

No true CpG island flanks the LTR promoter, but instead only stretches of spared CG dinucleotides sometimes forming short islands about 100 bp long [109]. On this basis, we compared by bisulfite sequencing the status of DNA methylation of stretches of CG dinucleotides spaced along the 5′ side of the LTR promoter in three cell lines: COLO-205 (LTR transcript: 15.1 fg/pg of β-actin), MKN-45 (LTR transcript: 3.2 fg/pg of β-actin), and Huh-7, not expressing the transcript at all even in the presence of high levels of HNF1 [133]. In the light of preliminary results, we focused on two DNA regions spanning about 1 kb from the LTR promoter, regions 1 and 2 (15 and 24 CG residues, respectively (Figure 23).

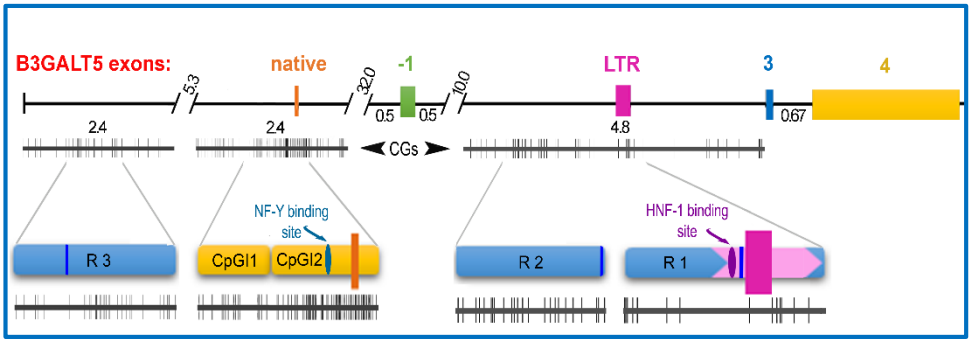


FIGURE 23: **Schematic representation of the B3GALT5 gene and DNA regions potentially involved in the epigenetic regulation of transcription**. DNA size is in Kbp. The relative location of the following B3GALT5 exons are drawn approximately to scale: native exon 1 A (23 bp), exon −1 (238 bp), LTR exon 1 (273 bp), exon 3 (166 bp), and exon 4 (2328 bp) that contains the whole coding sequence. Thin vertical bars represent CG dinucleotides and the enlargements show true CpG islands (CpGI) and relevant stretches of CG dinucleotides (CGs). R1 (1112 bp), R2 (821 bp) and R3 (891 bp), denote DNA regions 1, 2 and 3, respectively, subjected to bisulfite sequencing and ChIP assay. Transcription factor binding sites of both B3GALT5 promoters are indicated by arrows. The blue vertical lines in each region indicate the target sequences for ChIP quantification by qPCR. The entire LTR transposon (644 bp) present in region 1 is pink-arrowed.

Region 1, encompassing the LTR transposon, presented a methylation pattern unrelated to LTR transcription (Figure 24).

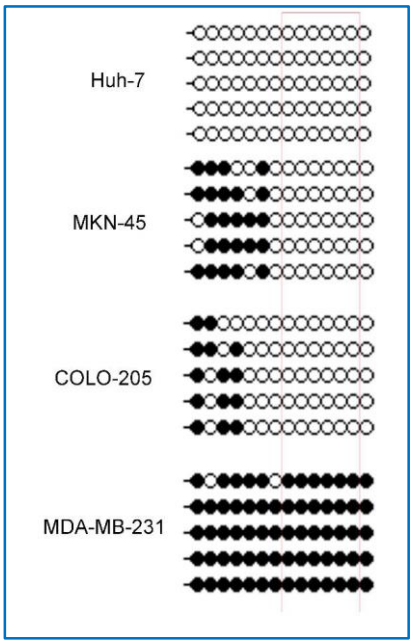


FIGURE 24: Lollipop representation of methylated CG residues in DNA region 1. Genomic DNA extracted from cell lines was processed as in Figure 22. Five independent clones were submitted to direct DNA sequencing and the obtained sequences analyzed using BiQ Analyzer software; results are presented as in Figure 22. Fourteen out of 15 CG pairs were analyzed for region 1 as depicted in Figure 23. The red lines comprise CG residues belonging to the LTR transposon. MDA-MB-231 cells expresses neither HNF1 nor B3GALT5 LTR transcript, Huh-7 expresses HNF1 only, but not the transcript; MKN-45 and COLO-205 cells express both, but transcript levels are much higher in COLO-205 [10].

Conversely, a surprisingly high degree of homogeneous methylation of DNA region 2 was found in COLO-205 cells, which was lower in MKN-45, while in Huh-7 the region was largely unmethylated. Such a methylation pattern was destroyed by 5AZA treatment (Figure 25 A). We thus analyzed two matched pairs of colon specimens, and found that the methylation pattern was somewhat recognizable in normal mucosa, but not in cancer (Figure 25 B). Conversely, in two matched specimens from the pancreas, the methylation pattern of the same region was scarcely preserved, without differences between normal and cancer tissues (Figure 25 B), in accordance with the low levels of expression of the LTR transcript irrespective of normal or cancer status.

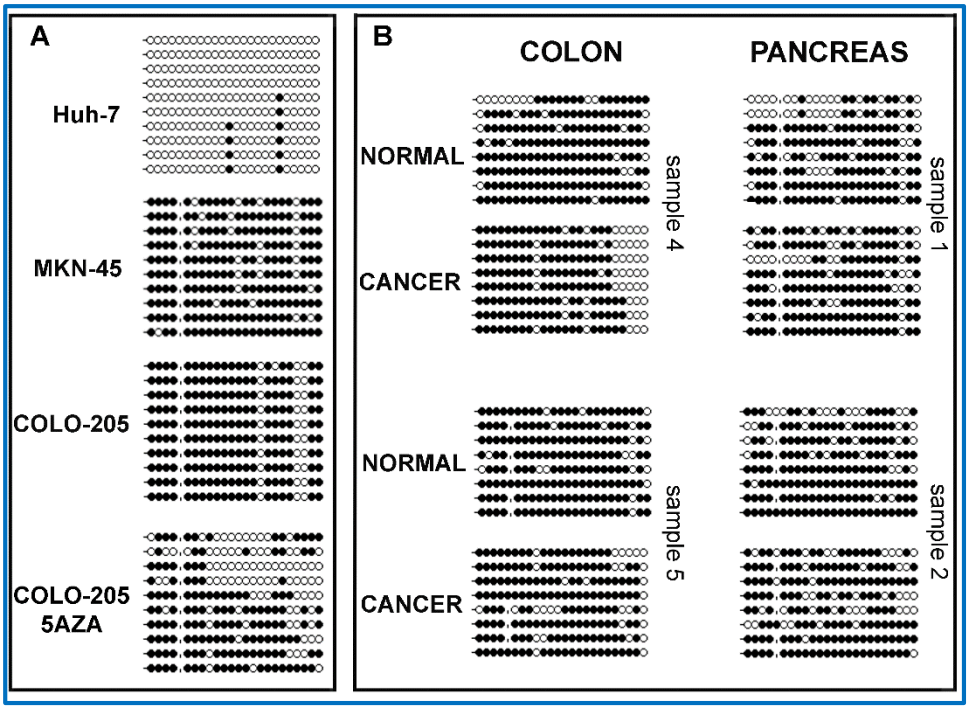


FIGURE 25: Lollipop representation of methylated CG residues in DNA region 2. Genomic DNA extracted from cell lines (A) or colon and pancreas tissues (B) was processed as in Figure 22. Tissue samples are those indicated in previous figures. Ten (A) or eight (B) independent clones were submitted to direct DNA sequencing and the obtained sequences analyzed using BiQ Analyzer software; results are presented as in Figure 22. Twenty-four CG pairs were analyzed for region 2 as depicted in Figure 23. Huh-7 cells express HNF1 but not B3GALT5 LTR transcript. MKN-45 and COLO-205 cells express both, but transcript levels are much higher in COLO-205, and drop down upon 5AZA treatment [69]. The corresponding levels of B3GALT5 LTR transcript in each tissue sample are presented in Figure 19.

To confirm the potential involvement of such a region in LTR promoter activation, we designed ChIP experiments. Unfortunately, no commercially available anti-HNF1 antibodies were found suitable for ChIP assays. We tried specific anti-HNF1α or -HNF1β antibodies, as well as other antibodies recognizing both isoforms, but none worked properly. We thus tried to find histone modifications associated with the HNF1 binding region of the LTR promoter by ChIP assays with various antibodies. Using chromatin prepared

from COLO-205 and MKN-45 (both expressing HNF1), or MDA-MB-231 cells (not expressing HNF1), we found that H3KAc was recruited to the HNF1 binding site and associated with LTR transcription (Figure 26 A). H3 acetylation is indeed commonly considered as a marker of active chromatin. Surprisingly, H4K20 trimethylation provided a slight enrichment, while H3K4 trimethylation did not. This was not expected, since H4K20 trimethylation and H3K4 trimethylation are considered markers of inactive and active chromatin, respectively. Interestingly, an overlapping pattern was found using PCR primers specific for region 2 (Figure 26 B), but not for region 3 (Figure 26 C). Recruitment of H3KAc to the HNF1 binding region of LTR (Figure 26 D), as well as to region 2 (Figure 26 E), was abolished when COLO-205 cells were treated with 5AZA, while it was totally lacking in Huh-7 cells (expressing HNF1 but lacking LTR transcription), and minimal in MDA-MB-231 cells expressing HNF1α (MDA-MB-231/HNF1α). Altogether, these data suggest that region 2 is probably involved in LTR promoter activation, which requires a permissive chromatin status characterized by a specific pattern of DNA methylation and a set of histone modifications unknown at present.

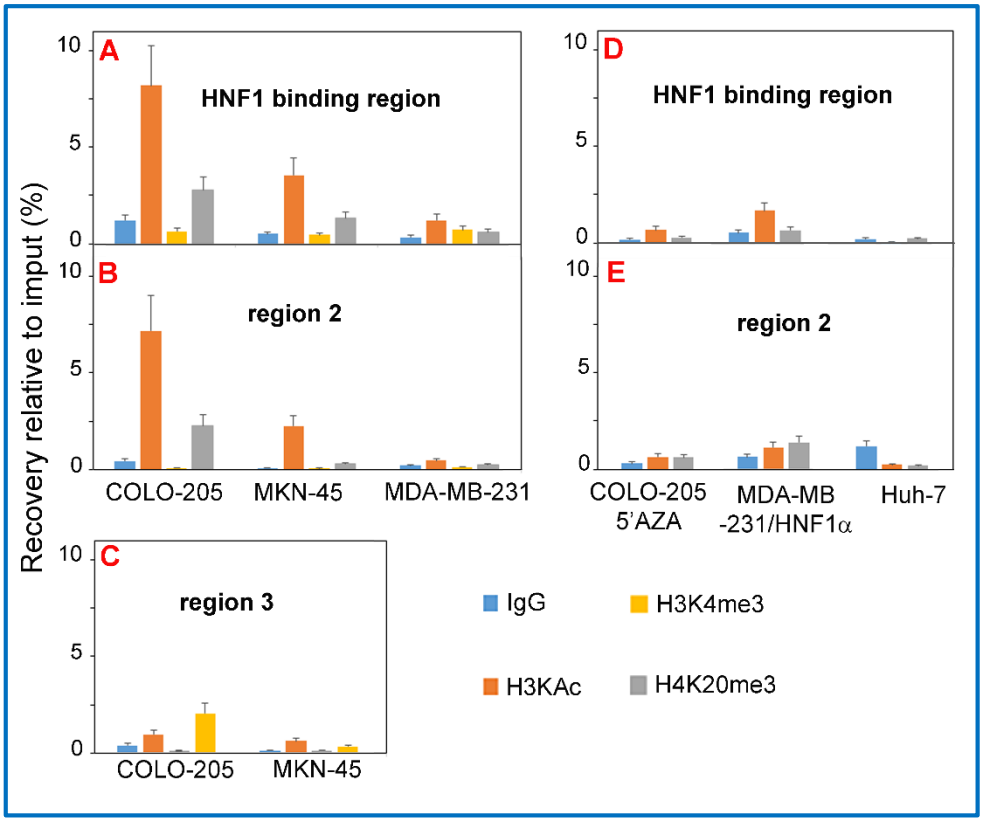


FIGURE 26: ChIP analysis of histone modifications of DNA regions upstream of the B3GALT5 LTR transcription starting site. Chromatin was prepared from different cell lines or clones, cross-linked, digested and immunoprecipitated with the indicated antibodies as reported in section 2.5. Aliquots of immunoprecipitated DNA and dilutions of non-precipitated (input) DNA were submitted to qPCR amplification using primer pairs specific for the HNF1 binding region and for regions 2 and 3. The exact location of target DNA is shown in Figure 23. The relative levels of each histone modification were calculated by a standard curve and normalized to the input DNA. Results are the mean ± standard deviation for at least two amplifications of two independent precipitations.

5. DISCUSSION

We already showed that mouse tissues, which lack CA19.9 and all type 1 chain Lewis antigens due to the absence of α1,4 fucosyltransferase activity [98, 138], are stained by IH with the anti-CA19.9 antibody NS-1116-19-9. Conversely, reference COLO-205 and SW-1116 cells lines, as well as xenografts obtained with recombinant HCT-15 cells expressing B3GALT5, are not, indicating that the antibody loses epitope specificity and acquires false cross-reactivity when used in combination with histochemistry [66]. Here, the comparative detection of CA19.9 by IF or IH on human tissue sections revealed that the true antigen is scarcely detectable in colon mucosa and virtually absent in colon cancer. An overlapping result was obtained by dot-blot staining of tissue lysates where transcript analysis indicated low levels of B3GALT5 and ST3GAL3. These data suggest that the levels of B3GALT5 and ST3GAL3 transcripts do affect Lewis antigen patterns and amounts, as previously established for FUT3 and FUT2 [67, 74]. In particular, low levels of B3GALT5 accompanied by low levels of ST3GAL3, which is reported as a frequent condition in colon cancer [79, 137, 139], do not allow significant synthesis of CA19.9, nor elevation of circulating antigen per se. The origin of circulating CA19.9 antigen in colon cancer thus remains an open question. However, since it is restricted to patients at the last stage of the disease [15, 49], we speculate that it may depend on the concurrent involvement of the biliary system [48], since bile contains huge amounts of the antigen [47, 140]. In the pancreas, the expression pattern of type 1 chain Lewis antigens was rather different than in the colon, as was that of cognate glycosyltransferase transcripts. The levels of B3GALT5 transcript were lower than in the colon, but maintained in cancer, and the levels of ST3GAL3 transcript were higher than in the colon. Concurrently, CA19.9 was present in both normal and cancer status, together with Leb and Lea antigens, and reactivity was always restricted to ducts or duct-like structures only, while the bulk of the other cancer cells were negative. The

relative amounts of each antigen appeared variable in different patients. In this regard, the FUT3 and FUT2 genotype of the patients may play a relevant role, as mentioned above. No qualitative differences appeared between normal and cancer status while dot-blot staining suggested that the amounts of Lewis antigens may be higher in cancer than in the normal counterpart, although the levels of glycosyltransferase transcripts are not. It has already been reported that multiple proteins carry the sLea epitope in pancreatic cancer but none appears responsible for cancer up-regulation [141]. Taking together all these observations, we speculate that Lewis antigen accumulation occurs as a consequence of obstructed and uncanalized neoformed ducts, and/or inverted polarity of transformed ductal cells [142], which in turn give rise to reabsorption into vessels and elevation in circulating levels. CA19.9 thus appears as a physiological product that acquires tumor marker properties, as already suggested 30 years ago [143]. According to our data, Leb may share the same properties in the pancreas and account for some of the patients where elevation of serum CA19.9 does not occur.

Due to the relevant role of B3GALT5, we tried to elucidate the mechanism of differential gene regulation in the two tissues. We found that the transcription factors required for the activation of native and LTR promoters, NF-Y and HNF1, respectively, were both expressed in pancreas tissues, as expected. However, HNF1 is faintly detected probably because its expression is restricted to some cell types, namely those forming the ducts. Interestingly, it was faintly detected even in the pancreatic cell lines. In the analyzed pancreas samples, we found that the two CpG islands flanking the native promoter were both largely unmethylated, and remained unmethylated in cancer, accounting for the persistent expression of the native B3GALT5 transcript in cancer. These findings corroborate the concept that DNA methylation is involved in tissue-specific gene expression [105, 114]. Based on the inhibitory effect of 5AZA on LTR transcription [133], we screened DNA

sequences upstream of the LTR promoter by bisulfite sequencing, searching for stretches of CG dinucleotides whose degree of methylation associates with LTR transcription. We found a candidate stretch of 24 CGs located about 1000 bp upstream of the LTR promoter (region 2). They were unmethylated in Huh-7 cells, lacking transcription even in the presence of HNF1, and methylated in cells expressing the transcript. In particular, we found a kind of methylation motive in COLO-205 cells whose disruption by 5AZA treatment leads to reduced LTR transcription, as found in pancreas tissues or colon cancers expressing the transcript from low to undetectable levels. Conversely, such a methylation motive was rather preserved in colon mucosa expressing moderate to high levels of LTR transcript. ChIP experiments indicated that such a DNA sequence, referred to as region 2 (Figure 22), is probably associated with the HNF1 binding elements of the LTR promoter in active chromatin, marked by histone modifications unknown at present. Demethylation through 5AZA treatment abolished the association to the HNF1 binding region, supporting the involvement of the sequence in transcription. Interestingly, such a region is immediately adjacent to the insertion of the retroviral sequence that contains the HNF1 binding site (Figure 22) and may play a role in stabilizing the transposon during evolution.

6. CONCLUSION

Our data suggest a novel conceptual frame underlying serum CA19.9 and type 1 chain antigen elevation in gastrointestinal cancers, where obstruction/reverse of polarity and individual glycosyltransferase pattern play the pivotal role in a tissue-specific manner. In pancreatic cancer, Lea and Leb appear candidate alternative markers for some CA19.9-negative patients. In colon cancer, the more probable origin is from outside the organ, and bile is the likely candidate. In gastric and bile duct cancers, information is lacking. In all cases, we suggest that for a rational clinical use and for research, the simple serum CA19.9 determination is not sufficient, and is potentially misleading. More experimental work is necessary, studying each patient from a personalized perspective. Multiple antigens should be detected in the available surgery specimens by proper techniques (no immunohistochemistry with NS- 1116-19-9 antibody), and in the blood of the same patients. FUT3 and FUT2 allelic status relative to common mutations should be determined, together with the levels of B3GALT5 and ST3GAL3 transcripts. Antigen determination in a bile sample when available would be very useful. These aspects can be studied using commonly available procedures in vast cohorts of patients, while the role and mechanisms of reverse of polarity in secretive cancer cells deserves more attention and requires elucidation at the molecular level. Although surprising after 30 years since their first use, it seems that the story of CA19.9 and related antigens has to be rewritten and promises fruitful clinical applications.

7. REFERENCES

[1] B. a. W. Stewart, C.P., "World Cancer Report 2014" *International Agency for Research on Cancer, WHO*, 2014.

[2] H. Marusawa, and B. J. Jenkins, "Inflammation and gastrointestinal cancer: an overview", *Cancer Lett,* vol. 345, no. 2, pp. 153-6, Apr 10, 2014.

[3] L. A. Torre, F. Bray, R. L. Siegel *et al.*, "Global cancer statistics, 2012", *CA Cancer J Clin,* vol. 65, no. 2, pp. 87-108, Mar, 2015.

[4] K. W. Jasperson, T. M. Tuohy, D. W. Neklason *et al.*, "Hereditary and familial colon cancer", *Gastroenterology,* vol. 138, no. 6, pp. 2044-58, Jun, 2010.

[5] K. W. Kinzler, and B. Vogelstein, "Lessons from hereditary colorectal cancer", *Cell,* vol. 87, no. 2, pp. 159-70, Oct 18, 1996.

[6] W. M. Grady, and S. D. Markowitz, "Genetic and epigenetic alterations in colon cancer", *Annu Rev Genomics Hum Genet,* vol. 3, pp. 101-28, 2002.

[7] M. Perucho, "Tumors with microsatellite instability: many mutations, targets and paradoxes", *Oncogene,* vol. 22, no. 15, pp. 2223-5, Apr 17, 2003.

[8] S. N. Thibodeau, G. Bren, and D. Schaid, "Microsatellite instability in cancer of the proximal colon", *Science,* vol. 260, no. 5109, pp. 816-9, May 7, 1993.

[9] K. Tamas, A. M. Walenkamp, E. G. de Vries *et al.*, "Rectal and colon cancer: Not just a different anatomic site", *Cancer Treat Rev,* vol. 41, no. 8, pp. 671-9, Sep, 2015.

[10] A. Jemal, T. Murray, E. Ward *et al.*, "Cancer statistics, 2005", *CA Cancer J Clin,* vol. 55, no. 1, pp. 10-30, Jan-Feb, 2005.

[11] S. Mudassar, M. S. Khan et al. "Possible Role of Proto-Oncogenes in Colorectal Cancer — A Population Based Study", Colorectal Cancer

- Surgery, Diagnostics and Treatment, J. Khan (Ed.), InTech, DOI: 10.5772/50593, Mar, 2014.

[12] R. Jaenisch, and A. Bird, "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals", *Nat Genet,* vol. 33 Suppl, pp. 245-54, Mar, 2003.

[13] M. L. Veigl, L. Kasturi, J. Olechnowicz *et al.*, "Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers", *Proc Natl Acad Sci U S A,* vol. 95, no. 15, pp. 8698-702, Jul 21, 1998.

[14] H. J. Freeman, "Early stage colon cancer", *World J Gastroenterol,* vol. 19, no. 46, pp. 8468-73, Dec 14, 2013.

[15] K. Yamashita, and M. Watanabe, "Clinical significance of tumor markers and an emerging perspective on colorectal cancer", *Cancer Sci,* vol. 100, no. 2, pp. 195-9, Feb, 2009.

[16] T. Kamisawa, L. D. Wood, T. Itoi *et al.*, "Pancreatic cancer", *Lancet,* vol. 388, no. 10039, pp. 73-85, Jul 2, 2016.

[17] R. Freelove, and A. D. Walling, "Pancreatic cancer: diagnosis and management", *Am Fam Physician,* vol. 73, no. 3, pp. 485-92, Feb 1, 2006.

[18] J. B. Kisiel, T. C. Yab, W. R. Taylor *et al.*, "Stool DNA testing for the detection of pancreatic cancer: assessment of methylation marker candidates", *Cancer,* vol. 118, no. 10, pp. 2623-31, May 15, 2012.

[19] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2015", *CA Cancer J Clin,* vol. 65, no. 1, pp. 5-29, Jan-Feb, 2015.

[20] L. Rahib, B. D. Smith, R. Aizenberg *et al.*, "Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States", *Cancer Res,* vol. 74, no. 11, pp. 2913-21, Jun 1, 2014.

[21] R. H. Hruban, M. Goggins, J. Parsons *et al.*, "Progression model for pancreatic cancer", *Clin Cancer Res,* vol. 6, no. 8, pp. 2969-72, Aug, 2000.

[22] P. Ghaneh, E. Costello, and J. P. Neoptolemos, “Biology and management of pancreatic cancer”, *Gut,* vol. 56, no. 8, pp. 1134-52, Aug, 2007.

[23] R. H. Hruban, A. Maitra, and M. Goggins, “Update on pancreatic intraepithelial neoplasia”, *Int J Clin Exp Pathol,* vol. 1, no. 4, pp. 306-16, 2008.

[24] S. Yonezawa, M. Higashi, N. Yamada *et al.*, “Precursor lesions of pancreatic cancer”, *Gut Liver,* vol. 2, no. 3, pp. 137-54, Dec, 2008.

[25] L. A. Brosens, W. M. Hackeng, G. J. Offerhaus *et al.*, “Pancreatic adenocarcinoma pathology: changing "landscape"“, *J Gastrointest Oncol,* vol. 6, no. 4, pp. 358-74, Aug, 2015.

[26] M. Malumbres, and M. Barbacid, “RAS oncogenes: the first 30 years”, *Nat Rev Cancer,* vol. 3, no. 6, pp. 459-65, Jun, 2003.

[27] E. Fokas, E. O'Neill, A. Gordon-Weeks *et al.*, “Pancreatic ductal adenocarcinoma: From genetics to biology to radiobiology to oncoimmunology and all the way back to the clinic”, *Biochim Biophys Acta,* vol. 1855, no. 1, pp. 61-82, Jan, 2015.

[28] R. H. Hruban, R. E. Wilentz, and S. E. Kern, “Genetic progression in the pancreatic ducts”, *Am J Pathol,* vol. 156, no. 6, pp. 1821-5, Jun, 2000.

[29] J. P. Morton, P. Timpson, S. A. Karim *et al.*, “Mutant p53 drives metastasis and overcomes growth arrest/senescence in pancreatic cancer”, *Proc Natl Acad Sci U S A,* vol. 107, no. 1, pp. 246-51, Jan 5, 2010.

[30] J. D. Day, J. A. Digiuseppe, C. Yeo *et al.*, “Immunohistochemical evaluation of HER-2/neu expression in pancreatic adenocarcinoma

and pancreatic intraepithelial neoplasms", *Hum Pathol,* vol. 27, no. 2, pp. 119-24, Feb, 1996.

[31] M. Goggins, R. H. Hruban, and S. E. Kern, "BRCA2 is inactivated late in the development of pancreatic intraepithelial neoplasia: evidence and implications", *Am J Pathol,* vol. 156, no. 5, pp. 1767-71, May, 2000.

[32] E. S. Calhoun, J. B. Jones, R. Ashfaq *et al.*, "BRAF and FBXW7 (CDC4, FBW7, AGO, SEL10) mutations in distinct subsets of pancreatic cancer: potential therapeutic targets", *Am J Pathol,* vol. 163, no. 4, pp. 1255-60, Oct, 2003.

[33] J. Q. Cheng, B. Ruggeri, W. M. Klein *et al.*, "Amplification of AKT2 in human pancreatic cells and inhibition of AKT2 expression and tumorigenicity by antisense RNA", *Proc Natl Acad Sci U S A,* vol. 93, no. 8, pp. 3636-41, Apr 16, 1996.

[34] B. A. Ruggeri, L. Huang, M. Wood *et al.*, "Amplification and overexpression of the AKT2 oncogene in a subset of human pancreatic ductal adenocarcinomas", *Mol Carcinog,* vol. 21, no. 2, pp. 81-6, Feb, 1998.

[35] A. Vincent, J. Herman, R. Schulick *et al.*, "Pancreatic cancer", *Lancet,* vol. 378, no. 9791, pp. 607-20, Aug 13, 2011.

[36] M. J. Duffy, "Tumor markers in clinical practice: a review focusing on common solid cancers", *Med Princ Pract,* vol. 22, no. 1, pp. 4-11, 2013.

[37] N. L. Henry, and D. F. Hayes, "Cancer biomarkers", *Mol Oncol,* vol. 6, no. 2, pp. 140-6, Apr, 2012.

[38] T. Malati, "Tumour markers: An overview", *Indian J Clin Biochem,* vol. 22, no. 2, pp. 17-31, Sep, 2007.

[39] S. E. Bates, "Clinical applications of serum tumor markers", *Ann Intern Med,* vol. 115, no. 8, pp. 623-38, Oct 15, 1991.

[40] M. J. Duffy, “Role of tumor markers in patients with solid cancers: A critical review”, *Eur J Intern Med,* vol. 18, no. 3, pp. 175-84, May, 2007.

[41] A. S. Schrohl, M. Holten-Andersen, F. Sweep *et al.*, “Tumor markers: from laboratory to clinical utility”, *Mol Cell Proteomics,* vol. 2, no. 6, pp. 378-87, Jun, 2003.

[42] H. Koprowski, Z. Steplewski, K. Mitchell *et al.*, “Colorectal carcinoma antigens detected by hybridoma antibodies”, *Somatic Cell Genet,* vol. 5, no. 6, pp. 957-71, Nov, 1979.

[43] J. L. Magnani, B. Nilsson, M. Brockhaus *et al.*, “A monoclonal antibody-defined antigen associated with gastrointestinal cancer is a ganglioside containing sialylated lacto-N-fucopentaose II”, *J Biol Chem,* vol. 257, no. 23, pp. 14365-9, Dec 10, 1982.

[44] D. Zopf, and G. C. Hansson, “The chemical basis for expression of the sialyl-Le(a) antigen”, *Adv Exp Med Biol,* vol. 228, pp. 657-76, 1988.

[45] K. A. Paschos, D. Canovas, and N. C. Bird, “The engagement of selectins and their ligands in colorectal cancer liver metastases”, *J Cell Mol Med,* vol. 14, no. 1-2, pp. 165-74, Jan, 2010.

[46] S. L. Ong, A. Sachdeva, G. Garcea *et al.*, “Elevation of carbohydrate antigen 19.9 in benign hepatobiliary conditions and its correlation with serum bilirubin concentration”, *Dig Dis Sci,* vol. 53, no. 12, pp. 3213-7, Dec, 2008.

[47] L. Mare, A. Caretti, R. Albertini *et al.*, “CA19.9 antigen circulating in the serum of colon cancer patients: where is it from?”, *Int J Biochem Cell Biol,* vol. 45, no. 4, pp. 792-7, Apr, 2013.

[48] D. V. Mann, R. Edwards, S. Ho *et al.*, “Elevated tumour marker CA19-9: clinical interpretation and influence of obstructive jaundice”, *Eur J Surg Oncol,* vol. 26, no. 5, pp. 474-9, Aug, 2000.

[49] Z. Vukobrat-Bijedic, A. Husic-Selimovic, A. Sofic *et al.*, “Cancer Antigens (CEA and CA 19-9) as Markers of Advanced Stage of Colorectal Carcinoma”, *Med Arch,* vol. 67, no. 6, pp. 397-401, Dec, 2013.

[50] U. K. Ballehaninna, and R. S. Chamberlain, “The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal”, *J Gastrointest Oncol,* vol. 3, no. 2, pp. 105-19, Jun, 2012.

[51] S. Scara, P. Bottoni, and R. Scatena, “CA 19-9: Biochemical and Clinical Aspects”, *Adv Exp Med Biol,* vol. 867, pp. 247-60, 2015.

[52] C. Galli, D. Basso, and M. Plebani, “CA 19-9: handle with care”, *Clin Chem Lab Med,* vol. 51, no. 7, pp. 1369-83, Jul, 2013.

[53] J. A. Ferreira, A. Magalhaes, J. Gomes *et al.*, “Protein glycosylation in gastric and colorectal cancers: Toward cancer detection and targeted therapeutics”, *Cancer Lett*, Jan 29, 2016.

[54] N. Furuya, S. Kawa, O. Hasebe *et al.*, “Comparative study of CA242 and CA19-9 in chronic pancreatitis”, *Br J Cancer,* vol. 73, no. 3, pp. 372-6, Feb, 1996.

[55] M. J. Duffy, C. Sturgeon, R. Lamerz *et al.*, “Tumor markers in pancreatic cancer: a European Group on Tumor Markers (EGTM) status report”, *Ann Oncol,* vol. 21, no. 3, pp. 441-7, Mar, 2010.

[56] R. Kannagi, “Carbohydrate antigen sialyl Lewis a--its pathophysiological significance and induction mechanism in cancer progression”, *Chang Gung Med J,* vol. 30, no. 3, pp. 189-209, May-Jun, 2007.

[57] M. Gion, R. Franceschini, C. Rosin *et al.*, “An epidemiology-based model to estimate the rate of inappropriateness of tumor marker requests”, *Clin Chem Lab Med,* vol. 52, no. 6, pp. 889-97, Jun, 2014.

[58] M. K. Accordino, J. D. Wright, S. Vasan *et al.*, “Serum Tumor Marker Use in Patients With Advanced Solid Tumors”, *J Oncol Pract,* vol. 12, no. 1, pp. 65-6, e36-43, Jan, 2016.

[59] J. W. Dennis, M. Granovsky, and C. E. Warren, “Glycoprotein glycosylation and cancer progression”, *Biochim Biophys Acta,* vol. 1473, no. 1, pp. 21-34, Dec 6, 1999.

[60] S. Tao, Y. Huang, B. E. Boyes *et al.*, “Liquid chromatography-selected reaction monitoring (LC-SRM) approach for the separation and quantitation of sialylated N-glycans linkage isomers”, *Anal Chem,* vol. 86, no. 21, pp. 10584-90, Nov 4, 2014.

[61] A. Varki, Sharon, N., "Historical Background and Overview.Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY). 2009".

[62] S. S. Goreta, S. Dabelic, and J. Dumic, “Insights into complexity of congenital disorders of glycosylation”, *Biochem Med (Zagreb),* vol. 22, no. 2, pp. 156-70, 2012.

[63] J. M. Rhodes, “Unifying hypothesis for inflammatory bowel disease and associated colon cancer: sticking the pieces together with sugar”, *Lancet,* vol. 347, no. 8993, pp. 40-4, Jan 6, 1996.

[64] S. Mereiter, M. Balmana, J. Gomes *et al.*, “Glycomic Approaches for the Discovery of Targets in Gastrointestinal Cancer”, *Front Oncol,* vol. 6, pp. 55, 2016.

[65] K. Kumamoto, Y. Goto, K. Sekikawa *et al.*, “Increased expression of UDP-galactose transporter messenger RNA in human colon cancer tissues and its implication in synthesis of Thomsen-Friedenreich antigen and sialyl Lewis A/X determinants”, *Cancer Res,* vol. 61, no. 11, pp. 4620-7, Jun 1, 2001.

[66] L. Terraneo, L. Avagliano, A. Caretti *et al.*, “Expression of carbohydrate-antigen sialyl-Lewis a on colon cancer cells promotes

xenograft growth and angiogenesis in nude mice", *Int J Biochem Cell Biol,* vol. 45, no. 12, pp. 2796-800, Dec, 2013.

[67] E. M. Vestergaard, H. O. Hein, H. Meyer *et al.*, "Reference values and biological variation for tumor marker CA 19-9 in serum for different Lewis and secretor genotypes and evaluation of secretor and Lewis genotyping in a Caucasian population", *Clin Chem,* vol. 45, no. 1, pp. 54-61, Jan, 1999.

[68] M. Valli, A. Gallanti, S. Bozzaro *et al.*, "Beta-1,3-galactosyltransferase and alpha-1,2-fucosyltransferase involved in the biosynthesis of type-1-chain carbohydrate antigens in human colon adenocarcinoma cell lines", *Eur J Biochem,* vol. 256, no. 2, pp. 494-501, Sep 1, 1998.

[69] J. L. Magnani, "The discovery, biology, and drug development of sialyl Lea and sialyl Lex", *Arch Biochem Biophys,* vol. 426, no. 2, pp. 122-31, Jun 15, 2004.

[70] A. Cazet, S. Julien, M. Bobowski *et al.*, "Consequences of the expression of sialylated antigens in breast cancer", *Carbohydr Res,* vol. 345, no. 10, pp. 1377-83, Jul 2, 2010.

[71] B. W. Weston, K. M. Hiller, J. P. Mayben *et al.*, "Expression of human alpha(1,3)fucosyltransferase antisense sequences inhibits selectin-mediated adhesion and liver metastasis of colon carcinoma cells", *Cancer Res,* vol. 59, no. 9, pp. 2127-35, May 1, 1999.

[72] J. Holgersson, and J. Lofling, "Glycosyltransferases involved in type 1 chain and Lewis antigen biosynthesis exhibit glycan and core chain specificity", *Glycobiology,* vol. 16, no. 7, pp. 584-93, Jul, 2006.

[73] M. Soejima, and Y. Koda, "Molecular mechanisms of Lewis antigen expression", *Leg Med (Tokyo),* vol. 7, no. 4, pp. 266-9, Jul, 2005.

[74] H. Narimatsu, H. Iwasaki, F. Nakayama *et al.*, "Lewis and secretor gene dosages affect CA19-9 and DU-PAN-2 serum levels in normal

individuals and colorectal cancer patients", *Cancer Res,* vol. 58, no. 3, pp. 512-8, Feb 1, 1998.

[75] K. Miyazaki, K. Ohmori, M. Izawa *et al.*, "Loss of disialyl Lewis(a), the ligand for lymphocyte inhibitory receptor sialic acid-binding immunoglobulin-like lectin-7 (Siglec-7) associated with increased sialyl Lewis(a) expression on human colon cancers", *Cancer Res,* vol. 64, no. 13, pp. 4498-505, Jul 1, 2004.

[76] A. S. Carvalho, A. Harduin-Lepers, A. Magalhaes *et al.*, "Differential expression of alpha-2,3-sialyltransferases and alpha-1,3/4-fucosyltransferases regulates the levels of sialyl Lewis a and sialyl Lewis x in gastrointestinal carcinoma cells", *Int J Biochem Cell Biol,* vol. 42, no. 1, pp. 80-9, Jan, 2010.

[77] M. Perez-Garay, B. Arteta, L. Pages *et al.*, "alpha2,3-sialyltransferase ST3Gal III modulates pancreatic cancer cell motility and adhesion in vitro and enhances its metastatic potential in vivo", *PLoS One,* vol. 5, no. 9, 2010.

[78] S. Isshiki, A. Togayachi, T. Kudo *et al.*, "Cloning, expression, and characterization of a novel UDP-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase (beta3Gal-T5) responsible for synthesis of type 1 chain in colorectal and pancreatic epithelia and tumor cells derived therefrom", *J Biol Chem,* vol. 274, no. 18, pp. 12499-507, Apr 30, 1999.

[79] S. Isshiki, T. Kudo, S. Nishihara *et al.*, "Lewis type 1 antigen synthase (beta3Gal-T5) is transcriptionally regulated by homeoproteins", *J Biol Chem,* vol. 278, no. 38, pp. 36611-20, Sep 19, 2003.

[80] A. Seko, T. Ohkura, H. Kitamura *et al.*, "Quantitative differences in GlcNAc:beta1-->3 and GlcNAc:beta1-->4 galactosyltransferase activities between human colonic adenocarcinomas and normal

colonic mucosa", *Cancer Res,* vol. 56, no. 15, pp. 3468-73, Aug 1, 1996.

[81] E. H. Holmes, G. K. Ostrander, H. Clausen *et al.*, "Oncofetal expression of Lex carbohydrate antigens in human colonic adenocarcinomas. Regulation through type 2 core chain synthesis rather than fucosylation", *J Biol Chem,* vol. 262, no. 23, pp. 11331-8, Aug 15, 1987.

[82] T. Ichikawa, J. Nakayama, N. Sakura *et al.*, "Expression of N-acetyllactosamine and beta1,4-galactosyltransferase (beta4GalT-I) during adenoma-carcinoma sequence in the human colorectum", *J Histochem Cytochem,* vol. 47, no. 12, pp. 1593-602, Dec, 1999.

[83] W. S. Chen, H. Y. Chang, C. P. Li *et al.*, "Tumor beta-1,4-galactosyltransferase IV overexpression is closely associated with colorectal cancer metastasis and poor prognosis", *Clin Cancer Res,* vol. 11, no. 24 Pt 1, pp. 8615-22, Dec 15, 2005.

[84] A. Bardoni, M. Valli, and M. Trinchera, "Differential expression of beta1,3galactosyltransferases in human colon cells derived from adenocarcinomas or normal mucosa", *FEBS Lett,* vol. 451, no. 1, pp. 75-80, May 14, 1999.

[85] L. Mare, and M. Trinchera, "Suppression of beta 1,3galactosyltransferase beta 3Gal-T5 in cancer cells reduces sialyl-Lewis a and enhances poly N-acetyllactosamines and sialyl-Lewis x on O-glycans", *Eur J Biochem,* vol. 271, no. 1, pp. 186-94, Jan, 2004.

[86] H. Laubli, and L. Borsig, "Selectins promote tumor metastasis", *Semin Cancer Biol,* vol. 20, no. 3, pp. 169-77, Jun, 2010.

[87] A. F. Chambers, G. N. Naumov, H. J. Varghese *et al.*, "Critical steps in hematogenous metastasis: an overview", *Surg Oncol Clin N Am,* vol. 10, no. 2, pp. 243-55, vii, Apr, 2001.

[88] R. Kannagi, M. Izawa, T. Koike *et al.*, “Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis”, *Cancer Sci,* vol. 95, no. 5, pp. 377-84, May, 2004.

[89] M. Fukuda, “Possible roles of tumor-associated carbohydrate antigens”, *Cancer Res,* vol. 56, no. 10, pp. 2237-44, May 15, 1996.

[90] M. Ugorski, and A. Laskowska, “Sialyl Lewis(a): a tumor-associated carbohydrate antigen involved in adhesion and metastatic potential of cancer cells”, *Acta Biochim Pol,* vol. 49, no. 2, pp. 303-11, 2002.

[91] Y. Kawarada, H. Ishikura, T. Kishimoto *et al.*, “The role of sialylated Lewis antigens on hematogenous metastases of human pancreas carcinoma cell lines in vivo”, *Pathol Res Pract,* vol. 196, no. 4, pp. 259-63, 2000.

[92] A. Tsuchida, T. Okajima, K. Furukawa *et al.*, “Synthesis of disialyl Lewis a (Le(a)) structure in colon cancer cell lines by a sialyltransferase, ST6GalNAc VI, responsible for the synthesis of alpha-series gangliosides”, *J Biol Chem,* vol. 278, no. 25, pp. 22787-94, Jun 20, 2003.

[93] Y. Suzuki, T. Ichihara, A. Nakao *et al.*, “High serum levels of DUPAN2 antigen and CA19-9 in pancreatic cancer: correlation with immunocytochemical localization of antigens in cancer cells”, *Hepatogastroenterology,* vol. 35, no. 3, pp. 128-35, Jun, 1988.

[94] S. V. Portela, C. V. Martin, L. M. Romay *et al.*, “sLea and sLex expression in colorectal cancer: implications for tumourigenesis and disease prognosis”, *Histol Histopathol,* vol. 26, no. 10, pp. 1305-16, Oct, 2011.

[95] G. Ohshio, K. Ogawa, H. Kudo *et al.*, “Immunohistochemical studies on the localization of cancer associated antigens DU-PAN-2 and CA19-9 in carcinomas of the digestive tract”, *J Gastroenterol Hepatol,* vol. 5, no. 1, pp. 25-31, Jan-Feb, 1990.

[96] Y. Ikeda, M. Mori, A. Kido *et al.*, “Immunohistochemical expression of carbohydrate antigen 19-9 in gastric carcinoma”, *Am J Gastroenterol,* vol. 86, no. 9, pp. 1163-6, Sep, 1991.

[97] B. C. Wolf, R. R. Salem, H. F. Sears *et al.*, “The expression of colorectal carcinoma-associated antigens in the normal colonic mucosa. An immunohistochemical analysis of regional distribution”, *Am J Pathol,* vol. 135, no. 1, pp. 111-9, Jul, 1989.

[98] F. Dupuy, A. Germot, M. Marenda *et al.*, “Alpha1,4-fucosyltransferase activity: a significant function in the primate lineage has appeared twice independently”, *Mol Biol Evol,* vol. 19, no. 6, pp. 815-24, Jun, 2002.

[99] V. B. Chachadi, M. F. Ali, and P. W. Cheng, “Prostatic cell-specific regulation of the synthesis of MUC1-associated sialyl Lewis a”, *PLoS One,* vol. 8, no. 2, pp. e57416, 2013.

[100] G. Zhang, and S. Pradhan, “Mammalian epigenetic mechanisms”, *IUBMB Life,* vol. 66, no. 4, pp. 240-56, Apr, 2014.

[101] P. A. Jones, and S. B. Baylin, “The epigenomics of cancer”, *Cell,* vol. 128, no. 4, pp. 683-92, Feb 23, 2007.

[102] S. Sharma, T. K. Kelly, and P. A. Jones, “Epigenetics in cancer”, *Carcinogenesis,* vol. 31, no. 1, pp. 27-36, Jan, 2010.

[103] Y. Okugawa, W. M. Grady, and A. Goel, “Epigenetic Alterations in Colorectal Cancer: Emerging Biomarkers”, *Gastroenterology,* vol. 149, no. 5, pp. 1204-1225 e12, Oct, 2015.

[104] A. P. Feinberg, and B. Vogelstein, “Hypomethylation distinguishes genes of some human cancers from their normal counterparts”, *Nature,* vol. 301, no. 5895, pp. 89-92, Jan 6, 1983.

[105] G. Lauc, A. Vojta, and V. Zoldos, “Epigenetic regulation of glycosylation is the quantum mechanics of biology”, *Biochim Biophys Acta,* vol. 1840, no. 1, pp. 65-70, Jan, 2014.

[106] J. Serpa, P. Mesquita, N. Mendes *et al.*, “Expression of Lea in gastric cancer cell lines depends on FUT3 expression regulated by promoter methylation”, *Cancer Lett,* vol. 242, no. 2, pp. 191-7, Oct 28, 2006.

[107] V. B. Chachadi, H. Cheng, D. Klinkebiel *et al.*, “5-Aza-2'-deoxycytidine increases sialyl Lewis X on MUC1 by stimulating beta-galactoside:alpha2,3-sialyltransferase 6 gene”, *Int J Biochem Cell Biol,* vol. 43, no. 4, pp. 586-93, Apr, 2011.

[108] M. L. Poeta, E. Massi, P. Parrella *et al.*, “Aberrant promoter methylation of beta-1,4 galactosyltransferase 1 as potential cancer-specific biomarker of colorectal tumors”, *Genes Chromosomes Cancer,* vol. 51, no. 12, pp. 1133-43, Dec, 2012.

[109] M. Trinchera, A. Zulueta, A. Caretti *et al.*, “Control of Glycosylation-Related Genes by DNA Methylation: the Intriguing Case of the B3GALT5 Gene and Its Distinct Promoters”, *Biology (Basel),* vol. 3, no. 3, pp. 484-97, 2014.

[110] Q. Du, P. L. Luu, C. Stirzaker *et al.*, “Methyl-CpG-binding domain proteins: readers of the epigenome”, *Epigenomics,* vol. 7, no. 6, pp. 1051-73, 2015.

[111] C. Vinson, and R. Chatterjee, “CG methylation”, *Epigenomics,* vol. 4, no. 6, pp. 655-63, Dec, 2012.

[112] M. Weber, I. Hellmann, M. B. Stadler *et al.*, “Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome”, *Nat Genet,* vol. 39, no. 4, pp. 457-66, Apr, 2007.

[113] T. S. Mikkelsen, M. Ku, D. B. Jaffe *et al.*, “Genome-wide maps of chromatin state in pluripotent and lineage-committed cells”, *Nature,* vol. 448, no. 7153, pp. 553-60, Aug 2, 2007.

[114] R. Chatterjee, and C. Vinson, “CpG methylation recruits sequence specific transcription factors essential for tissue specific gene

expression", *Biochim Biophys Acta,* vol. 1819, no. 7, pp. 763-70, Jul, 2012.

[115] J. A. Hackett, and M. A. Surani, "DNA methylation dynamics during the mammalian life cycle", *Philos Trans R Soc Lond B Biol Sci,* vol. 368, no. 1609, pp. 20110328, Jan 5, 2013.

[116] M. A. Dawson, and T. Kouzarides, "Cancer epigenetics: from mechanism to therapy", *Cell,* vol. 150, no. 1, pp. 12-27, Jul 6, 2012.

[117] R. J. Klose, and A. P. Bird, "Genomic DNA methylation: the mark and its mediators", *Trends Biochem Sci,* vol. 31, no. 2, pp. 89-97, Feb, 2006.

[118] K. Bardhan, and K. Liu, "Epigenetics and colorectal cancer pathogenesis", *Cancers (Basel),* vol. 5, no. 2, pp. 676-713, 2013.

[119] C. Stirzaker, P. C. Taberlay, A. L. Statham *et al.*, "Mining cancer methylomes: prospects and challenges", *Trends Genet,* vol. 30, no. 2, pp. 75-84, Feb, 2014.

[120] R. A. Irizarry, C. Ladd-Acosta, B. Wen *et al.*, "The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores", *Nat Genet,* vol. 41, no. 2, pp. 178-86, Feb, 2009.

[121] A. Doi, I. H. Park, B. Wen *et al.*, "Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts", *Nat Genet,* vol. 41, no. 12, pp. 1350-3, Dec, 2009.

[122] A. Hellman, and A. Chess, "Gene body-specific methylation on the active X chromosome", *Science,* vol. 315, no. 5815, pp. 1141-3, Feb 23, 2007.

[123] M. A. Hahn, X. Wu, A. X. Li *et al.*, "Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks", *PLoS One,* vol. 6, no. 4, pp. e18844, 2011.

[124] M. Kulis, A. C. Queiros, R. Beekman *et al.*, “Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer”, *Biochim Biophys Acta,* vol. 1829, no. 11, pp. 1161-74, Nov, 2013.

[125] A. K. Maunakea, R. P. Nagarajan, M. Bilenky *et al.*, “Conserved role of intragenic DNA methylation in regulating alternative promoters”, *Nature,* vol. 466, no. 7303, pp. 253-7, Jul 8, 2010.

[126] K. E. Varley, J. Gertz, K. M. Bowling *et al.*, “Dynamic DNA methylation across diverse human cell lines and tissues”, *Genome Res,* vol. 23, no. 3, pp. 555-67, Mar, 2013.

[127] J. O. Haerter, C. Lovkvist, I. B. Dodd *et al.*, “Collaboration between CpG sites is needed for stable somatic inheritance of DNA methylation states”, *Nucleic Acids Res,* vol. 42, no. 4, pp. 2235-44, Feb, 2014.

[128] C. Lovkvist, I. B. Dodd, K. Sneppen *et al.*, “DNA methylation in human epigenomes depends on local topology of CpG sites”, *Nucleic Acids Res,* vol. 44, no. 11, pp. 5123-32, Jun 20, 2016.

[129] C. A. Dunn, P. Medstrand, and D. L. Mager, “An endogenous retroviral long terminal repeat is the dominant promoter for human beta1,3-galactosyltransferase 5 in the colon”, *Proc Natl Acad Sci U S A,* vol. 100, no. 22, pp. 12841-6, Oct 28, 2003.

[130] L. Mare, and M. Trinchera, “Comparative analysis of retroviral and native promoters driving expression of beta1,3-galactosyltransferase beta3Gal-T5 in human and mouse tissues”, *J Biol Chem,* vol. 282, no. 1, pp. 49-57, Jan 5, 2007.

[131] A. Caretti, S. M. Sirchia, S. Tabano *et al.*, “DNA methylation and histone modifications modulate the beta1,3 galactosyltransferase beta3Gal-T5 native promoter in cancer cells”, *Int J Biochem Cell Biol,* vol. 44, no. 1, pp. 84-90, Jan, 2012.

[132] C. A. Dunn, L. N. van de Lagemaat, G. J. Baillie *et al.*, “Endogenous retrovirus long terminal repeats as ready-to-use mobile promoters: the case of primate beta3GAL-T5”, *Gene,* vol. 364, pp. 2-12, Dec 30, 2005.

[133] A. Zulueta, A. Caretti, P. Signorelli *et al.*, “Transcriptional control of the B3GALT5 gene by a retroviral promoter and methylation of distant regulatory elements”, *FASEB J,* vol. 28, no. 2, pp. 946-55, Feb, 2014.

[134] A. Caretti, P. Bianciardi, G. Sala *et al.*, “Supplementation of creatine and ribose prevents apoptosis in ischemic cardiomyocytes”, *Cell Physiol Biochem,* vol. 26, no. 6, pp. 831-8, 2010.

[135] L. Avagliano, E. Virgili, C. Garo *et al.*, “Autophagy and human parturition: evaluation of LC3 expression in placenta from spontaneous or medically induced onset of labor”, *Biomed Res Int,* vol. 2013, pp. 689768, 2013.

[136] M. D. Girgis, V. Kenanova, T. Olafsen *et al.*, “Anti-CA19-9 diabody as a PET imaging probe for pancreas cancer”, *J Surg Res,* vol. 170, no. 2, pp. 169-78, Oct, 2011.

[137] R. Salvini, A. Bardoni, M. Valli *et al.*, “beta 1,3-Galactosyltransferase beta 3Gal-T5 acts on the GlcNAcbeta 1-->3Galbeta 1-->4GlcNAcbeta 1-->R sugar chains of carcinoembryonic antigen and other N-linked glycoproteins and is down-regulated in colon adenocarcinomas”, *J Biol Chem,* vol. 276, no. 5, pp. 3564-73, Feb 2, 2001.

[138] P. G. Falk, L. Bry, J. Holgersson *et al.*, “Expression of a human alpha-1,3/4-fucosyltransferase in the pit cell lineage of FVB/N mouse stomach results in production of Leb-containing glycoconjugates: a potential transgenic mouse model for studying Helicobacter pylori infection”, *Proc Natl Acad Sci U S A,* vol. 92, no. 5, pp. 1515-9, Feb 28, 1995.

[139] T. Kudo, Y. Ikehara, A. Togayachi *et al.*, “Up-regulation of a set of glycosyltransferase genes in human colorectal cancer”, *Lab Invest,* vol. 78, no. 7, pp. 797-811, Jul, 1998.

[140] D. Baeckstrom, N. Karlsson, and G. C. Hansson, “Purification and characterization of sialyl-Le(a)-carrying mucins of human bile; evidence for the presence of MUC1 and MUC3 apoproteins”, *J Biol Chem,* vol. 269, no. 20, pp. 14430-7, May 20, 1994.

[141] T. Yue, K. Partyka, K. A. Maupin *et al.*, “Identification of blood-protein carriers of the CA 19-9 antigen and characterization of prevalence in pancreatic diseases”, *Proteomics,* vol. 11, no. 18, pp. 3665-74, Sep, 2011.

[142] G. Kloppel, G. Lingenthal, M. von Bulow *et al.*, “Histological and fine structural features of pancreatic ductal adenocarcinomas in relation to growth and prognosis: studies in xenografted tumours and clinico-histopathological correlation in a series of 75 cases”, *Histopathology,* vol. 9, no. 8, pp. 841-56, Aug, 1985.

[143] H. Kalthoff, C. Kreiker, W. H. Schmiegel *et al.*, “Characterization of CA 19-9 bearing mucins as physiological exocrine pancreatic secretion products”, *Cancer Res,* vol. 46, no. 7, pp. 3605-7, Jul, 1986.

8. SCIENTIFIC PRODUCTS

- ✓ **Publication on the thesis topic:**

  Aronica A., Avagliano L., Caretti A., Tosi D., Bulfamante G.P., Trinchera M. "Unexpected distribution of CA19.9 and other type 1 chain Lewis antigens in normal and cancer tissues of colon and pancreas: Importance of the detection method and role of glycosyltransferase regulation". Biochim Biophys Acta. 2016 Aug 14. pii: S0304-4165(16)30286-0. doi: 10.1016/j.bbagen.2016.08.005. [Epub ahead of print].

- ✓ **Poster presentations:**

  1) CA19.9 is not a tumor marker: a novel strategy for using Lewis antigens in the approach to gastrointestinal cancers. I Congress of DiSS, 2015 November 13.

  2) Mechanisms of B3GALT5 gene silencing in cancer. I Congress of DiSS, 2015 November 13.

- ✓ **Oral presentation:**

  CA19.9 and type 1 chain Lewis antigens: unraveling the molecular basis of expression in gastrointestinal tissues to improve the clinical effectiveness as tumor markers.II Congress of DiSS, 2016 November 11.

9. ACKNOWLEDGMENTS

The work reported on this thesis was performed at the Department of Health Sciences, Laboratory of Biochemistry, San Paolo Hospital, thanks to the financial support of a PhD program in “Molecular and Translational Medicine” of the University of Milan, Italy.

I would like to thank my supervisor, Prof. Riccardo Ghidoni, for his helpful advices and for giving me the opportunity to carry out my PhD studies in his research group.

Special thanks go to my co-supervisor, Dr. Marco Trinchera, whose patience, enthusiasm, constant support and constructive advices, have been of invaluable help for me.

I want to thank everyone in the “biochemistry lab” for being great colleagues and bringing joy into the daily work.

I would like to thank my wonderful family and lovely friends for all their love, support and patience.

Finally, I want to thank God…I never would have made it without him!