UNIVERSITÀ DEGLI STUDI DI MILANO
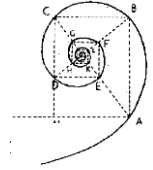
SCUOLA DI DOTTORATO IN MEDICINA MOLECOLARE

CICLO XXVI

Anno Accademico 2012/2013

TESI DI DOTTORATO DI RICERCA

BIO/10

Transcriptional regulation of the B3GALT5 gene

Dottorando Dr.ssa Aída Zulueta

Matricola R09232

TUTORE : Prof. Riccardo Ghidoni

CO-TUTORE: Dr. Marco Trinchera

DIRETTORE DEL DOTTORATO: Prof. Mario Clerici

A Diego,

que cuando mires dentro de ti, y de lo que haces,
te encuentres como la tierra por la mañana,
bañada de luz....
José Martí

ABSTRACT

Background. β1,3 galactosyltransferase (B3GALT5) is responsible for the synthesis of type 1 chain oligosaccharides, including Lewis antigens as sialyl-Lewis a, the epitope of tumor marker CA19.9 and an E-selectin ligand potentially involved in cancer malignancy. Transcription occurs through multiple promoters. In some epithelia it is driven by a weak promoter, known as the native promoter that is epigenetically modulated and sensitive to nuclear factor NF-Y. In some organs of the gastrointestinal tract (as the colon, stomach, pancreas and related cell lines) another stronger promoter is active and named the LTR promoter after its retroviral origin. It was supposed to be regulated through a set of homeoproteins: hepatocyte nuclear factor HNF1α/β and caudal-related homeobox Cdx1/2. Surprisingly, *B3GALT5* is strongly down regulated in colon cancer, the LTR transcript is not relevant in the small intestine, and Cdx1/2 were reported absent from a cell line expressing large amount of such transcript.

Aims. To elucidate the mechanisms controlling transcription of *B3GALT5* through its retroviral LTR promoter, in order to explain the tissue specificity and down-regulation in colon adenocarcinomas, and to understand the evolutionary stabilization of the transposon in some primates.

Methods. To this aim, we determined the expression levels of putative transcription factors by western blot and the amounts of B3GALT5 LTR transcript by competitive RT-PCR in cancer tissues and cell lines. Moreover, we silenced HNF1α or β in different cell lines, through an shRNA approach, expressed them in another by permanent cDNA transfection, and treated cells with the DNA demethylating agent 5'-AZA-2'-deoxycitydine and in all cases, we measure the effects on LTR transcript levels. We also evaluated the behavior of the LTR promoter *in vitro*, through electrophoresis mobility shift and reporter luciferase assays.

Results. We found that Cdx1/2 are not detectable in cells and tissues expressing high amount of B3GALT5 LTR transcript, while HNF1α/β are well detectable, but even in cells and cancers expressing very low or undetectable levels of the transcript, which is absent in all cells lacking HNF1α/β. Among them, the cell line MDA-MB-231, upon transfection with HNF1α or β, became able to express B3GALT5 LTR transcript, but a very low levels, similar to those found in colon cancers. Transient silencing of HNF1α in cells expressing both HNF1α and β, has no effect on LTR transcript, while similar silencing of HNF1β in cells expressing HNF1β only, determines strong reduction of the transcript. Cell lines expressing high levels of B3GALT5 LTR transcript are affected by the demethylating agent 5AZA that determines strong down regulation of the transcript, falling down to the amounts found in colon cancers, while HNF1 levels remain unaffected. *In vitro*, luciferase placed under the control of LTR promoter is more active in cells or clones expressing high HNF1 and low or no LTR transcript than in those expressing low HNF1 and high transcript. The same promoter, when used as a probe in EMSA, forms specific complex with nuclear protein extracted from all cells expressing HNF1, irrespectively of the levels of B3GALT5 LTR transcript.

Conclusion. Our results suggest that HNF1α and HNF1β are necessary but not sufficient to drive expression of LTR promoter, while Cdx1/2 are not involved. HNF1α/β play an interchangeable and not cumulative role and are not immediately responsible for cancer down-regulation, which depends on a distal regulatory element(s) active when methylated. The successful insertion and activation of B3GALT5 LTR promoter during evolution depended not only on its HNF1 binding site, but even on such distal element(s) unknown at present.

RIASSUNTO

Introduzione. La β1,3 galattosiltransferasi (B3GALT5) è responsabile della sintesi della catena oligosaccaridica di tipo 1, tra cui gli antigeni Lewis come il sialil-Lewis a, epitope del marcatore tumorale CA19.9 e ligando della E-selectina, potenzialmente coinvolto nella malignità tumorale. La sua trascrizione è regolata da molteplici promotori. In alcuni epiteli essa è sotto il controllo di un promotore debole, chiamato nativo, modulato epigeneticamente e tramite il fattore nucleare NF-Y. In alcuni organi e cellule di origine gastrointestinale è attivo inoltre un altro promotore, più forte e chiamato LTR per la sua origine retrovirale, che secondo la letteratura dovrebbe essere regolato attraverso il fattore nucleare epatocitario HNF1α/β e quello analogo al *Caudale* di drosofila Cdx1/2. Tuttavia la B3GALT5 è repressa nel cancro del colon, il trascritto LTR non è rilevante nell'intestino tenue, e Cdx1/2 risultano assenti in una linea cellulare che esprime grandi quantità di tale trascritto.

Scopi. Scoprire i meccanismi che controllano la trascrizione di *B3GALT5* attraverso il suo promotore LTR, con l'obiettivo di spiegare la specificità tissutale e la repressione negli adenocarcinomi del colon, nonché di capire il processo di stabilizzazione evolutiva del trasposone in alcuni primati.

Metodi. A questo scopo abbiamo quantificato l'espressione di HNF1α/β e Cdx1/2, tramite Western Blot, e quella del trascritto B3GALT5 LTR, tramite RT-PCR competitiva, in tessuti tumorali e linee cellulari. Inoltre, abbiamo silenziato HNF1α o β in diverse cellule, tramite shRNA, e li abbiamo espressi in un'altra, tramite trasfezione del cDNA. Abbiamo quindi trattato delle cellule con l'agente demetilante il DNA 5'-AZA-2'-desossicitidina, misurandone poi i livelli di trascritto LTR. Abbiamo infine valutato il promotore LTR *in vitro*, mediante saggi di EMSA e di luciferasi.

Risultati. Cdx1/2 risultano immisurabili in cellule e tessuti che pur

esprimono alti livelli di trascritto LTR, mentre HNF1α/β sono presenti anche in cellule e tumori che esprimono una quantità bassa o nulla del trascritto. Nelle cellule che non esprimono HNF1α/β però non c'è espressione alcuna del trascritto LTR. Tra queste, le MDA-MB-231, dopo transfezione con HNF1α o β, esprimono il trascritto LTR, ma a bassi livelli, paragonabili a quelli dei tumori del colon. Il silenziamento di HNF1α in una linea cellulare che esprime entrambi i fattori HNF1α e β non ha effetti nell'espressione del trascritto LTR, mentre quello di HNF1β in un'altra linea che esprime solo HNF1β produce forte riduzione del trascritto. Pure il trattamento con 5'-AZA-2'-desossicitidina riduce il trascritto in cellule che ne esprimono alti livelli, portandolo ai livelli dei tumori del colon, e senza effetto sulla quantità di HNF1. Usando i saggi delle luciferasi, abbiamo visto che la luciferasi sotto il controllo del promotore LTR è più attiva in cellule e cloni che esprimono alte quantità di HNF1, anche se non esprimono o esprimono poco trascritto LTR, che in quelle cellule che esprimono poco HNF1 ma magari una grande quantità di trascritto. Usando la sequenza del promotore LTR in saggi di EMSA, abbiamo visto che forma dei complessi specifici con estratti di proteine nucleari di tutte le linee cellulari che esprimono HNF1, indipendentemente dei livelli di espressione del trascritto B3GALT5 LTR.

Conclusione. I nostri risultati suggeriscono che HNF1α e -β sono necessari ma non sufficienti a regolare l'espressione del promotore LTR, mentre Cdx1/2 non sono coinvolti. HNF1α/β svolgono un ruolo intercambiabile e non cumulativo, e non sono immediatamente responsabili della regolazione negativa che avviene nel cancro, che invece dipende da elementi regolatori distanti attivi solo se metilati. Il successo della inserzione e l'attivazione del promotore B3GALT5 LTR durante l'evoluzione dipendono quindi non solo dal suo sito di legame a HNF1, ma anche da questi elementi distanti attualmente sconosciuti.

INDEX

SIMBOLS LIST

B3GALT5	β1,3 galactosyltransferase 5
Le(a), Lewis a	Galβ1-3[Fucα1-4]GlcNAc
Le(b), Lewis b	Fucα1-2Galβ1-3[Fucα1-4]GlcNAc
sLe(a), sialyl-Lewis a	NeuAcα2-3Galβ1-3[Fucα1-4]GlcNAc
GlcNAc	N-acetylglucosamine
B3GALTs	β1,3-galactosyltransferase
Gal	Galactose
Lacto-N-biose	Galβ1-3GlcNAc
ST3GALTs	Galactose-α2,3-sialyltransferases
sialyl-lacto-N-biose	NeuAc-α2-3Galβ1-3GlcNAc
FUT	fucosyltransferase
Fuc	Fucose
GalNAc	N-acetylgalactosamine
LacNAc	N-acetyllactosamine
sLe(x), sialyl-Lewis x	NeuAcα2-3Galβ1-4[Fucα1-3]GlcNAc
GnT5	β1,6 N-acetylglucosaminyltransferase V
MUC	Mucin
LAMP	Lysosomal membrane glycoproteins
NK	Natural killer cells
SabA	pylori sialic acid-binding adhesin
HNF1	Hepatocyte nuclear factor 1
Cdx	Caudal related homeobox
LTR	Long terminal repeat
UTR	Untranslated region
NF-Y	CCAAT-binding factor
TE	Transposable elements
ERV	Endogenous retrovirus
LINE	Long interspersed nuclear elements

SINE	Short interspersed nuclear elements
MITE	Miniature inverted-repeat transposable elements
TFBS	Transcription factor binding site
CTCF	CCCTC-binding factor
Pou5F1	POU domain, class 5, transcription factor 1
Sox2	SRY-box containing gene 2
ESR1	Estrogen receptor 1
GBP	Guanylate binding protein
BAAT	Bile acid CoA: amino acid N-acyltransferase
MSLN	Mesothelin
ADH1C	Alcohol dehydrogenase 1C
HSD17B1	Hydroxysteroid 17-beta dehydrogenase 1
CYP19A1 polypeptide 1	Cytochrome P450, family 19, subfamily A
IL2RB	Interleukin 2 receptor, beta
NOS3	Nitric oxide synthase 3
PTN	Pleiotrophin
PAPPA2	Pappalysin 2
EDNRB	Endothelin receptor type B
ENTPD	Ectonucleoside triphosphate diphosphohydrolase
MID1	Midline 1
SEMA4	Semaphorin IV
APOCI	Apolipoprotein C-I
MBD	Methyl-CpG binding domain
ncRNA	non-coding RNA
miRNA	micro RNA
lncRNA	long non-coding RNA
LASS6	Ceramide synthase 6

eRNA	non-coding RNA from enhancer sequences
MODY	Maturity-onset diabetes mellitus
GATA-4	GATA binding protein 4
RT-PCR	reverse transcription-polymerase chain reaction
shRNA	small hairpin RNA
TSA	Trichostatin A
5AZA	5'-aza-2'-deoxycytidine
EMSA	Electrophoresis mobility shift assay
LPH	Lactase-phlorizin hydrolase

FIGURES LIST

1. INTRODUCTION

B3GALT5 gene codes for β1,3 galactosyltransferase 5 (B3GALT5), an enzyme responsible for the synthesis of type 1 chain carbohydrates in mammals. In humans, in particular, it participates in the biosynthesis of the histo-blood group antigens Lewis a (Le(a)), Lewis b (Le(b)) and sialyl-Lewis A (sLe(a)) [1]. sLe(a) is the tetrasaccharide antigenic epitope (NeuAcα2-3Galβ1-3[Fucα1-4]GlcNAc) that constitutes the epitope of CA19.9 antigen [2, 3], a tumor marker widely used in the clinical practice.

1.1 B3GALT5 in the biosynthesis of Lewis-type antigens

At least four glycosyltransferases are required for the synthesis of the sLe(a) epitope (Figure 1.1). First, one of the N-acetylglucosaminyltransferase enzymes catalyses the addition of N-acetylglucosamine (GlcNAc) to N- or O-linked chains [4] of glycoproteins, or even to glycolipids [5]. After that, N-acetylglucosamine-β1,3-galactosyltransferase (B3GALTs) transfers a galactose (Gal) to GlcNAc with a β1,3-linkage, resulting in the synthesis of the type 1 chain, Galβ1-3GlcNAc (lacto-N-biose), and then one of Galactose-α2,3-sialyltransferases (ST3GALTs) transfers sialic acid to the Gal residue of the type 1 chain with an α2,3- linkage, resulting in the synthesis of sialyl-type 1 chain, NeuAc-α2-3Galβ1-3GlcNAc (sialyl-lacto-N-biose). Finally, α1,3/4-fucosyltransferase (FUTIII) transfers fucose (Fuc) to the GlcNAc residue of the sialyl-type 1 chain with an α1,4-linkage to complete the structure NeuAc-α2-3Galβ1-3[Fucα1-4]GlcNAc. Le(a) and Le(b) antigens are originated by the mono- or di- fucosyl substitution of lacto-N-biose, respectively.
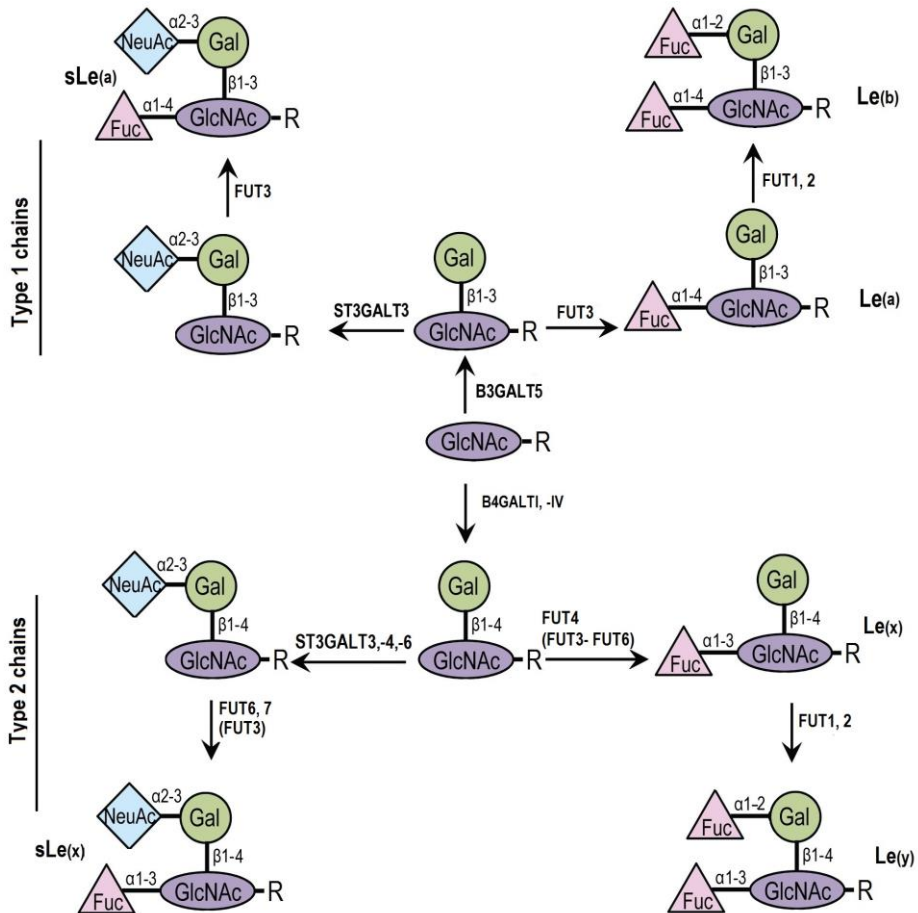
Figure 1.1 Biosynthetic pathway and structures of Lewis-type antigens. Substitution of GlcNAc by a β1,3-linked galactose leads to lacto-N-biose, the basic unit of type 1 chain, while substitution with a β1,4-linked galactose leads to N-acetyllactosamine (LacNAc), the basic unit of type 2 chain. The addition of a fucose linked either via α1,4 or α1,3 to N-acetylglucosamine in type 1 or 2 chains respectively, leads to the formation of Le(a) and Le(x), respectively. These antigens can be further elongated by a α1,2 fucose, leading to the formation of Le(b) and Le(y) antigens, respectively. 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 Le(a) and sialyl Le(x) antigens, respectively.

The first cloning of a B3GALT gene was performed in 1996 starting from human WM266–4 melanoma cells and using an expression cloning approach [6]. Later on, three putatively novel human B3GALT genes homologous to the original one were cloned [7, 8]. Such B3GALTs were named B3GALT1 to -T4 [8]. Expression studies on the four human B3GALTs demonstrated that two of them, B3GALT1 and T2, transfer Gal to GlcNAc in a β1,3-linkage, resulting in type 1 chain synthesis, while B3GALT4 transfers Gal to an N-acetylgalactosamine (GalNAc) residue, resulting in the synthesis of the type 3 chain, Galβ1-3GalNAc [8]. The activity of human B3GALT3 has not been detected toward any of the acceptor substrates used [8]. The tissue distribution of the four B3GALTs was determined by Northern analysis [7, 8], and it was found that neither B3GALT1 nor -T2 is expressed in the pancreas, which indicated that there was at least another unknown B3GALT synthesizing the type 1 chain in the pancreas. In 1999, the B3GALT5 cDNA was first cloned from the human colon adenocarcinoma cell line COLO-205 by Isshiki *et al.*, and the enzyme was proposed as the most probable candidate participating in the synthesis of the sLe(a) epitope in gastrointestinal and pancreatic cells [1].

1.2. Role of Sialyl Lewis antigens

Substitution of GlcNAc by a β1,4-linked galactose (instead of a β1,3-linked galactose) leads to LacNAc, the basic unit of type 2 chains. The α2,3-sialylation of type 2 chain, followed by the addition of α1,3-linked fucose, leads to the biosynthesis of sialyl Lewis x (sLe(x)) antigen [9]. An aberrant expression of Lewis-type antigens appears to be a general cancer-associated phenomenon, reported in carcinomas of the lung [10], colon [5, 11, 12], stomach [13] and kidney [14]. sLe(a) is the epitope of CA19.9 antigen, which is present in human serum and found elevated in various diseases [15-17] including cancers of the digestive tract, as pancreas [16,

18], bile ducts [19], stomach [20], and colon [17, 21]. The interest in the expression of sialyl Lewis antigens in cancer increased enormously after the discovery that sLe(x) and sLe(a) acted as ligands for E- and P-selectin cell adhesion molecules expressed on activated endothelial cells [22, 23]. The physiological role of E- and P-selectins is to mediate leukocyte extravasation at the sites of tissue damage or injury [24]. However, these molecules may also regulate the metastatic cascade by forming emboli of cancer cells and platelets and favouring their arrest on endothelia [25-28]. The relationship between expression of sialylated Lewis antigens and hematogenous metastasis is suggested by many clinical and experimental studies [29] (Figure 1.2). In colon cancer patients, increased expression of sLe(x) and sLe(a) antigens correlated with metastasis and poor survival [30, 31]. sLe(x) correlated with malignancy also in renal cell carcinoma [32] and breast cancer [33] although in the latter the survival did not appear to be related with sLe(x) expression [34].

sLe(x) antigen structures are usually present at the terminal non-reducing end of polylactosaminic chains of glycoproteins and glycolipids [4, 5], preferentially mounted on the β1,6-branching in the case of N-linked glycoproteins. The β1,6-branching of N-linked chains consists in the addition of an antenna whose first GlcNAc is β1,6-linked to a core mannose residue. This antenna is preferentially elongated by polylactosaminic sequences and is frequently terminated by antigens of the Lewis type. Although the association of β1,6-branching with metastasis has long been known [35], the conclusive evidence about the causative role played by these structure in metastasis formation came from studies in mice in which the enzyme responsible for this modification, β1,6 N-acetylglucosaminyltransferase V (GnT5 product of the Mgat5 gene), was knocked down (Mgat5-/-). Mice expressing the polyomavirus middle T antigen (PyMT) from a transgene in mammary epithelium, spontaneously

develop mammary tumors. When these mice were crossed with Mgat5-/-
mice, the tumors grew slower than in the PyMT-transgenic littermate
expressing Mgat5 and metastasis formation was almost completely
inhibited [36]. The relationship between β1,6-branching and increased
growth and metastasis is probably due to more than one mechanism [37].

In some cell lines, the major glycoproteins carrying sialyl Lewis antigens
have been identified as the hyaluronate receptor CD44 [26, 38, 39], mucin
1 (MUC1) [40-42] and lysosomal membrane glycoproteins 1 and 2 (LAMP-1
and LAMP-2) [42].

Several studies have reported that downregulation of sialyl Lewis antigen
expression by knockdown of key glycosyltransferases in cancer cell lines
resulted in reduced selectin binding and reduced metastatic ability [43-46],
while cancer cells forced to express sialyl Lewis antigens by gene transfer
exhibited increased adhesion to selectins *in vitro* and increased metastatic
ability in vivo [47]. Consistently, populations of cancer cells selected for
their increased metastatic potential often displayed increased expression of
sialyl Lewis antigens [48, 49]. The role of selectins in the metastatic
process was confirmed by the findings that the formation of experimental
pulmonary metastases could be inhibited by the use of peptides mimicking
sLe(a) and were inhibited in E-selectin- knock-out mice [50].

Apart from the role as selectin ligands, sialyl Lewis antigens can play a role
in cancer progression in at least two other key steps of invasion:
angiogenesis and immune recognition of cancer cells. The role of sLe(x) in
angiogenesis is supported by the finding that when epidermoid cancer cells
were co-cultured with endothelial cells, the former produced nests of
growing cells surrounded by tube-like networks consisting of endothelial
cells. These phenomena could be reproduced in vivo and could be inhibited
by antibodies against sLe(x) [51]. The ability of sLe(x)-expressing cancer
cells to promote angiogenesis was confirmed by the fact that inhibition of

sLe(x) biosynthesis in hepatocarcinoma HepG2 cells resulted an impairment of their ability to induce angiogenesis [52].
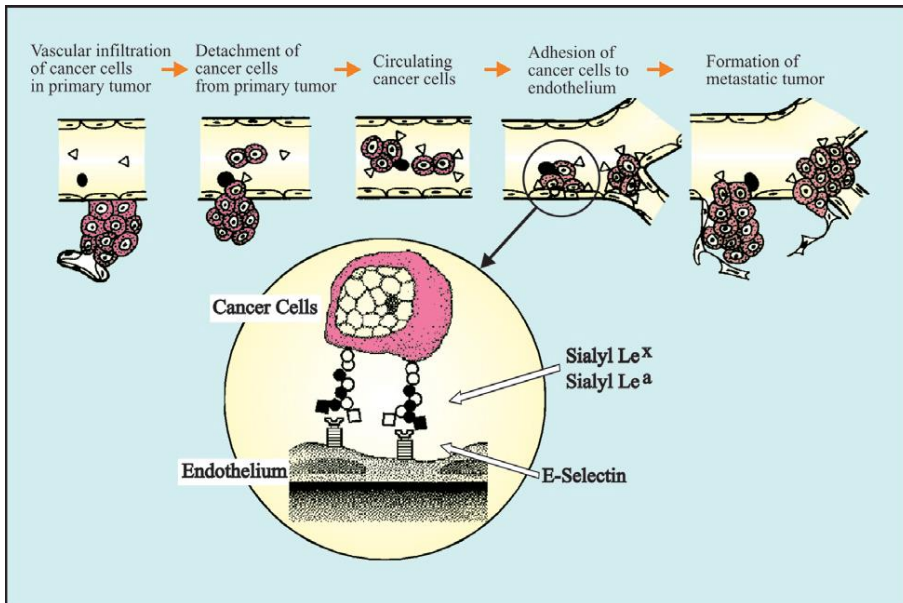
Figure 1.2. Representation of the multi-step process of hematogenous metastasis of cancer [53]. The process starts with the intravasation of cancer cells into the bloodstream in the primary tumor lesion. Cancer cells then travel through the bloodstream, where they interact with various blood cells such as leukocytes and platelets, finally adhering to endothelial cells somewhere in the peripheral vessel walls. The interaction of E-selectin on endothelial cells and sialyl Lewis a/x determinant on cancer cells is involved in cell adhesion.

The role of sLe(x) in the recognition of cancer cells by natural killer (NK) cells stemmed from the unexpected observation that melanoma cells expressing high sLe(x) levels were less metastatic than cells expressing moderate levels of the antigen [54, 55]. This striking behavior was explained by the finding that high sLe(x)-expressing cells were a better target of NK cells than cells expressing moderate levels of the antigen [55]. Altogether, these findings indicate that sialyl Lewis antigens are important

in mediating key steps of the metastatic process, in particular the adhesion of emboli of cancer cells to endothelia and neoangiogenesis. Nevertheless, very high expression levels of sLe(x) can trigger a strong NK-mediated tumor rejection [56].

The terminal steps of the biosynthesis of sialyl Lewis antigens proceeds from the α1,3/4 fucosylation of α2,3-sialylated type 1 (sLe(a)) or type 2 (sLe(x)) chains. On this basis the role of α1,3/4 fucosyltransferases and α2,3 sialyltransferases in the cancer-related over-expression of sialylated Lewis antigens has been the focus of intense investigation. It should be noted that the forced expression or down-regulation of α2,3 sialyltransferases [47, 57] or of α1,3 fucosyltransferases [43, 45] or of core 2 β1,6 N-acetylglucosaminyltransferase (C2GnT) [58] could modulate the expression of sialyl Lewis antigens in experimental systems. However, this does not necessarily imply the regulatory role of each mentioned glycosyltransferase in vivo. There are at least four enzymes which can mediate the addition of fucose in α1,3 linkage to an α2,3-sialylated type 2 chain: fucosyltransferases III, V, VI and VII (FUT3, FUT5-7), while only one (FUT3) can add fucose efficiently in α1,4-linkage to an α2,3-sialylated type 1 chain. The expression of sLe(x) appears to be regulated mainly by FUT6 in breast tumors [59], while in lung tumors it is regulated by a coordinate up-regulation of FUT3 and FUT6 [60]. In gastrointestinal tumors, such as pancreatic cancer, the over-expression of sLe(x) antigen seems to correlate with an augmented expression of α2,3-sialyltransferases ST3Gal III and ST3Gal IV [61]. In colon cancer, the molecular basis of the over-expression of sialyl Lewis antigens are particularly complex. An investigation on the level of activity of the fucosyltransferases synthesizing sLe(a) or sLe(x) concluded that an altered activity of fucosyltransferases could not explain the increased expression of sLe(a)/sLe(x) antigens in colon cancer tumors [62]. Consistently, other investigations reported that the mRNA level of

different fucosyltransferases and sialyltransferases involved in the biosynthesis of sialylated Lewis antigens could not explain their increased expression in colon cancer tissues [63, 64].

However, the biosynthesis of sialyl Lewis antigens is a complex process involving the coordinate expression of several glycosyltransferases, which might be different depending on the nature of the glycoconjugate (N- or O-linked chains of glycoproteins or glycolipids) carrying the antigen. In fact, the expression of both sLe(x) and sLe(a) antigens expressed by glycolipids in colon cancer tissues has been related to the activation of a β1,3GlcNAc transferase which synthesizes a sugar chain which is a precursor for both type 1 and 2 Lewis structures [65]. Interestingly, this enzyme is activated by Helicobacter pylori infection, leading in stomach cells to increased expression of sLe(x), which is a ligand for H. pylori sialic acid-binding adhesin SabA [66].

On the other hand, the relative abundance of type 1 and type 2 chains is an important factor in determining the relative level of expression of sLe(x)/sLe(a) antigens. An up regulation of lactosaminic chains [67] and of their biosynthetic enzymes β1,4-galactosyltransferase I [68] and -IV [69] and a down-regulation of the B3GALT5 which synthesizes type 1 chains in epithelia [1, 70] (Figure 1.1), has been reported in colon cancer [71-73] indicating a switch towards the synthesis of type 2 chains in the transformation of colonic tissues. The key role of B3GALT5 in the regulation of the balance between type-1 and -2 chains was also indicated by the finding that suppression by anti-sense DNA of B3GALT5 resulted in down-regulation of sLe(a) and up-regulation of sLe(x) and lactosaminic chains in the pancreatic cancer cell line BxPC3 [74].

The down-regulation of the biosynthesis of type 1 chains in colon cancer tissues leaves unanswered the question on the origin and the nature of the circulating CA19.9 present in the blood of several patients affected by

various cancers of digestive organs. Recent data from our lab demonstrated that no antigen is detectable in colon cancer homogenates, by dot-blot or enzyme immunoassay. 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, which appear much higher than those measured in colon cancer specimens. In cancer samples, B3GALT activity was closely related with B3GALT5 transcript down-regulation. Since colon cancers appear unable to synthesize relevant amount of CA19.9, it was suggested that the metabolic origin of the circulating antigen should be searched in other tissues actually able to synthesize and secrete relevant amounts of CA19.9 [17].

1.3 B3GALT5 transcription is driven by multiple promoters

The mRNA corresponding to the *B3GALT5* coding sequence was found expressed in epithelia of various gastrointestinal tissues and some related cell lines [1], and strongly down-regulated in colon cancer [72]. A single transcription starting site was recognized, and then a 5' flanking region was found able to act as a promoter [71]. HNF1 and Cdx transcription factors were proposed to control *B3GALT5* expression [71] according to a model reported for the typical intestinal enzyme sucrase-isomaltase. Very interestingly, it was found soon later that the entire exon 1 of this B3GALT5 transcript, as well as the putative HNF1/Cdx binding motive in the 5'flanking promoter region, belonged to a retroviral long terminal repeat (LTR) sequence [75] also present in genomes of Old World monkeys, but not in New World monkeys or mice. The sequence of the B3GALT5 LTR proximal promoter region was found highly similar in all primate species tested. Indeed, non-transcribed LTR nucleotides 1-174 were 97.1% identical between the human and baboon sequence. The sequence identity of the full-length LTR was 92.2%, comparable to the expected identity between

human and Old World monkey non-coding DNA. These findings dated the insertion of the transposon to 25-30 million years ago, and suggested the existence of another ancestral promoter [76].

Later on, Mare and Trinchera characterized multiple *B3GALT5* transcription initiation sites and cognate 5'UTRs, as well as distinct 5' flanking regions active as promoters [77] (Figure 1.3). Among them, the type A promoter, sensitive to CCAAT-binding factor (also named NF-Y), was proposed as the native promoter. This ancestral promoter conserved in mouse, where it acts as a unique B3GALT5 promoter, has substantial gastrointestinal specificity in this specie, as previously suggested [76]. This type A promoter provides low *B3GALT5* expression in both humans and mice. In humans, it is located about 42 Kbp upstream of the LTR sequence in the context of CpG islands, appears active in many epithelia, such as mammary gland, thymus and trachea, but weak and down-regulated in cancer through epigenetic mechanisms [78]. A common role of *B3GALT5* in some general mammalian cell function may be proposed when it is transcribed under the control of this promoter. Little information is currently available concerning type 1 chain expression in the mouse [79, 80], where α1,4 fucosylation does not occur [81]. Moreover, it is not presently clear how B3GALT5 predominates in the murine gastrointestinal tract, as it happens in the human counterpart, simply using the native promoter despite the absence of the LTR promoter in this species.

The type B promoter, located about 27 Kbp upstream of the LTR in humans, appeared as a defective weak promoter having the same HNF1/Cdx binding sequence as the LTR promoter, but placed in the opposite orientation. Because of such a perfect sequence match, it was supposed that the type B promoter helped stabilize the LTR promoter at the time of its transposition. A suggestive hypothesis is that the type B sequence represents a remnant of a true native promoter that was relevant

at the time of the transposon insertion but has been overwhelmed by the mobile element [77].

The type C promoter was found to be a rather strong promoter active in the small intestine only. It is placed about 17 Kbp upstream the LTR sequence and its binding sites for transcription factors remain unknown. On the other side, the brain-specific promoter appeared to be a very weak promoter whose relevance is uncertain because the expression levels of *B3GALT5* in brain are extremely low. In fact, other β1,3-galactosyltransferases, such as B3GALT1 [8, 70] and B3GALT2 [7], are expected to play a major role in tissues and cells of neuroectodermal origin. All B3GALT5 transcripts driven by such promoters share a common 3' sequence including exons 3 (untranslated) and 4 (spanning the entire coding sequence), and differ for their 5'UTRs only (Figure 1.3).
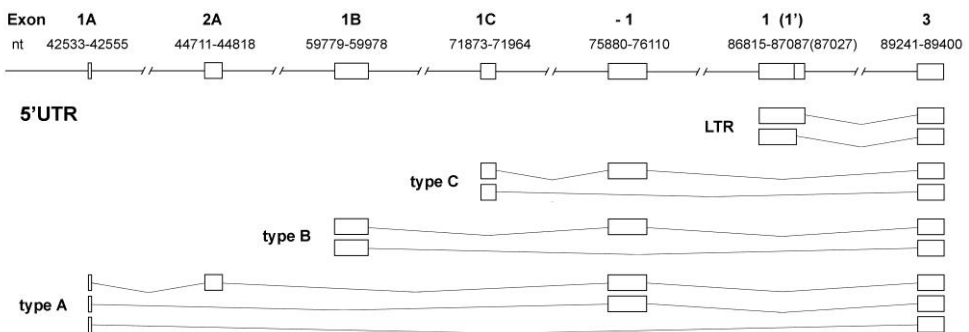
Figure 1.3. Genomic structure of the human B3GALT5 gene [77]. The positions of the exons in the context of human chromosome 21 are shown following the nucleotide (nt) numeration of the available sequence (GenBankTM accession number AF064860). The structure of the 5'-UTRs of B3GALT5 mRNA with their splice variants is presented together with the reported LTR-derived 5'-UTR.

1.4 Transposable elements

Remarkably, almost half of eukaryotic genomes comprise transposable, or transposed, elements (TE). These are repeated and mobile DNA sequences, with the capacity to replicate, move across genomes and invade them.

TEs can be separated into two major classes depending on their transposition mechanisms (Figure 1.3). Class I elements, called retrotransposons, move via a reverse-transcribed RNA intermediate and are represented by long terminal repeat (LTR)/endogenous retrovirus (ERV) elements, non-LTR retrotransposons (such as long interspersed nuclear elements (LINEs)) and non-autonomous elements (short interspersed nuclear elements (SINEs) such as Alu and the composite primate family of elements termed SVA [82, 83]. Class II elements replicate without an RNA intermediate, either by a cut-and-paste mechanism (DNA transposons containing transposases), by rolling-circle DNA replication (helitrons), or by mechanisms that remain unknown (polintons/mavericks). Class II elements also include non-autonomous TEs, such as truncated DNA transposons and miniature inverted-repeat TEs (MITEs) [84]. Each family can be further separated into subfamilies based on TE structure and evolutionary aspects (such as DNA sequence homology).

TEs are genetic units bringing positive, neutral, or negative effects to the host. They serve as recombination hot spots and may acquire specific cellular functions, such as controlling protein translation and gene transcription [85]. One of the most direct influences of transposable elements on the host genome is their role in modulating the structure and expression of "resident" genes.
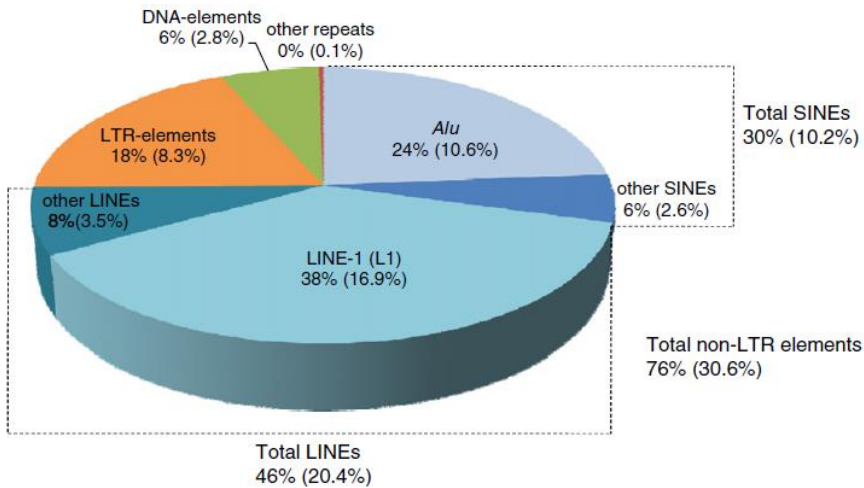
Figure 1.4. Schematic overview of the TE repartition in the human genome [83]. Transposable element as a whole represents about 45% of the human genome. The diagram represents the percentage of each TE category compared to all transposable elements (first %) and over the entire human genome (% in brackets).

Many TEs have been described in the last decade that can add a variety of functions to their targeted genes. These include alteration of splicing patterns, premature transcription termination, increased recombination, alternative promoter use [86, 87], polyadenylation sites, enhancer and silencers effects [88]. It seems that a sizable fraction of eukaryotic gene-associated regulatory elements arose in this modular fashion by insertion of TEs, and not only by point mutations of static neighboring sequences. When a TE is inserted upstream from a gene, a few short motifs can be conserved if they were subjected to selective pressure as promoters or enhancers of transcription. Even though the rest of the TE sequence might evolve beyond recognition due to absence of functional constraints, TEs are hence exapted into a novel function [89]. Retrotransposition activity itself is a source of raw genetic material for evolution, resulting in variability within and among species [90, 91].

1.4.1 LTR elements as promoters

After discovery that LTRs carry promoter and enhancer motifs, it became clear that integration of such elements in proximity of a host gene must have an influence on this gene expression [92]. Although LTRs are widely distributed in the human genome, their distributions along the human chromosome are not only non-random but also associated with gene density [93]: they are enriched in the regions of transcription units. LTRs can be located in the sense and antisense orientation of its adjacent gene and at any region, including 5'UTR, intron, exon and 3'UTR. These distributions provide favorable conditions for LTRs regulating the expression of their neighboring genes in different ways [94]. LTR elements can benefit from genes by surviving from one generation to the next, while genes can benefit from LTR characteristics. For example, genes could obtain their promoter (sometimes bidirectional) [75, 76, 95, 96], enhancer [95], polyadenylation sites [97, 98], alternative splicing pattern [99] from LTRs.

The LTR promoters active in the modern human genome are likely to have beneficial or neutral effects because if an LTR insertion results in a detrimental change in gene expression it will be selected against and will not become fixed in the population [100]. A significant role for ERVs in shaping gene regulatory networks is starting to be revealed through whole genome analyses that show functional transcription factor binding sites (TFBSs) are commonly present within LTR sequences. Binding sites for the transcription factors p53, CTCF, Pou5F1-Sox2 and ESR1 are enriched within individual LTR sequences and are also present in LTR consensus sequences [101, 102].

Major LTR promoter-induced expression changes are actually the exception to the rule. In most cases, the LTR promoter exerts a relatively subtle effect on the tissue specificity of expression and is likely to only augment expression from the native promoter [100]. Deviations from this

trend are found in placenta, a tissue in which ERVs are known to be more transcriptionally active [103, 104] and where LTR promoters do indeed tend to confer tissue-specific expression.

1.4.1.1 Alternative promoters

Individual cases of alternative gene promoters have long been described and more recent genome-wide studies have estimated that up to 75% of human genes use alternative promoters, perhaps far more than anticipated [105-107]. The varying genetic structure of alternative promoters results in different functional effects [108, 109]. In most cases, the two or more alternative promoters each contribute an alternative first exon containing distinct transcription start sites. The translation initiation codon is within a common downstream exon so the same protein is formed regardless of promoter use [110-112]. The main functional consequence of such an alternative promoter is to drive different transcription patterns, for example in different tissues, at different developmental stages or at different levels of expression [113, 114]. However, if two or more alternative first exons each contain unique translation initiation sites, they will encode proteins with different N-terminal but identical C-terminal sequences [115, 116]. Alternative promoter use can also regulate alternative splicing products, resulting in the production of different protein isoforms [117, 118]. These changes in protein sequence may result in loss or gain of alternative functional domains or may change the subcellular localization due to the inclusion or exclusion of signal peptides. Alternative promoters can even drive translation of different open reading frames producing completely different proteins [119, 120] and can affect the translation efficiency of different isoforms. The mechanisms that control alternative promoter use in different cellular states are not very well studied but probably involve either availability of regulatory proteins, such as transcription factors, or epigenetic mechanisms such as DNA methylation [108].

1.4.1.2 LTR exaptation as primary promoter

In most cases, the LTR is one of two or more alternative promoters, but there are examples where the LTR is the sole or primary promoter.

The LTR12C (HERV-9) promoter of *GBP5* (guanylate binding protein 5) confers an expression pattern in lymphocytes, endothelial cells and a lymphoma cell line that is similar to that of its paralog, *GBP1* [121, 122]. There is evidence that *BAAT* (bile acid CoA: amino acid N-acyltransferase) and *MSLN* (mesothelin) have non-TE alternative promoters as well as their characterized LTR promoters [123]. For *BAAT*, the alternative promoter also appears to be liver-specific. The contribution of the *MSLN* native promoter is unknown, but the LTR promoter has been characterized and acts as a core promoter with additional non-LTR enhancers required for specific, functional expression in mesothelium [124, 125]. *ADH1C* (alcohol dehydrogenase 1C) involved in the alcohol metabolism, and *HSD17B1* (hydroxysteroid 17-beta dehydrogenase 1) involved in estrogen synthesis in breast cancer are other examples of LTR exapted as primary promoters [100].

1.4.1.3 Placental specificity of LTRs

In many species ERVs are highly transcribed in germ line cells, testis and placenta [103, 104, 126] so it is not surprising to find that LTR promoters are also more active in these tissues. *CYP19A1* (cytochrome P450, family 19, subfamily A), *IL2RB* (interleukin 2 receptor, beta), *NOS3* (nitric oxide synthase 3) and *PTN* (pleiotrophin) are all examples of genes whose expression in the placenta is solely due to the presence of the LTR promoter [100]. The placental expression of genes with LTR promoters may be an artifact of the evolutionary mechanism of endogenization since retroviruses that were more active in reproductive tissues may have been more likely to be able to infect germ cells. Assuming this hypothesis, we might predict that active LTRs will preferentially contain binding sites for

placental or reproductive tissue-specific transcription factors. An alternative theory, not mutually exclusive, is that ERV LTRs are less methylated in germ line cells and reproductive tissue [127], although in a recent study this could not be concluded to be the sole mechanism of ERV LTR transcriptional regulation in the placenta [128]. Further studies are required to understand fully the regulatory mechanisms by which LTR promoters are more active in the placenta (and potentially other reproductive tissues) than other somatic cells.

A further point to consider is whether the LTR-driven gene expression in the placenta has any functional consequence, regardless of the mechanism by which expression is achieved. Expressed ERV genes are known to have been exapted in the placenta, for example the HERV-W Env gene Syncytin has acquired a functional role in placental syncytiotrophoblast formation [129, 130]. In several cases (*PTN*, *CYP19A1*, and *PAPPA2*) orthologs in other mammals are also expressed in the placenta, but not due to retroviral insertion [131]. In these cases the gene function, rather than the regulatory mechanism by which it is expressed, has been conserved, suggesting that their placental expression is actually a result of convergent evolution rather than a novel function of the LTR insertion. For *EDNRB* (endothelin receptor type B), *ENTPD1* (ectonucleoside triphosphate diphosphohydrolase) and *MID1* (midline 1) the LTR promoter augments placental expression rather than providing novel specificity, and these genes are known to have functional effects in the placenta. Similarly, the LTR promoter of *INSL4* (insulin-like 4) has been exapted as the primary promoter and confers the expression that enables *INSL4* to have a role in placental morphogenesis [131, 132]. However, there are a number of cases (*IL2RB*, *NOS3* and *HSD17B1*) for which the functional effects of placental expression remain unknown.

1.4.2 Evolutionary models for LTR insertion

There are some possible evolutionary scenarios in which an LTR insertion upstream of a cellular gene results in the LTR developing into a promoter (Figure 1.5). Models (a) and (b) show LTR insertions that evolve as alternative promoters under two basic evolutionary scenarios. Upon integration LTRs will naturally contain transcriptional regulatory motifs, such as TFBSs, necessary for the retroviral life cycle. The first case (a) describes a model in which TFBSs present in the original LTR insertion that maintains functional roles are conserved more than non-functional TFBSs. Dunn *et al.* have proposed this model for the case of *B3GALT5* on the basis of sequence comparison between the specific LTR copy, the LTR consensus sequence and insertions in other primates [76]. This model is also supported by genome-wide studies analyzing the distribution of TFBSs within LTRs that show the presence of regulatory motifs within consensus LTR sequences [101, 102]. In the second scenario (b), TFBSs have been acquired through novel mutations in the LTR and consequently are only present in individual LTR sequences, not in the consensus sequence for that family. Scenarios (c) and (d) depict evolutionary outcomes that relate to other examples described above. In model (c), transcription from the LTR promoter gives the same expression pattern as the main promoter leading to two possible outcomes; (i) the LTR is exapted as the primary promoter; (ii) the LTR promoter activity is lower than the native promoter so its effect is minor. Sometimes, important functional binding sites may lie upstream of the LTR in which case the LTR contains core promoter elements and tissue specific or additional activating factors are contributed by the host genome (Model d). Although not specifically depicted, this scenario could again result in the LTR becoming the primary promoter of the gene, or one of two alternative promoters.

(a) LTR-dependent change in tissue specificity, TFBSs present in original insertion e.g. *B3GALT5*

(b) LTR-dependent change in tissue specificity, TFBSs acquired through mutation

(c) LTR exaptation resulting in (i) LTR as primary promoter e.g. *BAAT* (ii) minor effect e.g. *SEMA4D*

(d) LTR contributes core promoter activity, upstream TFBSs provide tissue-specify e.g. *APOCI*
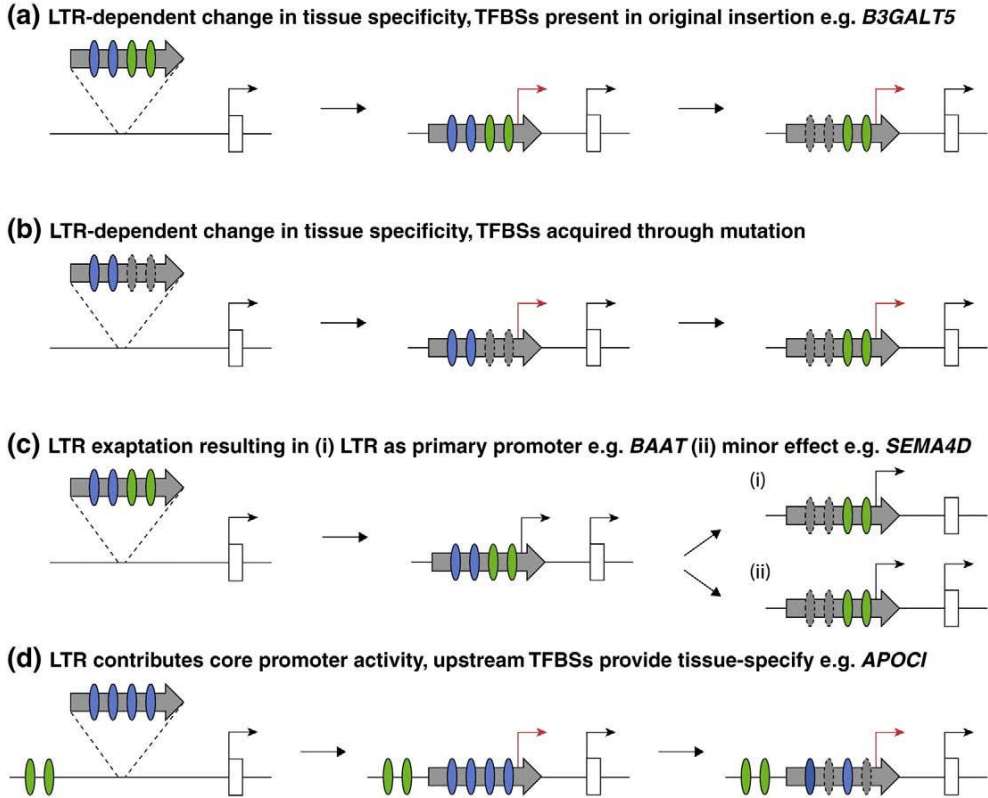
Figure 1.5 Evolutionary models of LTR promoter generation and maintenance modified from [100] (a) Insertion of an LTR (grey arrow) with pre-existing tissue-specific or general TFBSs (green and blue ovals) upstream of a native promoter (black bent arrow) results in novel tissue-specific expression from the LTR promoter (red bent arrow). Over evolutionary time (moving from left to right), only functional TFBSs are maintained. (b) A similar scenario to (a), except functional TFBSs are not present in the original LTR insertion but evolve over time. (c) An LTR insertion contributes an alternative promoter with the same specificity as the native promoter, resulting in (i) loss of the native promoter or (ii) a minor effect. (d) An LTR insertion contributes core promoter activity and upstream TFBSs contribute tissue-specific effects.

1.5 Genes expression regulated through DNA methylation

Until recently, the majority of DNA methylation studies focused on the analysis of CpG islands associated to promoter regions. Experimental evidence demonstrated that gene expression can be regulated by DNA methylation levels of CpG islands in the proximity of transcription start sites [133]. Based on these data, the majority of cancer-related DNA methylation studies focused on the role of CpG islands hypermethylation as a mechanism of tumor suppressor gene silencing [134, 135]. For explaining this phenomenon, some models are been proposed. One of the most accepted models is that gene silencing is mediated by proteins containing a methyl-CpG binding domain (MBD). MBD binding to methylated CpG islands is followed by recruitment of histone deacetylases, chromatin compaction and gene silencing [136, 137] (Figure 1.6a). Another possible model involves DNA methylation dependent transcription factor binding to their recognition sites; if the DNA is methylated, the transcription factor cannot bind to the promoter region and therefore, the gene is inactive [138, 139].

However, already in the 1980s, pioneer studies in cancer cells reported that neoplastic transformation was also associated with global and gene specific loss of DNA methylation [140, 141]. Hypomethylation of transcription regulatory regions in cancer seems to be much less frequent than hypermethylation of CpG islands overlapping promoters [141-143]. Nonetheless, some of the cancer-associated loss of DNA methylation encompasses gene regions, including transcription control sequences [144-146].

With the upcoming possibilities to study DNA methylation in a genome-wide context, this epigenetic mark can now be studied in an unbiased manner. As a result, recent studies have shown that DNA methylation does not occur exclusively at CpG islands. The term CpG island shores, referring to regions of lower CpG density that lie in close proximity (~2 kb) of CpG islands, has recently been described. The methylation of these CpG island shores is closely associated with transcriptional inactivation (Figure 1.6b). Most of the tissue-specific DNA methylation seems to occur not at CpG islands but at CpG island shores [147, 148]. Differentially methylated CpG island shores are sufficient to distinguish between specific tissues and are conserved between human and mouse.
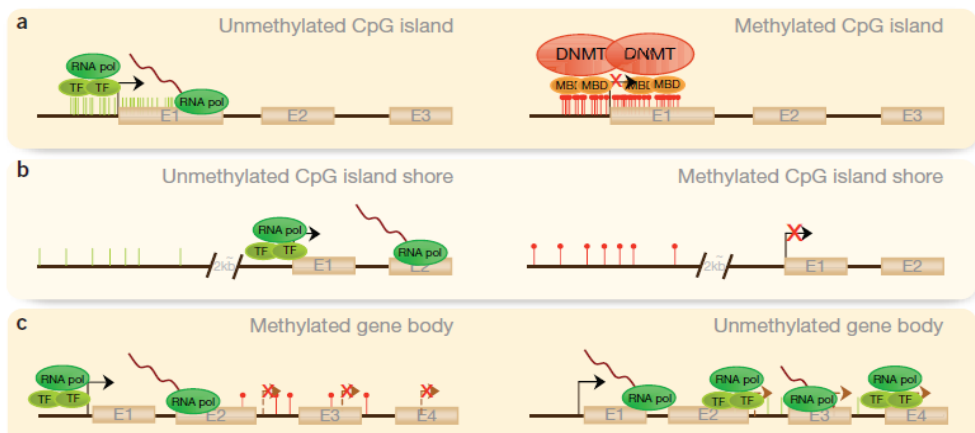
Figure 1.6 Effects of DNA methylation in different regions of the genome modified from [149].(a) CpG islands at promoters of genes are normally unmethylated, allowing transcription. Aberrant hypermethylation leads to transcriptional inactivation. (b) The same pattern is observed when studying island shores, which are located up to 2 kb upstream of the CpG island. (c) However, when methylation occurs at the gene body, it facilitates transcription, preventing spurious transcription initiations. In disease, the gene body tends to demethylate, allowing transcription to be initiated at several incorrect sites.

Although DNA methylation mainly occurs in the CpG dinucleotide context in mammals, non-CG methylation has recently been described in humans at CHG and CHH sites (where H is A, C or T) [149]. CHG and CHH methylation has been found in stem cells and seems to be enriched in gene bodies directly correlated with gene expression and to be depleted in protein binding sites and enhancers [150]. The levels of non-CpG methylation decrease during differentiation and are restored in induced pluripotent stem cells, suggesting a key role in origin and maintenance of pluripotent state [150, 151]. Mechanisms of non-CpG methylation remain unclear [151].

In addition to 5-methylcytosines, 5-hydroxymethyl-2′-deoxycytidine has also been observed. So far, 5-hydroxymethyl-2′-deoxycytidine has been reported in Purkinje cells constituting less than 1% of total nucleotides, but it seems not to be present in cancer cell lines [149]. These new DNA modifications need to be further studied to determine their implications for normal and diseased epigenetic regulation.

Additionally, not only promoters but also intragenic and intergenic regions are widely modulated during physiological processes and disease. In particular, it is becoming clear that DNA methylation in the gene body widely change during cell differentiation and carcinogenesis. It seems to be actively involved in multiple gene regulation processes, such as transcript elongation, expression of intragenic coding and non-coding transcripts, alternative splicing and enhancer activation.

1.5.1 Intragenic DNA methylation and gene expression

The negative correlation between gene expression and CpG islands methylation at the transcription start sites is well established. However, this association cannot be extrapolated to other genomic contexts such as CpGs located in the gene body. Gene body methylation is common in ubiquitously expressed genes and is positively correlated with gene

expression [152]. It has been proposed that it might be related to elongation efficiency and prevention of spurious initiations of transcription (Figure 1.6c). 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, seem to recruit DNA methyltransferases, thus facilitating the methylation of intragenic DNA [153]. Therefore, even if the gene body is highly methylated, which is a frequent finding in normal undifferentiated cells, the gene may be transcribed. This can lead to an apparent contradiction, as DNA methylation in the promoter can be negatively associated with gene expression whereas the methylation status in the gene body of the same gene can show a positive correlation. A positive correlation between intragenic DNA methylation and gene expression has been recently observed in multiple genome-wide epigenomic studies, both in the context of cell development and differentiation as well as in cancer cells [154, 155].

An interesting study observed a context-dependent correlation of CpG gene-body methylation, related to whether the CpGs were located in 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 either be positively or negatively correlated with gene expression levels. Only for approximately 15% of the intragenic CpG islands presenting a negative correlation, an association with specific gene regulatory processes, e.g., alternative promoter usage or intragenic enhancer activity, was found [156]. DNA methylation at intragenic CpG islands seems to regulate alternative promoter usage and to interfere with expression of the main transcripts [157]. It seems that DNA demethylation at alternative promoters results in an accessible chromatin status and binding of transcription factors as well as transcription initiation factors.
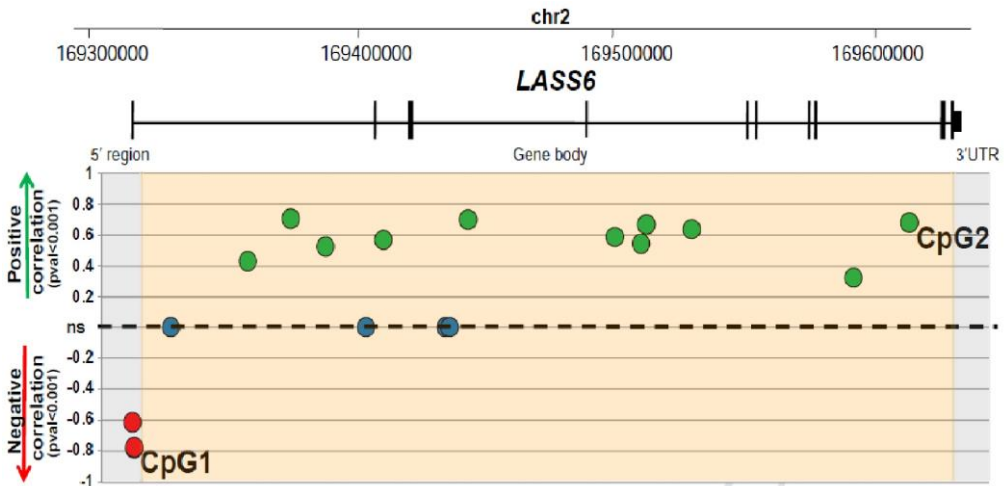
Figure 1.7 Association between intragenic DNA methylation and gene expression [157]. Correlation between DNA methylation and expression across the promoter and gene body of the LASS6 gene. Vertical bars in the schematic representation of LASS6 represent exons. CpGs located at the 5' region (gray shadow) show a clear negative correlation whereas most of the CpGs in the gene body (pink shadow) show a positive (or not significant (ns)) correlation between DNA methylation and gene expression levels. The data for this figure are derived from DNA methylation and gene expression arrays performed in 125 patients with chronic lymphocytic leukemia [158]

Another role of gene-body DNA methylation is the regulation of intragenic non-coding RNAs (ncRNAs) present in the intronic region of numerous genes. They can be co-expressed with the host gene or can have their own promoter. Expression of many miRNAs [159, 160] and lncRNAs [161] has been described to be regulated by DNA methylation in a similar way as for protein-coding genes. Hence, alteration of DNA methylation patterns can be responsible for the deregulation of miRNAs or lncRNAs expression and consequently can alter the expression of their target genes. In addition to short and long ncRNAs, the expression of transposable elements has also been shown to be regulated by DNA methylation. It is widely accepted that

DNA methylation is crucial in preventing the activation of LINEs, SINEs, retroviruses and other transposable elements, thereby ensuring genomic integrity and stability [141, 142]. An additional function of intragenic DNA methylation is related to RNA processing mechanisms such as alternative splicing and alternative polyadenylation [162]. It is commonly accepted that alternative splicing is regulated by splicing enhancers, silencers and specific binding of splicing factors [163]. Furthermore, several studies have identified a link between epigenetics and alternative splicing [164, 165].

1.5.2 DNA methylation at enhancers

Different genome-wide and single gene studies demonstrate that a large number of enhancers show a reverse correlation between DNA methylation and gene expression levels or enhancer activity [166, 167]. Aran *et al.* have found that a subset of the predicted enhancers was hypo- or hypermethylated in the cancer cells compared to the normal cells, thus enhancer methylation is remarkably altered in cancer [167]. Moreover, they showed that enhancer methylation is more closely related to changes in gene expression than promoter methylation. This indicates that in some cases, surprisingly, the methylation status of distal regulatory regions might be more predictive for expression levels than promoter methylation itself [167]. Similar results were obtained in a large study, in which the strongest correlation between gene expression and DNA methylation was found in intragenic CpGs rather than for those located at 5′ regions. Furthermore, the negative correlation between intragenic DNA methylation and gene expression was frequently related to enhancer regions [158]. A rather new phenomenon that sheds light on enhancer activity is expression of non-coding RNAs from enhancer sequences (eRNAs) [168]. Expression of these eRNAs is correlated with transcription levels of nearby located genes. Interestingly, it was recently shown by Melo *et al.* that induction of eRNA expression by p53 may occur at loci with a pre-programmed chromosome

conformation state. At these loci, p53 responsive enhancer elements are already located in close proximity to their target genes, while p53 binding and consequent eRNA expression are the final, essential steps that facilitate expression of p53 target genes [169].

Many factors are involved in transcription regulation at the enhancer level, e.g., protein binding, formation and stabilization of long range promoter-enhancer interactions and expression of eRNAs. Hence, the effect of DNA methylation changes on transcription potential may vary in a similar context-dependent manner.

2. RATIONALE

HNF1 is important for liver specific expression of a variety of genes including albumin, α1-antitrypsin, and β-fibrinogen [170, 171]. Although originally described as liver-specific [172], HNF1 has subsequently been shown to be expressed not only in liver but also in kidney, intestine, stomach and pancreas [173-175]. Studies on HNF1- knocked out mice have revealed that this transcription factor is crucial for the transcriptional activation of genes that play key roles in liver (in phenylalanine catabolism), pancreas (in β-cell glucose-sensing), and in kidney (in the renal proximal tubular reabsorption of glucose) [176]. Additionally, mutations in the human HNF1 gene, in the heterozygous state, have been found to be responsible for particular forms of diabetes mellitus termed maturity-onset diabetes mellitus of the young (MODY-3 and MODY-5) [177].

Sucrase-isomaltase is another example of a non-hepatic gene regulated by HNF1 [178] that is expressed in a strict tissue, position and cell lineage-dependent pattern. Absorptive enterocytes located on villi in the small intestine are the only cells in the body that express mRNA [179] or protein [180]. Not only HNF1, but also Cdx1 and Cdx2 [181] and GATA-4 [182] have been described as primary transcription factors that can cooperate in the regulation of this gene. Sucrase-isomaltase expression pattern has been extensively investigated [178-182] and extrapolated to explain a general model of regulation of intestinal genes. However, re-evaluation of the tissue distribution of the various B3GALT5 mRNAs [77] indicated that the LTR transcript was very highly expressed only in normal colon mucosa, while in the other organs of the gastrointestinal tract, including the small intestine, it is present in lower amounts. Such expression pattern was thus opposite to that reported for sucrase-isomaltase, which instead is active in the small intestine [178-180].

Moreover, according to such expression model depending on HNF1 and Cdx, we should expect: 1) Not to find HNF1 expressed on colon cancers

biopsies which express low levels of B3GALT5 LTR transcript. However, in previous experiments we determined HNF1 in such samples and we found appreciable levels of HNF1. 2) If the LTR expression depends on Cdx, then this factor should be expressed on cell lines expressing high levels of the LTR transcript. Nevertheless, it was reported that COLO-205 cells, expressing extremely high levels of B3GALT5 mRNA, including the LTR transcript [1, 74], were devoid of any Cdx expression [183]. These findings appeared not compatible with the proposed model of *B3GALT5* transcription and prompted us to re-evaluate it in details.

Strategy

To determine the role of the putative transcription factors HNF1α/β and Cdx1/2 in *B3GALT5* transcription, we compared the expression levels of the B3GALT5 LTR transcript, quantitated by competitive RT-PCR, with those of transcription factors, determined by western blot analysis, in colon cancer biopsies, in various cell lines, and in cell models serving as controls. They included cells transfected with plasmids coding HNF1α/β cDNAs or shRNAs in order to corroborate if the overexpression or silencing of those factors can modulate LTR transcript expression. Although the LTR promoter and the proximal sequences do not have CpG islands, we treated cells expressing the LTR transcript with the demethylating agent 5AZA to understand the effect of this epigenetic feature on the activation of B3GALT5 LTR promoter. Finally, we evaluated the behavior of the LTR promoter *in vitro*, through electrophoresis mobility shift and reporter luciferase assays.

3. AIMS OF THE STUDY

This study is aimed to comprehend the regulation of B3GALT5 LTR promoter in order to:

 I. Elucidate the role of candidate transcription factors and the presence of potential epigenetic marks, to explain tissue specific expression and regulation, mainly in the process of colon tumorigenesis.

II. Understand the evolutionary stabilization of the transposon occurred in humans and some primates.

4. MATERIALS AND METHODS

4.1 Materials

4.1.1 Cell lines

Human breast cancer cell lines MCF-7 and MDA-MB-231, human gastric cell line MKN-45 and KATOIII, human bile duct carcinoma cell line HuCC-T1, and human colon cancer cell lines HT-29, HCT-15, COLO-205 and SW-1116 were cultured as reported [74, 77, 78, 184]. Human breast cancer cell line MDA-MB-361, a gift of Dr. Cristina Razzari (University of Milan, Italy), human embryonic kidney HEK-293T/17 (ATCC CRL-11268), a gift of Dr. Anna Menini (SISSA, Trieste, Italy) and human hepatoma cell lines Huh-7 and Hep-3B, a gift of GIM microscopy group (University of Milan, Italy), were cultured in DMEM containing 10% fetal bovine serum, 100 U/ml penicillin, 1 mg/ml streptomycin, and 2 mM l-glutamine.

4.1.2 Human biopsies

Human colon biopsies were as reported [72, 184]. This study was approved by the Senior Committee Board regulating non interventional studies, comparable to an Institutional Pre-review Board. All patients underwent surgery for colorectal cancer at Bologna University Hospital. After careful examination by the Pathologist, a sample of normal and cancer tissue was dissected for biochemical studies. Each specimen was immediately snap frozen in liquid nitrogen and conserved at -80 °C.

4.1.3 Plasmid DNAs

The coding sequences of HNF1α, HNF1β, Cdx1 and Cdx2 were obtained by PCR using the below mentioned primer pairs, harboring HindIII and XbaI restriction sites, starting from RNA extracted from COLO-205 (HNF1α), MKN-45 cells (HNF1β), and from commercially available (Stratagene #540009, Agilent Technologies Italia) human normal mucosa RNA (Cdx1 and Cdx2), using the reported procedure [184]. Amplified fragments were

restriction digested, cloned into the corresponding sites of the pcDNA3 vector, and sequenced. Since we obtained from our source the variants P130R [185] and S291P [186] of Cdx1 and Cdx2, respectively, we mutated the obtained plasmids with the QuikChange Lightning mutagenesis kit (Agilent Technologies, Italia) to obtain the wild-type sequences, according to the manufacture's protocol.

Primers sequences used for cloning were:

HNF1α Forward	GCAAGCTTGCCATGGTTTCTAAACTGAGCC
HNF1α Reverse	GCTCTAGATGGTTACTGGGAGGAAGAGGCC
HNF1β Forward	GCAAGCTTGAAAATGGTGTCCAAGCTCACG
HNF1β Reverse	GCTCTAGAGGCATCACCAGGCTTGTAGAGG
Cdx1 Forward	CTAAGCTTACCATGTATGTGGGCTATGTGC
Cdx1 Reverse	CTTCTAGAGCTATGGCAGAAACTCCTCTTT
Cdx2 Forward	CTAAGCTTACCATGTACGTGAGCTACCTCC
Cdx2 Reverse	CTTCTAGATCACTGGGTGACGGTGGGGTTT

Primers sequences used for mutagenesis of Cdx1 and Cdx2 were:

Cdx1 Forward	CGGAGCGCAGAGGCCGACGCCCTACGAG
Cdx1 Reverse	CTCGTAGGGCGTCGGCCTCTGCGCTCCG
Cdx2 Forward	CTGCAAGCCTCAGTGTCTGGCTCTGTCCCTGG
Cdx2 Reverse	CCAGGGACAGAGCCAGACACTGAGGCTTGCAG

For gene silencing, we used Human MISSION® shRNA Plasmid DNA (Sigma Aldrich Italia), based on pLKO vector, clone TRCN0000017509NM_000458.1-800s1c1TRC 1 (targeting HNF1β), clone TRCN0000017194NM_000545.3-1039s1c1TRC 1 (targeting HNF1α), or pLKO vector alone.
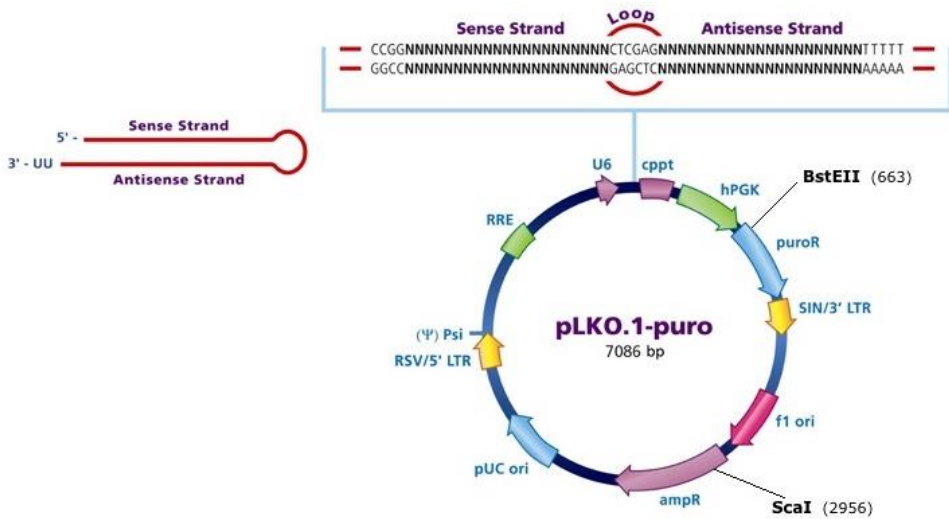
Figure 4.1 Vector Map of pLKO.1-puro (Human Mission shRNA Plasmid DNA)

4.2 Cell treatments

4.2.1 Cell treatments with drugs

Treatments of cells with drugs affecting DNA methylation and histone deacethylation were performed as reported (11). Cells were plated in 6-well plates, incubated 24h with regular medium that was replaced was replaced with medium containing different amount of 5AZA (Sigma, dissolved in DMSO as 10 mM stock solution) and/or TSA (Sigma, dissolved in ethanol as 1 mg/ml stock solution). Media were replaced every 24 h with media containing freshly diluted drugs. At the end of treatment cells were harvested by trypsinization and processed for extraction of total RNA or nuclear protein.

4.2.2 Permanent transfections

For permanent transfection, $3x10^6$ MDA-MB-231 cells were seeded in 100 mm plates 24 h in advance, and 20 µg of ScaI linearized pcDNA3-HNF1α, pcDNA3-HNF1β, or pcDNA3 alone were mixed with 1 µg of XhoI linearized pLKO-puro empty vector (Sigma Aldrich Italia) bearing the puromycin resistance gene. Transfection solutions were prepared by diluting the linearized DNA with 2 ml of serum-free medium and then adding 2 ml of serum-free medium containing 60 µl Lipofectamine 2000 (Life Technologies Italia). Liposomes were allowed to form for 20 min at room temperature. Cells were washed twice with serum-free medium, fed with transfection solution, and incubated under regular growth conditions for 3-4 h. At the end, 8 ml of standard complete medium was added and the incubation continued. Two days after transfection the cells were trypsinized and divided into multiple plates. Selection started after an additional 24 h in the presence of 1 µg/ml puromycin. Resistant colonies were picked up, grown, and harvested to extract nuclear protein and total RNA from the same cell suspensions.

4.2.3 Transient transfections

For transient transfection, $4x10^6$ HEK-293T cells were seeded in 60 mm plates 24 h in advance, and 5 µg of pcDNA3-Cdx1 or pcDNA3-Cdx2 were transfected in 1.5 ml of transfection solutions prepared as above for permanent transfection. The day after they were trypsinized and placed in 100 mm plates in the presence of 2 mg/ml G418. After additional 5 days (changing media twice), resistant cells were harvested and used for nuclear protein extraction.

4.2.4 Gene silencing

For gene silencing, plasmids were linearized with BstEII (to destroy puromycin resistance gene) and mixed with 1/20 of the same plasmid

linearized with ScaI. Transfection of MKN-45 or MDA-MB-361 cells and puromycin selection were carried out as above described for MDA-MB-231 cells. Individual or pooled colonies obtained upon selection were allow to grow 3 weeks, and the obtained cell pellets (about 0.5-1.0x10^6 cells) were washed with sterile PBS and divided in three fractions, one (0.1-0.2x10^6 cells) lysed for RNA extraction, one (0.4-0.6x10^6 cells) mixed with protease inhibitor cocktail HALT (Thermo Scientific) for direct western blot analysis, and the remaining plated again for further growing.

4.3 Measurements

4.3.1 Protein extraction and Western blot

Freshly collected cell pellets were processed to obtain nuclear extracts using a commercially available kit (NE-PER, Thermo scientific) as reported [78]. Frozen biopsies from colon were dounce homogenized and submitted to nuclear extraction as for cell pellets. Aliquots of nuclear extracts (5-10 µg of protein) were separated by 10% SDS-PAGE, transferred to a nitrocellulose membrane (Trans-Blot SD Semi Dry Transfer Cell, Bio-Rad Laboratories S.r.l, Italy) and blotted with rabbit polyclonal anti HNF1α/β (sc-8986, 1:200, Santa Cruz Biotechnology Italia), rabbit polyclonal anti H3 (D2B12 #4620, 1:500, Cell Signaling Technology, Italian distributor), rabbit polyclonal anti Cdx1 (PAB4713, 1:200, Abnova, Italian distributor), or mouse monoclonal anti Cdx2 antibody (M01, 1:200, Abnova) as the published protocol [187].

4.3.2 RNA extraction and Competitive RT-PCR

Quantification of transcripts was performed by competitive RT-PCR, since it was find very effective for distinguishing different 5'UTRs without underestimating splice variants [77]. Total RNA, prepared and DNase-treated using a commercially available kit (SV total RNA isolation system, or ReliaPrep cell RNA miniprep, both from Promega Italia) was quantitated

by fluorometry with Qubit (Life Technologies Italia), and reverse transcribed as reported [78]. cDNAs were amplified in a volume of 25 µl in the presence of the indicated amounts of competitors, for 35 cycles (B3GALT5 LTR, HNF1α and HNF1β) or 5 pg competitor for 25 cycles (β-actin). Amplification program included a single treatment at 94°C for 3 min followed by cycles consisting of 1 min at 94 °C (melting) and 3.5 min at 72 °C (annealing plus extension) and a final extension step at 72 °C for 8 min (B3GALT5 and β-actin), or 1 min at 94 °C (melting), 1 min at 66 °C (annealing) and 2.5 min at 72 °C (extension) and a final extension step at 72 °C for 8 min (HNF1α and HNF1β). B3GALT5 LTR and β-actin competitors and primers were exactly those already reported [72, 77]. HNF1β competitor was prepared subcloning the coding sequence in pCDM8 vector, digesting with restriction enzymes BspHI and MscI, and re-ligating. HNF1α competitor was prepared digesting pcDNA3-HNF1α with restriction enzymes XhoI and Bsp1407I, blunting with Klenow, and re-ligating.

Primers sequences used were:

HNF1α sense	5'-GCCATGGTTTCTAAACTGAGCCAG-3'
HNF1α antisense	5'-GTCCATAGCGCACACCGTGGAC-3'
HNF1β sense	5'-CAGAGCCATGGGCCTGGGCAGTC-3'
HNF1β antisense	5'-CTGCTGGGCCATGTGGCTGCCTG-3

Parallel PCR amplifications were performed on known amounts of standard cDNAs premixed with the competitors. Standard cDNAs were the original cloned sequences quantitated and diluted as for the competitors. Aliquots of PCR reactions were analyzed on 1% agarose gels stained with ethidium bromide. Quantification was performed by densitometric scanning of the gel and the amounts of amplified target cDNAs were calculated from their respective standard curves and normalized by those for β-actin. The

target/competitor ratios were proved to make PCR results quantitative [188].

4.3.3 Luciferase assay

The DNA fragment containing the sequence from -148 to -128 from the LTR transcription initiation site, encompassing part of the LTR transposon was amplified by PCR using LA *Taq* and GC II buffer (Takara) following the manufacturer's recommendations. The 25µl reaction contained 150 ng of human placenta genomic DNA as the template and primers having restriction sites at their 5' -ends. The resultant fragment was cloned in the corresponding sites of the vector pGL3 (Promega) upstream of the firefly luciferase coding sequence [77] and was used as the LTR promoter (pGL3-LTR).

For transfection, 50,000-80,000 host cells were plated 20 h in advance in 96-well plates with 0.1 ml of culture medium. Transfection solutions were prepared by mixing 200 ng of test DNA (pGL3-LTR alone or 100 ng pGL3-LTR plus 100 ng of pcDNA3-based plasmids) with 10 ng of the *Renilla* luciferase expression vector pRL-CMV (Promega Italia) in 25 µl of serum-free medium. Twenty-five µl of serum-free medium containing 0.6 µl of Lipofectamine 2000 were then added to the DNA dilution, and the complete solution used as above reported for all other transfections. After the addition of 0.1 ml of standard medium, the incubation was continued for 20 h. The cells were then washed, lysed, and assayed for luciferase activity as reported [77].

4.3.4 EMSA assays

EMSAs were performed with the Lightshift chemiluminescence kit (Thermo Scientific) following the manufacturer's recommendations, but the binding reaction volume was scaled down to 10 µl. DNA probes biotinylated at the 3′-end were synthesized by Eurofinsdna-MWG Operon (Ebersberg,

Gemany) and diluted to be 1 nM in the assay. Nuclear extracts were prepared as above described for western blotting and 1 μl of extract (3–4 μg of protein) was added to each binding reaction. Anti-HNF-1 antibody was incubated 30 min on ice in the binding reaction before adding the probes. Competitors oligonucleotides were: the unlabeled oligonucleotide probe (self), Cdx1/2 competitor, the 24-bp deduced from the sucrase-isomaltase promoter (SIF1) 5'-GGGTGCAATAAAACTTTATGAGTA-3' from Suh *et al.* [181] and the consensus sequence of AP2α 5'-GATCGAACTGACCGCCCGCGGCCCGT-3' used as an irrelevant DNA control. They were present in 100-fold molar excess in the incubation mixture, unless differently indicated. DNA/protein complexes were separated by 5% native PAGE, transferred to a nylon membrane, cross-linked under UV light, and visualized with chemiluminescence as reported [77].

5. RESULTS

5.1 Detection of HNF1α/β, Cdx1/2 and B3GALT5 LTR transcript in colon cancer biopsies and surrounding normal mucosa

Firstly, we determined the expression of the B3GALT5 LTR transcript by competitive RT-PCR, in matched pairs of biopsies representing colon cancers and surrounding normal mucosae (Figure 5.1).
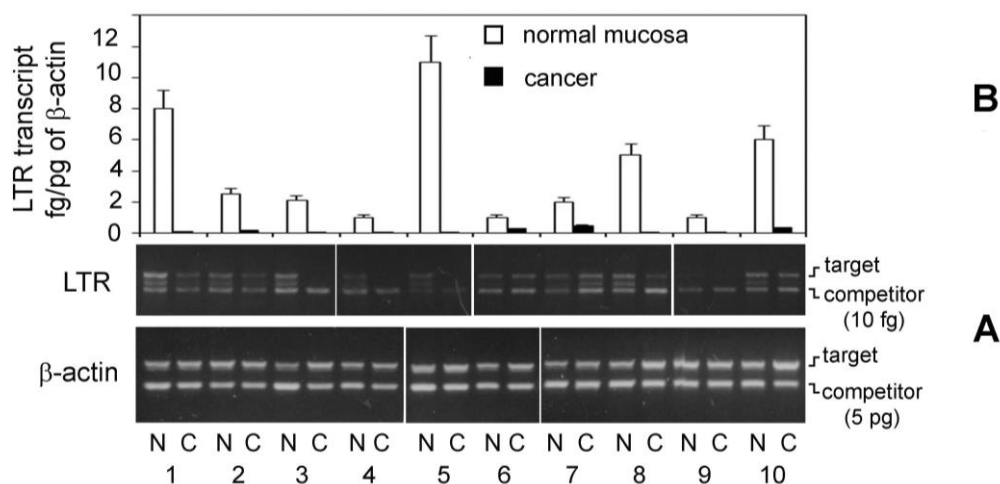
Figure 5.1 Quantification of B3GALT5 LTR transcript expressed in matched pairs of colon cancer and adjacent normal mucosa by competitive RT-PCR. A) RNA extracted from the biopsies was reverse transcribed, and normalized amounts of the resultant first-strand cDNA were mixed with the indicated amounts of competitor (truncated) cDNAs and subjected to PCR using primers specific to LTR transcript and β-actin, respectively.The target doublet corresponds to the alternative splicing previously reported [77]. Samples were not all run on the same gel, as indicated by vertical white spaces. B) Quantification of B3GALT5 LTR transcript: Densitometric scanning of bands obtained by gel images was performed to quantitate B3GALT5 LTR transcript. Results are the mean ± standard deviation for three determinations. N: normal colon mucosa; C: colon cancer.

As expected, owing to the strong prevalence of such transcript in these tissues, the results were very similar to those previously obtained measuring the B3GALT5 coding sequence [71, 72], which is common to all transcripts. Expression levels, in fact, ranged from 11 to 1 fg/pg of β-actin (average 3.9) in normal colon mucosae and from 0.5 fg/pg of β-actin to undetectable values (average 0.13) in colon cancers (Figure 5.1). Referred to the matched pairs, the amount of LTR transcript decreased in cancer from 4- to >200-folds.

To assess the relationship between LTR transcript and candidate transcription factors HNF1α/β and Cdx1/2, we determined their amounts in nuclear protein extracted only from two normal mucosa specimens (Figure 5.2), since we had not enough material available for extracting nuclear protein from the others. We detected HNF1α/β, but their amounts were not related to those of B3GALT5 LTR transcript. We did not find Cdx1/2 in the mucosa samples analyzed.

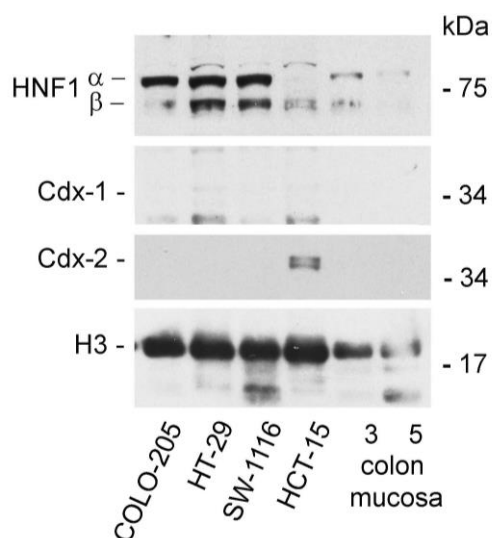
Figure 5.2 Detection of transcription factors HNF1α/β, Cdx1/2, and histone H3 in two normal mucosa samples and in some cell lines. Nuclear extracts (5-10 µg of protein) were separated by 10% SDS-PAGE and transferred to

a nitrocellulose membrane that was blotted with anti HNF1 antibody (recognizing both HNF1α and HNF1β), or anti-Cdx1, anti-Cdx2, or anti-histone H3 antibodies followed by HRP-labeled secondary antibody and chemiluminescence detection, as detailed under "Materials and Methods". Longer exposures were necessary for Cdx1 and Cdx2 to detect a visible spot.

5.2 Detection of HNF1α/β and Cdx1/2 in colon cancer biopsies and cell lines expressing different amounts of B3GALT5 LTR transcript

Since we had enough material available for extracting nuclear protein from eight cancer biopsies but from only two normal mucosa specimens, we then compared cancer samples with various cell lines presenting a wide range of expression of the LTR transcript. Surprisingly, we found that in all cancer samples tested HNF1α/β were easily detected, while Cdx1/2 were almost undetectable, irrespective of the amount of LTR transcript expressed (Figure 5.3). We also found that COLO-205 and SW-1116 cells, expressing the highest amount of LTR transcript (about 15 and 6 fg/pg of β-actin, respectively), expressed only HNF1α, at levels similar to those of samples expressing low or no LTR transcript, such as hepatoma cell lines Huh-7 and Hep-3B. MKN-45 and HT-29 cells expressed both HNF1α and HNF1β, but the transcript levels were lower than in COLO-205 or SW-1116 cells. On the other hand, MDA-MB-361 cells expressed only HNF1β, and a little more LTR transcript than MKN-45 or HT-29 cells. Interestingly, the cells lacking HNF1α/β also lacked LTR transcript, but not *vice versa*.
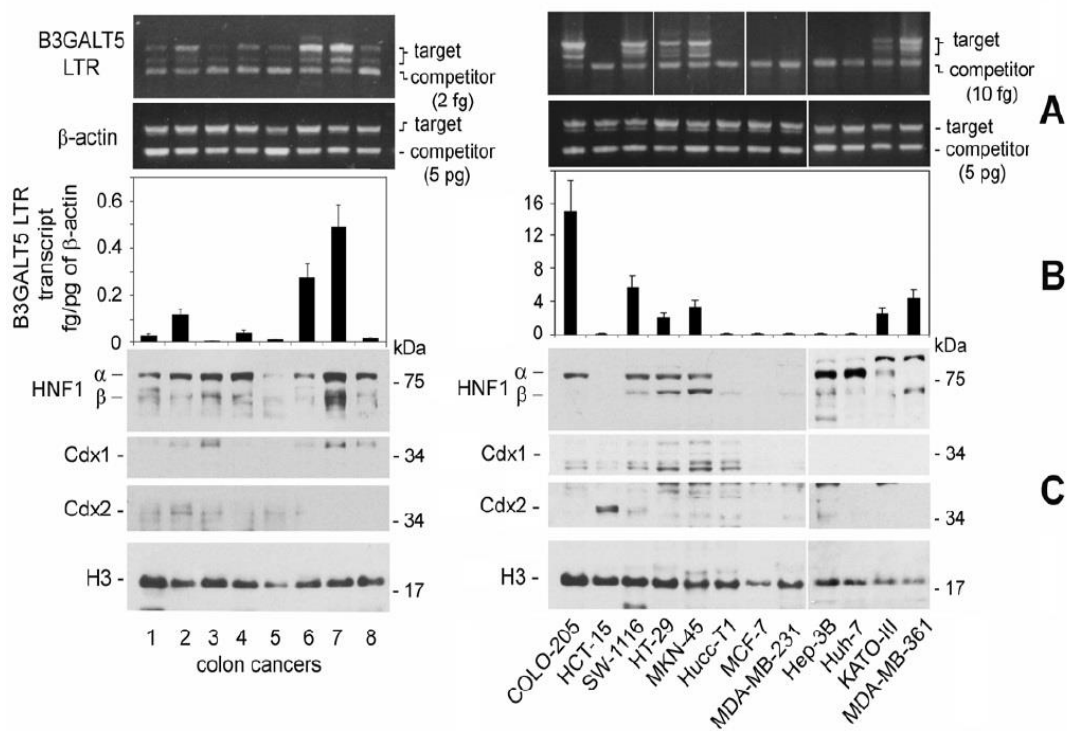
Figure 5.3 Detection of transcription factors HNF1α/β and Cdx1/2 in human colon cancers and in cell lines expressing different amounts of B3GALT5 LTR transcript. (A) Detection and (B) quantification of B3GALT5 LTR transcript by competitive RT-PCR was performed as in figure 5.1. Note the different amounts of B3GALT5 LTR competitor and scales used for colon cancer and cell lines.(C) Western blot detection of HNF1α/β, Cdx1/2 and histone H3, was performed as in figure 5.2. Longer exposures were necessary for Cdx1 and Cdx2 to detect a visible spot.

Cdx1 and Cdx2 remained undetectable under reference conditions in all tested samples. A weak Cdx2 spot became detectable after a longer exposure in HCT-15 cells and in a couple of cancer biopsies expressing no LTR transcript at all. To assess the sensitivity of anti-Cdx antibodies, we analyzed the nuclear protein extracted from HEK-293T cells transfected with Cdx1 or Cdx2 cDNAs and found very strong signals under the conditions used for detection (Figure 5.4).
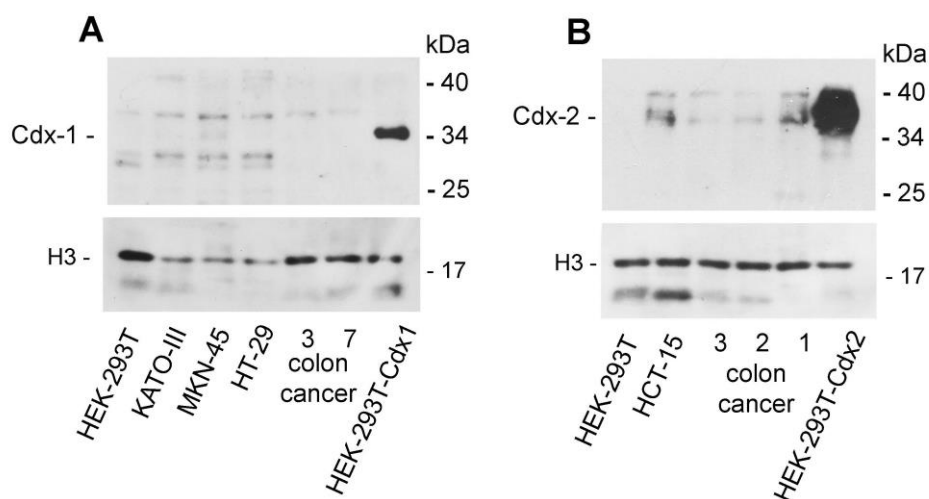
Figure 5.4 Detection of Cdx1, Cdx2 and histone H3 proteins in human colon cancers, some reference cell lines and in HEK-293Tcells transfected with Cdx1 or Cdx2 plasmids, was performed as reported in figure 5.2.

5.3 Detection of B3GALT5 LTR transcript in cell models expressing various amounts of HNF1α or HNF1β

To elucidate the specific role of transcription factors involved in the regulation of *B3GALT5* transcription, we compared the expression levels of B3GALT5 LTR transcript, quantitated by competitive RT-PCR, with those of candidate transcription factors HNF1α/β determined by western blot. For this purpose, we used various cell models serving as controls in which the amounts of HNF1α or HNF1β were modulated either by transfection with their cDNA or by inhibition with shRNA.

5.3.1 Effect of HNF1α/β overexpression on the levels of B3GALT5 LTR transcript

To better assess the role of HNF1α/β, we planned to overexpress such factors in a suitable cell model. First, we determined the expression levels of HNF1α/β transcripts in cell lines not expressing the protein by Western blotting, such as HCT-15 and HuCC-T1 cells (Figure 5.5).

Figure 5.5 Expression of HNF1α/β transcripts in various cell lines. Detection (lower segments) and quantification (upper segments) of HNF1α/β transcripts by competitive RT-PCR was performed and presented as in figure 5.1. β-actine samples for cDNA normalization are those presented in figure 5.3

Surprisingly, we found that both lines expressed HNF1β transcript at remarkable levels. We thus extended the analysis to many other cell lines, and found that HNF1β RNA was commonly expressed even in cells where the protein is undetectable by Western blotting (Figure 5.5), suggesting a crucial post-transcriptional regulation of the gene. This phenomenon was not evident in the case of HNF1α (Figure 5.5). The breast cancer cell line MDA-MB-231, lacking both HNF1α/β transcripts, was thus chosen for transfection experiments.
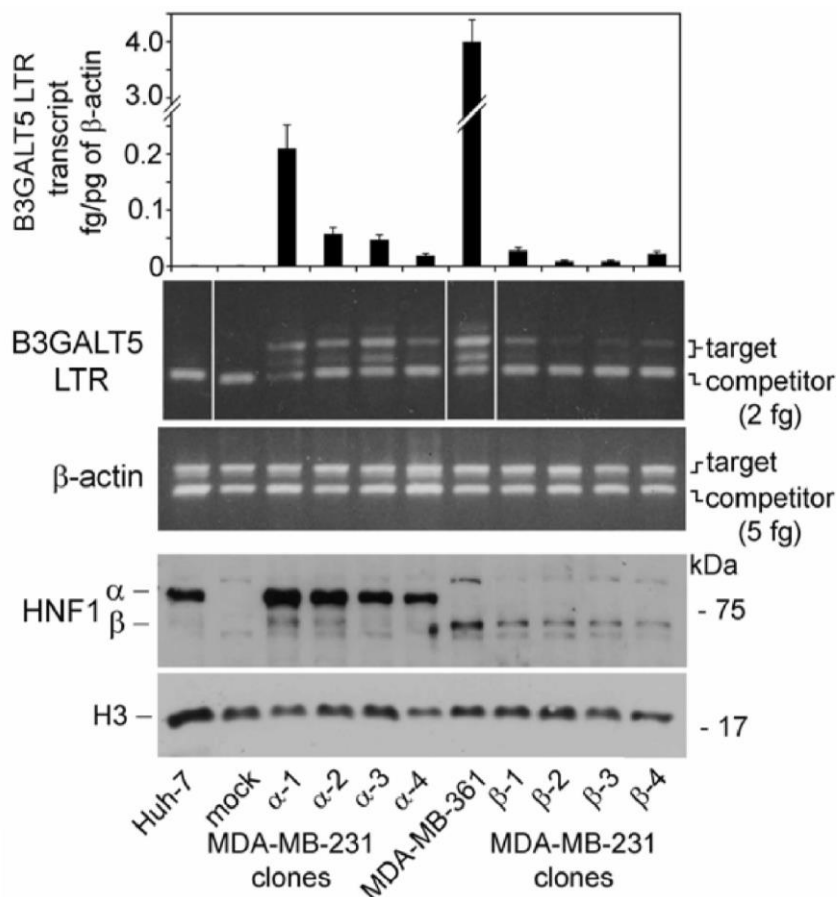
Figure 5.6 Effect of HNF1α/β over-expression on the levels of B3GALT5 LTR transcript. MDA-MB-231 cells, not expressing either HNF1α or HNF1β, were transfected with either factor and the clones were screened by Western blotting with anti-HNF1 antibody. RNA and nuclear proteins were extracted from positive clones and used for B3GALT5 LTR transcript quantification, and HNF1α/β detection, as reported in previous figures. A mock-transfected clone, obtained using the resistance plasmid mixed with pcDNA3 vector alone, was used as control. Huh-7, expressing high levels of HNF1α but not LTR transcript, and MDA-MB-361 cells, expressing HNF1β only and moderate amount of LTR transcript, were also used as references

The obtained clones, analyzed by Western blotting (Figure 5.6), were found to express varied but rather high levels of HNF1α, compared with Huh-7 cells, and varied but low levels of HNF1β, compared with MDA-MB-361 cells. B3GALT5 LTR transcript became detectable, although at low levels, in all the clones expressing either HNF1α or HNF1β (Figure 5.6). The amount of LTR transcript in fact ranged from 0.05 to 0.5 fg/pg of β-actin. However, even a weak expression of HNF1β was able to turn on transcription of LTR mRNA. In particular, clones β-1 and α-4 expressed similar levels of LTR transcript, while the amount of HNF1β in the former was much lower than that of HNF1α in the latter (Figure 5.6).

Notwithstanding the low levels of HNF1β protein in the clones, the corresponding amount of transcript was very high, the highest detected among tested cells (Figure 5.7), confirming that regulation of the HNF1β gene is largely post-transcriptional.
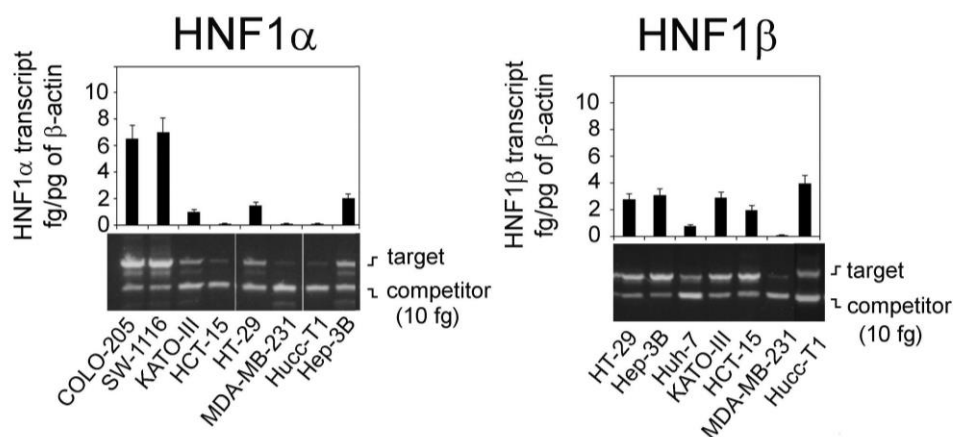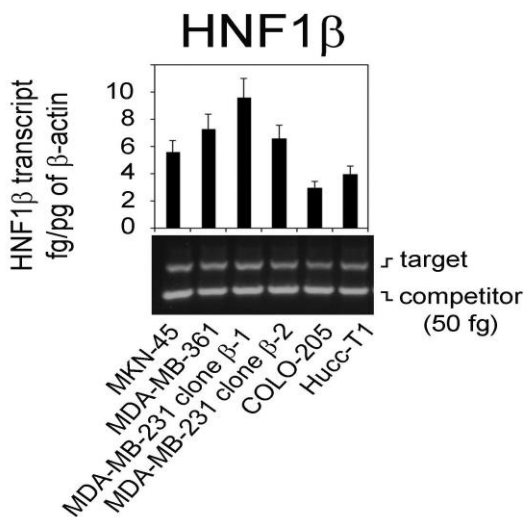

Figure 5.7 Expression of HNF1β transcripts in various cell lines and clones. Detection (lower segments) and quantification (upper segments) of HNF1β transcripts by competitive RT-PCR was performed and presented as in figure 5.1. β-actin samples for cDNA normalization are those presented in figure 5.3

5.3.2 Effect of HNF1α/β silencing on the levels of B3GALT5 LTR transcript

To focus on the role of HNF1β, we silenced HNF1α in MKN-45 cells, which express both factors, and HNF1β in MDA-MB-361, which express HNF1β only. Using plasmids encoding shRNAs targeting either factor, we found that gene silencing was lost after prolonged cell replications (over 4 weeks). Nevertheless, we obtained a transient but strong reduction of expression of HNF1α in MKN-45 cells (Figure 5.8 upper panel) and of HNF1β in MDA-MB-361 cells (Figure 5.9, upper panel), respectively.
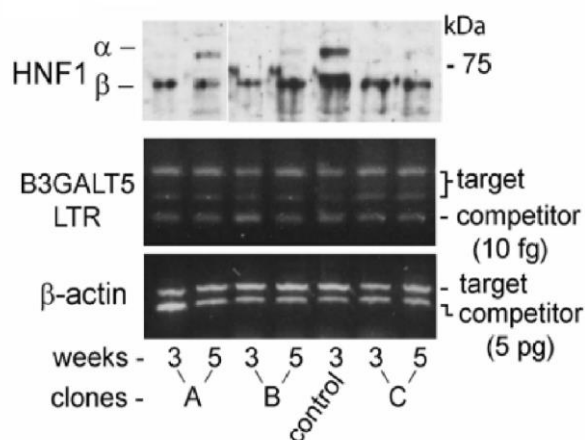
Figure 5.8 Effect of HNF1α silencing on MKN-45 cells on the levels of B3GALT5 LTR transcript. MKN-45 cells, expressing both HNF1α and HNF1β, as well as B3GALT5 LTR transcript, were transfected with pLKO plasmid bearing a shRNA sequence specifically targeting HNF1α. Individual clones resistant to puromycin selection were allowed to grow for a total of 3 weeks and the obtained cells were used for Western blot analysis and RNA extraction, as detailed under "Materials and Methods". A control clone was prepared in parallel transfecting cells with the empty pLKO vector. Western blotting (upper panel) and competitive RT-PCR analyses (lower panel) were then performed as reported in previous figures. Detection of HNF1β was used as internal reference to evaluate HNF1α silencing and re-expression.
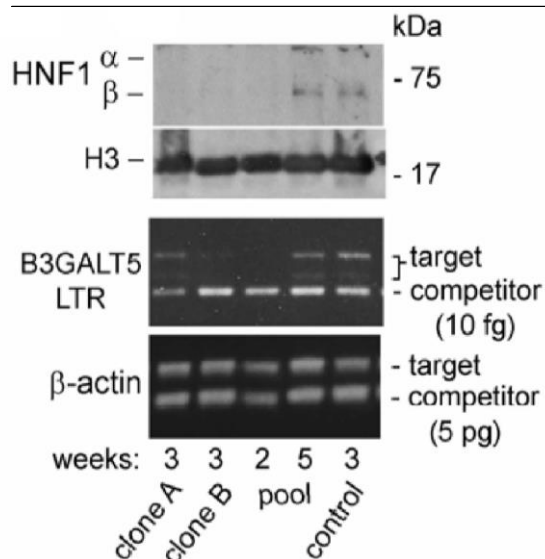
Figure 5.9 Effect of HNF1β silencing on MDA-MB-361 cells on the levels of B3GALT5 LTR transcript. MDA-MB-361 cells, expressing HNF1β only, as well as B3GALT5 LTR transcript, were transfected with pLKO plasmid bearing a shRNA sequence specifically targeting HNF1β. Individual clones resistant to puromycin selection were allowed to grow for a total of 3 weeks and the obtained cells were used for Western blot analysis, RNA extraction, as detailed under "Materials and Methods". A control clone was prepared in parallel transfecting cells with the empty pLKO vector. Western blotting (upper panel) and competitive RT-PCR analyses (lower panel) were then performed as reported in previous figures. Detection of histone H3 was used as internal reference to evaluate HNF1β silencing and re-expression.

Silencing of HNF1α in MKN-45 cells had no effect on LTR transcript levels (Figure 5.8, lower panel), while the silencing of HNF1 β in MDA-MB-361 cells was accompanied by a dramatic reduction of the transcript (Figure 5.9, lower panel). After longer times following transfection, recovery of HNF1α expression in MKN-45 cells did not affect LTR transcript levels, while recovery of HNF1β expression in MDA-MB-361 cells restored the amount of the transcript. These data indicate that HNF1α and HNF1β play an interchangeable and not cumulative role, and suggested that, although necessary for *B3GALT5* transcription, they are not sufficient to explain transcript modulation.

5.4 Effect of TSA and 5AZA treatment on the expression of B3GALT5 LTR transcript

To remove potential epigenetic constraints, we performed TSA and 5AZA treatments, starting with MKN-45 cells since they were found to be sensitive to such treatments in the case of the native B3GALT5 transcript [78].
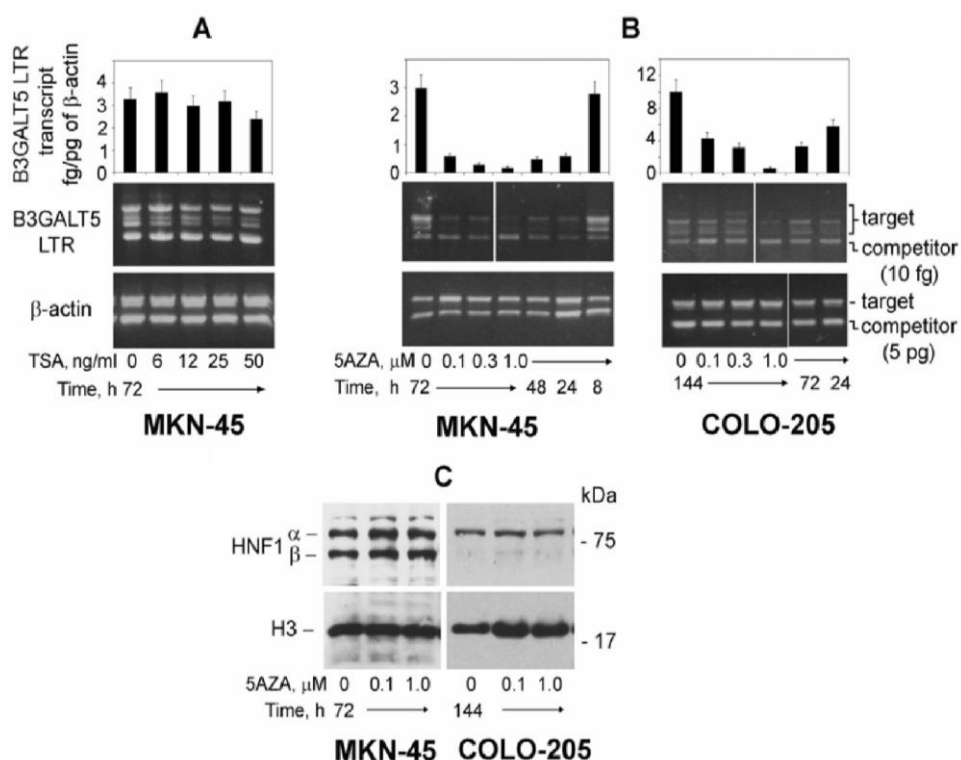
Figure 5.10 Effect of TSA and 5AZA treatment on the expression of B3GALT5 LTR transcript in cultured cells. A) MKN-45 cells were treated with various amounts of histone deacetylase inhibitor TSA. B) MKN- 45 and COLO-205 cells were treated with various amounts of DNA methyltransferase inhibitor 5AZA for different times, as detailed under "Materials and Methods". In both panels, detection and quantification of B3GALT5 LTR transcript by competitive RT-PCR were performed as in previous figures. Treatments were performed in duplicate. At the end of the treatments, cells were harvested, washed with PBS, and processed to

extract RNA and nuclear protein from the same cell suspension. Results are the mean ± standard deviation for two determinations performed on each duplicate. C) Nuclear protein extracted from representative data-points were submitted to Western blot analysis as reported in previous figures.

We found that TSA treatment (Figure 5.10A) had no effect on the LTR transcript, while 5AZA, surprisingly, strongly impaired expression (Fig. 5.10B). We repeated the treatment on COLO-205 cells, since they express LTR transcript at the highest levels found, and obtained similar results (Figure 5.10B). In fact, even in these cells 5AZA treatment impaired LTR expression in a dose- and time-dependent manner. In both treated cell lines B3GALT5 LTR transcript dropped down to the levels measured in some colon cancer biopsies or HNF1α/β-transfected MDA-MB-231 clones, while the levels of HNF1 proteins remained almost unaffected (Figure 5.10C). Since the 650-bp long LTR transposon contains only seven dispersed CG pairs, and no CpG island is present in the proximal sequences, this result suggests that the regulatory effect of methylation resides outside and far from the LTR promoter.

5.5 In vitro evaluation of the B3GALT5 promoter

To confirm the hypothesis that the B3GALT5 LTR promoter although necessary is unable to regulate transcription *per se*, we performed luciferase assays with a reporter plasmid in which luciferase is placed under the control of the LTR sequence, and EMSAs using the LTR sequence as a probe. Taking advantage of the availability of cell lines and clones showing various combinations of the expression levels of the B3GALT5 LTR transcript and HNF1α/β proteins, we used them as the host cells for transfecting the LTR/luciferase plasmid or as the sources for extracting nuclear protein to be tested with the LTR probe by EMSA.

5.5.1 Evaluation of the activation properties of B3GALT5 LTR promoter by luciferase assays

The results of luciferase activity promoted by B3GALT5 LTR promoter are show in figure 5.11. The pGL3 vector carrying the sequence -148-28 from the transcription initiation site of B3GALT5 LTR mRNA and firefly luciferase gene (pGL3-LTR), was transfected together with a Renilla luciferase reporter expression plasmid in various cell lines expressing different amounts of HNF1α/β and B3GALT5 LTR transcript.
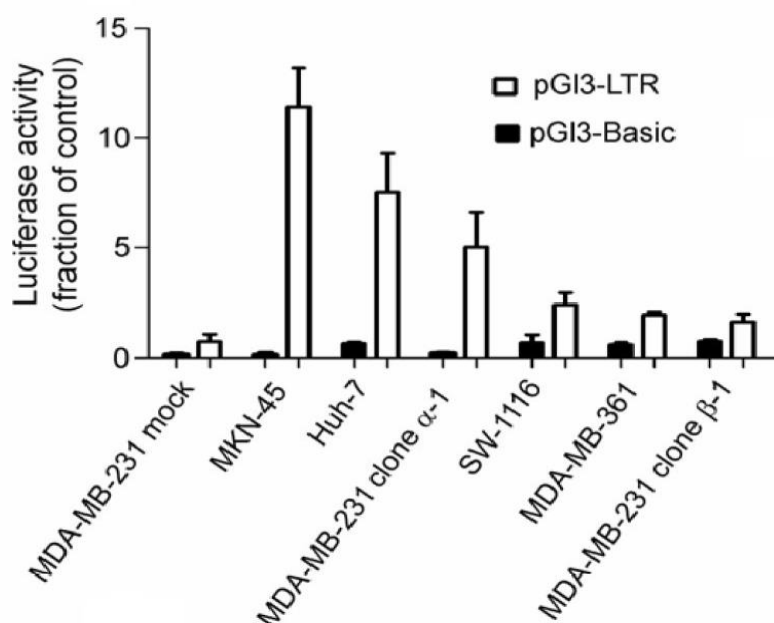
Figure 5.11 Luciferase activity assay of B3GALT5 LTR promoter in various host cells. The sequence -148-28 from the transcription initiation site of B3GALT5 LTR mRNA, was cloned in the pGL3 vector carrying the firefly luciferase gene (pGL3-LTR), and transfected together with a Renilla luciferase reporter expression plasmid. Host cells were cell lines or clones expressing different amounts of HNF1α/β and B3GALT5 LTR transcript used in previous experiments. pGL3-control vector, bearing the SV40 promoter, and pGL3-basic vector, lacking any promoter sequence, were also transfected for normalization and control. Values are the means ± standard variations for two experiments in triplicate.

66

By luciferase reporter assays the highest activity was found in MKN-45 cells, expressing high levels of both HNF1α and HNF1β and moderate levels of LTR transcript, followed by Huh-7 cells and MDA-MB-231 clone α-1, which expressed high levels of HNF1α only and no or low LTR transcripts. Conversely, much lower luciferase activity was detected in SW-1116 cells, expressing high levels of transcript and moderate levels of HNF1α. In MDA-MB-361 cells expressing moderate levels of transcript but low HNF1β alone, luciferase activity was minimal, and almost undetectable in MDA-MB- 231 clone β-1, which expresses minimal amounts of both LTR transcript and HNF1β.

We also transfected HEK-293T cells with pGL3-LTR and Renilla luciferase plasmids in the presence of transcription factor cDNAs cloned in the pcDNA3 vector (Figure 5.12). HEK-293T cells lack expression of HNF1α/β and Cdx1/2, and are able to replicate plasmids with the SV40 origin of replication (as pcDNA3). Addition of each plasmid coding HNF1α/β or Cdx1/2 enhanced luciferase activity, with preference for HNF1α followed by HNF1β and Cdx2, while Cdx1 was less effective.
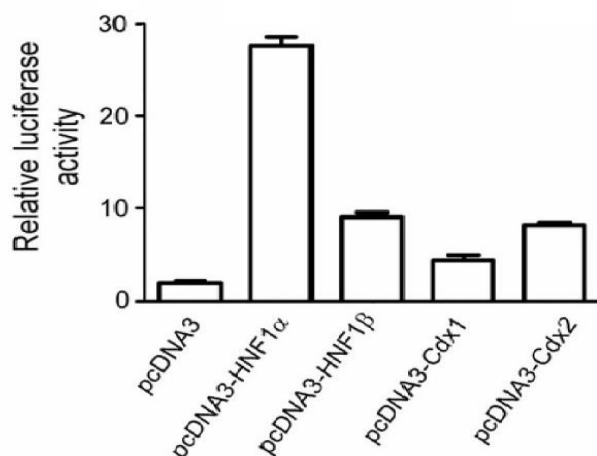
Figure 5.12 Luciferase activity assay of B3GALT5 LTR promoter co-transfected with the candidate transcription factors. pGL3-LTR and Renilla luciferase plasmids were mixed with expression plasmids having the coding

region of transcription factors cloned in the pcDNA3 vector, which bears the SV40 origin of replication and were transfected into HEK-293T cells. Firefly luciferase activity was measured 24 h later and expressed relative to the Renilla luciferase activity determined for each sample. Transfection with pGL3-LTR and Renilla luciferase plasmids mixed with pcDNA3 vector alone was performed as control. Values are the means ± standard variations for two experiments in triplicate.

5.5.2 Evaluation of the binding properties of B3GALT5 LTR promoter by EMSA assays

By EMSA, we found that the LTR sequence, when used as a probe, formed one specific complex plus one or two other non-specific complexes (Figure 5.13). The one specific complex was the most retarded and appeared as a doublet, corresponding to HNF1α/β binding. In fact, it was found only with nuclear protein extracted from cell lines expressing HNF1α/β, was affected by anti-HNF1 antibody, and was competed out by an excess of the unlabeled probe sequence (Figure 5.13A). Such a complex was also evident with nuclear proteins extracted from cells expressing very low to undetectable levels of LTR transcript but very high levels of HNF1α, such as Huh-7, Hep-3B, and MDA-MB-231 clones α-1 or α-3. The complex was detectable but much less evident with the protein extracted from MDA-MB-361 cells, expressing moderate levels of LTR transcript and low levels of HNF1β, and undetectable with MDA-MB-231 clone β-1, expressing very low levels of both. The other complexes were not competed out by the unlabeled probe, and should be considered non-specific (Fig. 5.13B).
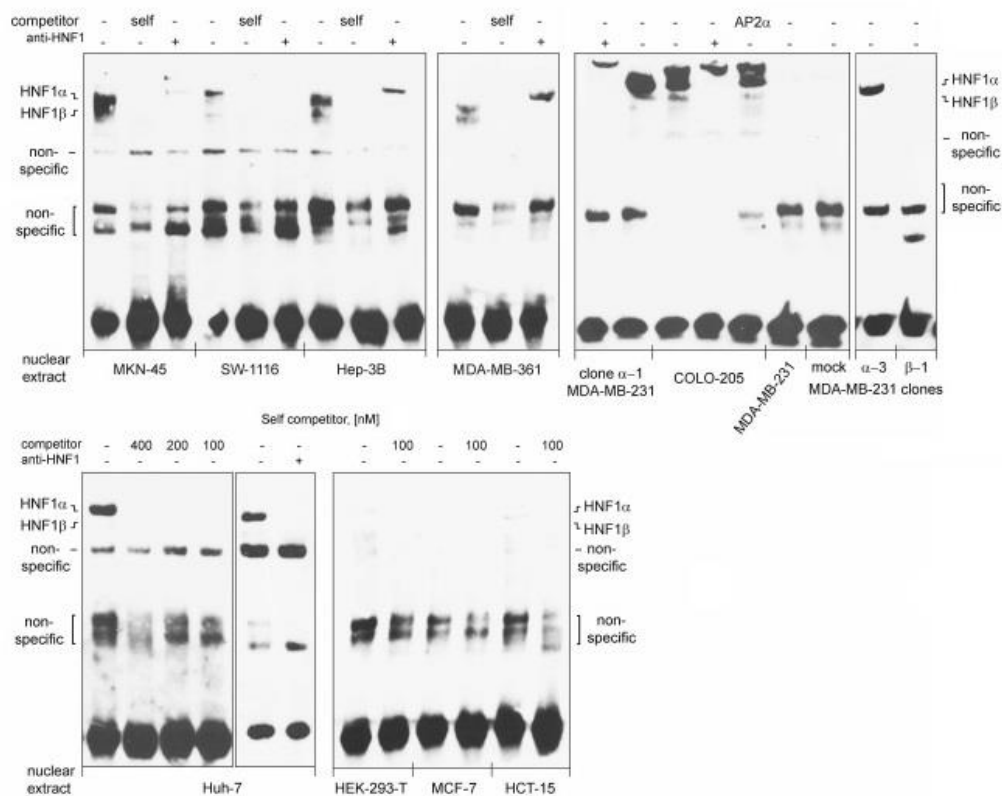
Figure 5.13 Characterization of the binding properties of B3GALT5 LTR promoter using cell lines or clones expressing different amounts of HNF1α/β and B3GALT5 LTR transcript. EMSAs were performed incubating a biotinylated oligonucleotide probe (nucleotides -151-112 calculated from the transcription initiation site of B3GALT5 LTR mRNA) with the nuclear protein extracted from the mentioned cell lines or clones.

In particular, the complex migrating as the less-retarded doublet was formed by almost any nuclear extract, including those prepared from cells not expressing B3GALT5 LTR and unable to drive luciferase activity upon transfection with the LTR construct. It migrated much faster than true Cdx1 or Cdx2 complexes formed by the same LTR probe with authentic Cdx1 or Cdx2 protein, and was not competed out by the SIF1 sequence, as

occurred with the true Cdx1 or Cdx2 complexes formed by recombinant factors (Figure 5.14). Although the LTR promoter was able to bind Cdx1/2 *in vitro*, taken together the data indicate that the HNF1 binding site is the only functional sequence brought by the insertion of the LTR transposon.
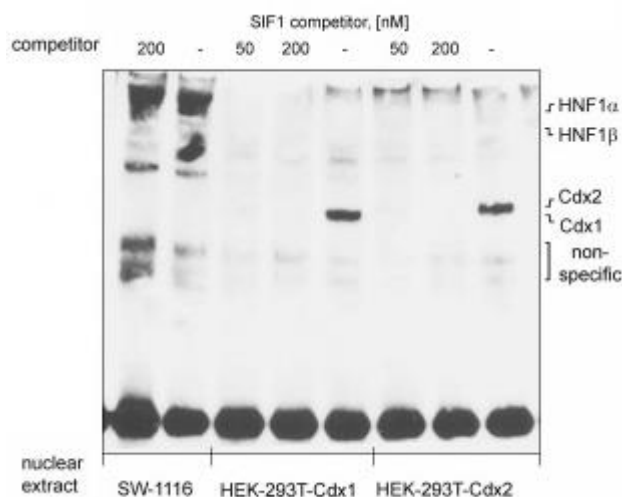
Figure 5.14 Characterization of the binding properties of B3GALT5 LTR promoter using HEK-293T cells expressing recombinant transcription factors Cdx1 or Cdx2 and SW-1116 cells as control. EMSAs were performed incubating a biotinylated oligonucleotide probe (nucleotides -151-112 calculated from the transcription initiation site of B3GALT5 LTR mRNA) with the nuclear protein extracted from the mentioned cell lines or clones.

6. DISCUSSION

Literature data [71] reported that transcriptional regulation of *B3GALT5* gene was under the cooperative control of HNF1 and Cdx as for other genes specifically expressed in intestine like sucrose-isomaltase [182], lactase-phlorizin hydrolase (LPH) [189] and claudin-2 [190]. Isshiki *et al.* [71], in their study attributed a positive role for HNF1α and Cdx2 in regulating B3GALT5 transcript whereas HNF1β and Cdx1 did not affect the levels of the transcript. Conversely, we observed that Cdx1/2 are not involved, while HNF1α or HNF1β are necessary to activate the B3GALT5 LTR promoter, playing an interchangeable and not cumulative role, but they are unable to modulate transcription, which instead depends on distal regulatory element(s) active when methylated, and which are unknown at present.

The complex role played by HNF1 was defined by several lines of evidence. First of all, in any cell line lacking both HNF1α and HNF1β the B3GALT5 LTR transcript was always undetectable; moreover, transfection of one of these cells with either HNF1α or HNF1β cDNA induced LTR transcription. In addition, shRNA-mediated silencing of HNF1β in cells expressing only HNF1β, strongly impaired LTR expression. These results clearly indicate that HNF1 is necessary for transcription. Since the transcript was detected in cell lines and clones expressing either HNF1α or HNF1β alone, and shRNA-mediated silencing of HNF1α in cells expressing both HNF1α and HNF1β does not affect LTR expression, the two forms appear to not be cooperative, but interchangeable, as reported in the gut [191] but not in the kidney [176]. HNF1α and HNF1β bind DNA with the same sequence specificity and can form homo- or heterodimers [176]. The fact that these two genes share the same DNA binding specificity and overlapping expression pattern, together with their ability to form heterodimers, suggests that these proteins could play complementary roles. However, HNF1α seems to be a more potent transactivator than is

HNF1β in transient transfection assays [176], as we also corroborate with our study.

On the other hand, the expression levels of either HNF1 form, or of both forms when expressed together, cannot explain the wide expression range of the LTR transcript. In fact, we detected comparable amounts of HNF1 in COLO-205 or SW-1116 cells, which expressed the highest levels of LTR transcript, and in cells, clones or colon cancer biopsies which expressed 10 or 100-fold less, or even undetectable, LTR transcript. This evidence prompts the question about what other factor or mechanism affects B3GALT5 LTR transcription and is responsible for cancer down-regulation.

We found that Cdx1/2 were not involved, although they were able to bind and activate the LTR promoter *in vitro*. In fact, the amounts of Cdx1/2 in cells or tissues expressing B3GALT5 LTR transcript were negligible and the few samples expressing detectable amounts of Cdx2 (HCT-15 cells and colon cancer biopsies 1 and 2) lacked B3GALT5 LTR transcript. Moreover, nuclear protein extracted from cells expressing the transcript never formed Cdx1/2 complexes in EMSA. The less-retarded doublet detected by EMSA with the LTR probe, previously proposed to depend on Cdx1/2 binding [71] did not behave as true Cdx1/2 complexes, but was found to be a non-specific artifact, formed by any nuclear protein extract, including those from cells not expressing the transcript at all. Altogether, these data indicate that the HNF1 binding site is the only functional part of the LTR promoter, and that no other binding sites, for stimulatory or inhibitory factors, including Cdx1/2, are physiologically relevant.

Conversely, DNA demethylation obtained through 5AZA treatment of cells expressing the B3GALT5 LTR transcript reproduced *in vitro* the down-regulation of the transcript observed among cell lines and cancer biopsies *in vivo.* In fact, in treated cells, the levels of B3GALT5 LTR transcript decreased from 3-10 to less than 0.2 fg/pg β-actin, while the amounts of

HNF1 remained almost unchanged. Since LTR transposon contain only seven CG pairs (Figure 6.1) and proximal sequences do not contain CpG islands, we conclude that methylation-sensitive DNA sequences represent element(s) involved in transcriptional regulation residing outside the LTR sequence, probably distant from the promoter.

```
TGTGATGGTT ACTTTTAGGT GTCAACTTGG CTGGATTAAT AAATACCTAG AGAACTGGTA
AAGCATTATT TCTGGGTGTG TTTGTGAAGG TGTTTCCAGA GGAGATTGGC TGTGAGTCAG
TGGGCTGAGT GGGGAGGAGC TGCCCTCCAT GTGGGCAGGC ACCATCCATT GACTGGGCCC
AGATAGAACA AGAAGGCAGA AGAAATGTGA ATTCCTCTTT CTCTGCTGGA GCTGGGATAT
TCTTCTTCTC CTGCCCTTGG ACATCAGAGC TGCAGGCTCT CTGGCCTTTG GACCCGAGGA
TTTATACCAA GCAGGTTTCT GGGTTCTCAG GCCTTTGGCC TTGGACTGAT AGTTACACCA
TTGGCATATC TGGTTCTGAG GCTCTTGGTC TTGGACTGAG CCACACTCCT GGCATCCCAG
CGTCTCCAGC TTGCATGGCC TGTCACGGTA TTTCCCAACC TCCGTAATCA CGCTAGCCAA
TTCTTCTAAG AAATTTCTTC TCATCTATCT GTCTGTCTAT CTATCTATCT GTCTACCTAC
CGACTTACCT ACCTACCTGC CTATCTATCT TTTGATTAAT CTACCTATCA ATCTTTCTAT
CTATCCATAA CCTGTTGATT CGATCTCTCT AGAGAACCCT GACTAATACA
```

Figure 6.1 Nucleotide sequence of B3GALT5 LTR transposon. Exon 1 of B3GALT5, framed by a red line; CpG dinucleotides, green boxes and putative HNF1/Cdx binding site, yellow box.

Alignment of the LTR sequence and the whole *B3GALT5* gene in the context of chromosome 21 revealed a single typical promoter-associated CpG island, already characterized in detail and responsible for regulating transcription of the native B3GALT5 mRNA [78]. In addition, a number of very short CpG islands were detected, one in an intron and the others in the intergenic regions (Figure 6.2). Unfortunately, due to the extremely high homology (> 98%) of this human sequence with that of the other primates sharing the LTR transposon, no prediction can be made *in silico* about the relevance of any of such island. Alternatively, we could suppose that the region involved is over 0.1 Mbp away from the LTR sequence.

We thus propose a model of *B3GALT5* transcription predicting that a distant DNA region, methylated in normal colon mucosa and cell lines such as COLO-205 or SW-1116, is demethylated in colon cancer, in several cell lines and even in other tissues still expressing HNF1; the degree of

demethylation downregulates transcription even to complete silencing, as in Hep-3B and Huh-7 cells.

DNA hypomethylation was indeed the first epigenetic abnormality detected in human cancers about 30 years ago [192]. However, it received much less attention in the last decade than the opposite modification, hypermethylation, frequently associated with silencing of tumor-suppressor and other genes. Only recently, high-resolution genome-wide analyses revealed an independent and relevant role of hypomethylation in cancer formation and progression [141, 193]. Among various DNA sequences affected by cancer-associated hypomethylation, transcription control elements appeared to be of special interest. In particular, it was found that genes associated with hypermethylated distant control elements are more frequently downregulated in cancer [167].

No information about the nature, exact location, or mechanism of action of such distant sequence is available for *B3GALT5* transcription at present. It may represent a typical CpG island, or a "shore" as recently proposed [147] (see after), a non-CpG methylation site [150] or instead include stretches of CpG dinucleotides shorter than CpG islands associated with promoters.
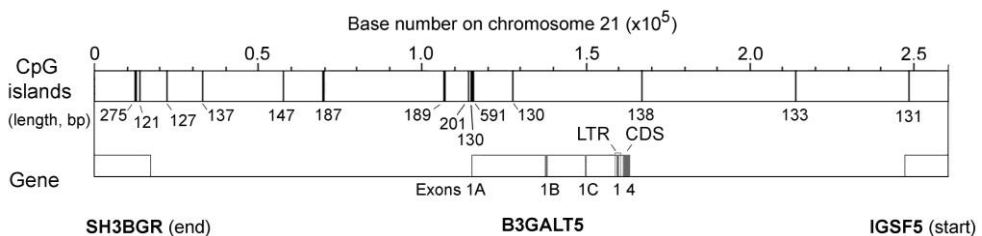
Figure 6.2 CpG islands detected in the context of the B3GALT5 gene. The genomic sequence of chromosome 21 (0.26 Mbp) encompassing B3GALT5 and intergenic regions was analyzed with EMBOSS Cpgplot software (EMBL), with the following parameters: observed:expected ratio >0.6, C+G >50%, and minimum length, 120 bp. One CpG island was present near exon 1A, which regulates the cognate promoter [78]. Several shorter stretches of CpG dinucleotides were instead detectable in the intergenic regions.

Many studies on cancer methylation have assumed that functionally important DNA methylation occurs in promoters and that most DNA methylation changes in cancer occur in CpG islands. However, it was proved that many methylation alterations in colon cancer do not occur in promoters or in CpG islands, but in sequences up to 2 kb distant termed 'CpG island shores' [147]. CpG island shore methylation is strongly related to gene expression, and it is highly conserved in mouse, discriminating tissue types regardless of species of origin.

Particularly, methylation of stretches of CpG dinucleotides, as those present in the intergenic regions of *B3GALT5*, is emerging as a relevant aspect of transcriptional control [157], being responsible for recruitment of alternative promoters, regulation of non-coding RNA synthesis, or modulation of enhancer activity. In particular, hypomethylation of enhancer sequences is reported to negatively regulate transcription in cancer and during tissue differentiation [157]. The occurrence of distal regulatory elements binding transcription factors in a methylation-dependent manner was recently reported in breast cancer [167]. Other types of genome-wide analyses predicted that HNF1 is able to play such a role in the kidney [194] and in the liver [195]. *B3GALT5* transcription thus represents a promising model to address such novel issues, since hypomethylation of distant sequences, acting on the LTR transcript, and promoter hypermethylation, acting on the native transcript [78], cooperate on one gene to obtain full cancer-associated silencing.

It is generally accepted that transposable elements are common targets of DNA methylation, and that their transcription is increased in hypomethylated cellular environments [103, 196]. It is striking that the LTR-driven transcripts of many mammalian genes are expressed early in development, in tumors, and in tissues such as the placenta, where the level of DNA methylation is generally lower [87, 127, 196]. This suggests

that they may be regulated in part by DNA methylation. In contrast, B3GALT5 LTR promoter is the first described example of LTR promoter having a positive correlation between promoter expression and DNA methylation.

Our present findings also contribute to better understanding of the mechanism of evolutionary stabilization of the LTR transposon with respect to the models hypothesized for other mobile elements [197]. In fact, previous evidence suggested that the transcription factors binding sites in the B3GALT5 LTR promoter were present in the original consensus sequence for this class of LTRs, maintaining the functional roles [76]. Such hypothesis is based on the sequence comparison between the specific LTR copy, the LTR consensus sequence, and insertions in other primates. It was proposed a model by which the LTR promoter was already active at the time of insertion [76] and may have taken advantage of the analogous sequence present in the type B promoter, which is much less active due to its opposite orientation [77]. Now we propose that effective interaction with a distant element played a crucial stabilizing effect, probably due to the precise localization of insertion, which in turn allowed not only stronger HNF1-associated expression, but even the finest tissue specificity obtained through the overlapping epigenetic control operated by the distal element.

7. CONCLUSIONS

These results appear self-consistent concerning the interchangeable but not cumulative role of HNF1α and HNF1β in the regulation of *B3GALT5* transcription. HNF1α/β are necessary but not sufficient to drive expression of LTR transcript, although they are both able to bind and activate the LTR promoter *in vitro.* They are not immediately responsible for cancer down-regulation, which instead depends on distal regulatory elements active when methylated and unknown at present. Moreover, the HNF1 binding site is the only functionally relevant binding site present in the LTR promoter since the Cdx1/2 binding site do not appear involved.

We also propose that the successful insertion and activation of B3GALT5 LTR promoter during evolution depended on its HNF1 binding site as well as on the distal regulatory element(s) sensitive to methylation, perhaps playing a positional role. The data thus suggest the existence of an unprecedented regulatory mechanism involving methylation of distal DNA elements as a requirement for transcriptional activation. Since the LTR retrotransposon is present in some primates only, which shares over 95% identity in the flanking DNA regions, no information from evolutionary conserved sequences is available. As future strategies, we are trying to recognize potential CpG islands and/or shores in the flanking regions (+/- 100 kBp from the LTR sequence) with the aim to perform bisulfite sequencing using model cell lines as COLO-205 (expected hypermethylated) and MDA-MB-231 (expected hypomethylated). We are also trying to determine the potential enhancer activity of selected distal DNA region through cloning in luciferase reporter vectors and assaying upon transfection in model cell lines.

8. REFERENCES

[1]   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.

[2]   J. L. Magnani, Z. Steplewski, H. Koprowski et al., "Identification of the gastrointestinal and pancreatic cancer-associated antigen detected by monoclonal antibody 19-9 in the sera of patients as a mucin," *Cancer Res,* vol. 43, no. 11, pp. 5489-92, Nov, 1983.

[3]   T. Yue, K. A. Maupin, B. Fallon *et al.*, "Enhanced discrimination of malignant from benign pancreatic disease by measuring the CA 19-9 antigen on specific protein carriers," *PLoS One,* vol. 6, no. 12, pp. e29180, 2011.

[4]   K. Shimodaira, J. Nakayama, N. Nakamura *et al.*, "Carcinoma-associated expression of core 2 beta-1,6-N-acetylglucosaminyltransferase gene in human colorectal cancer: role of O-glycans in tumor progression," *Cancer Res,* vol. 57, no. 23, pp. 5201-6, Dec 1, 1997.

[5]   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.

[6]   K. Sasaki, Sasaki, E., Kawashima, K., Hanai, N., Nishi, T. & Hasegawa, M., Japanese Patent, 1994.

[7]   F. Kolbinger, M. B. Streiff, and A. G. Katopodis, "Cloning of a human UDP-galactose:2-acetamido-2-deoxy-D-glucose   3beta-galactosyltransferase catalyzing the formation of type 1 chains," *J Biol Chem,* vol. 273, no. 1, pp. 433-40, Jan 2, 1998.

[8]   M. Amado, R. Almeida, F. Carneiro *et al.*, "A family of human beta3-galactosyltransferases. Characterization of four members of a UDP-galactose:beta-N-acetyl-glucamine/beta-nacetyl-galactosamine beta-1,3-galactosyltransferase family," *J Biol Chem,* vol. 273, no. 21, pp. 12770-8, May 22, 1998.

[9]   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.

[10]   D. Jordon, J. Jagirdar, and M. Kaneko, "Blood group antigens, Lewisx and Lewisy in the diagnostic discrimination of malignant mesothelioma versus adenocarcinoma," *Am J Pathol,* vol. 135, no. 5, pp. 931-7, Nov, 1989.

[11]   H. S. Cooper, M. J. Malecha, C. Bass *et al.*, "Expression of blood group antigens H-2, Le(y), and sialylated-Le(a) in human colorectal carcinoma. An immunohistochemical study using double-labeling techniques," *Am J Pathol,* vol. 138, no. 1, pp. 103-10, Jan, 1991.

[12]   F. G. Hanisch, C. Hanski, and A. Hasegawa, "Sialyl Lewis(x) antigen as defined by monoclonal antibody AM-3 is a marker of dysplasia in the colonic adenoma-carcinoma sequence," *Cancer Res,* vol. 52, no. 11, pp. 3138-44, Jun 1, 1992.

[13]   S. Sakamoto, T. Watanabe, T. Tokumaru *et al.*, "Expression of Lewisa, Lewisb, Lewisx, Lewisy, siayl-Lewisa, and sialyl-Lewisx blood group

antigens in human gastric carcinoma and in normal gastric tissue," *Cancer Res,* vol. 49, no. 3, pp. 745-52, Feb 1, 1989.

[14] C. Cordon-Cardo, V. E. Reuter, C. L. Finstad *et al.*, "Blood group-related antigens in human kidney: modulation of Lewis determinants in renal cell carcinoma," *Cancer Res,* vol. 49, no. 1, pp. 212-8, Jan 1, 1989.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] S. Kikkawa, K. Sogawa, M. Satoh *et al.*, "Identification of a Novel Biomarker for Biliary Tract Cancer Using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry," *Int J Proteomics,* vol. 2012, pp. 108609, 2012.

[20] D. H. Kim, S. J. Oh, C. A. Oh *et al.*, "The relationships between perioperative CEA, CA 19-9, and CA 72-4 and recurrence in gastric cancer patients after curative radical gastrectomy," *J Surg Oncol,* vol. 104, no. 6, pp. 585-91, Nov 1, 2011.

[21] 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.

[22] M. P. Bevilacqua, and R. M. Nelson, "Selectins," *J Clin Invest,* vol. 91, no. 2, pp. 379-87, Feb, 1993.

[23] G. Mannori, P. Crottet, O. Cecconi *et al.*, "Differential colon cancer cell adhesion to E-, P-, and L-selectin: role of mucin-type glycoproteins," *Cancer Res,* vol. 55, no. 19, pp. 4425-31, Oct 1, 1995.

[24] J. D. Marth, and P. K. Grewal, "Mammalian glycosylation in immunity," *Nat Rev Immunol,* vol. 8, no. 11, pp. 874-87, Nov, 2008.

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

[26] 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.

[27] I. P. Witz, "The selectin-selectin ligand axis in tumor progression," *Cancer Metastasis Rev,* vol. 27, no. 1, pp. 19-30, Mar, 2008.

[28] R. Renkonen, P. Mattila, M. L. Majuri *et al.*, "In vitro experimental studies of sialyl Lewis x and sialyl Lewis a on endothelial and carcinoma cells: crucial glycans on selectin ligands," *Glycoconj J,* vol. 14, no. 5, pp. 593-600, Aug, 1997.

[29]   R. Kannagi, "Carbohydrate-mediated cell adhesion involved in hematogenous metastasis of cancer," *Glycoconj J,* vol. 14, no. 5, pp. 577-84, Aug, 1997.

[30]   D. Ichikawa, K. Kitamura, N. Tani *et al.*, "Molecular detection of disseminated cancer cells in the peripheral blood and expression of sialylated antigens in colon cancers," *J Surg Oncol,* vol. 75, no. 2, pp. 98-102, Oct, 2000.

[31]   N. Yamada, Y. S. Chung, K. Maeda *et al.*, "Increased expression of sialyl Lewis A and sialyl Lewis X in liver metastases of human colorectal carcinoma," *Invasion Metastasis,* vol. 15, no. 3-4, pp. 95-102, 1995.

[32]   K. Tozawa, T. Okamoto, N. Kawai *et al.*, "Positive correlation between sialyl Lewis X expression and pathologic findings in renal cell carcinoma," *Kidney Int,* vol. 67, no. 4, pp. 1391-6, Apr, 2005.

[33]   J. Wei, L. Cui, F. Liu *et al.*, "E-selectin and Sialyl Lewis X expression is associated with lymph node metastasis of invasive micropapillary carcinoma of the breast," *Int J Surg Pathol,* vol. 18, no. 3, pp. 193-200, Jun, 2010.

[34]   P. Sozzani, R. Arisio, M. Porpiglia *et al.*, "Is Sialyl Lewis x antigen expression a prognostic factor in patients with breast cancer?," *Int J Surg Pathol,* vol. 16, no. 4, pp. 365-74, Oct, 2008.

[35]   J. W. Dennis, S. Laferte, C. Waghorne *et al.*, "Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis," *Science,* vol. 236, no. 4801, pp. 582-5, May 1, 1987.

[36]   M. Granovsky, J. Fata, J. Pawling *et al.*, "Suppression of tumor growth and metastasis in Mgat5-deficient mice," *Nat Med,* vol. 6, no. 3, pp. 306-12, Mar, 2000.

[37]   K. S. Lau, and J. W. Dennis, "N-Glycans in cancer progression," *Glycobiology,* vol. 18, no. 10, pp. 750-60, Oct, 2008.

[38]   W. D. Hanley, M. M. Burdick, K. Konstantopoulos *et al.*, "CD44 on LS174T colon carcinoma cells possesses E-selectin ligand activity," *Cancer Res,* vol. 65, no. 13, pp. 5812-7, Jul 1, 2005.

[39]   K. T. Lim, K. Miyazaki, N. Kimura *et al.*, "Clinical application of functional glycoproteomics - dissection of glycotopes carried by soluble CD44 variants in sera of patients with cancers," *Proteomics,* vol. 8, no. 16, pp. 3263-73, Aug, 2008.

[40]   C. Hanski, K. Drechsler, F. G. Hanisch *et al.*, "Altered glycosylation of the MUC-1 protein core contributes to the colon carcinoma-associated increase of mucin-bound sialyl-Lewis(x) expression," *Cancer Res,* vol. 53, no. 17, pp. 4082-8, Sep 1, 1993.

[41]   J. J. Ho, B. Siddiki, and Y. S. Kim, "Association of sialyl-Lewis(a) and sialyl-Lewis(x) with MUC-1 apomucin ina pancreatic cancer cell line," *Cancer Res,* vol. 55, no. 16, pp. 3659-63, Aug 15, 1995.

[42]   J. L. Tomlinson, and J. C. Johnson, "Quantification of measurement of femoral head coverage and Norberg angle within and among four breeds of dogs," *Am J Vet Res,* vol. 61, no. 12, pp. 1492-500, Dec, 2000.

[43]   K. M. Hiller, J. P. Mayben, K. M. Bendt *et al.*, "Transfection of alpha(1,3)fucosyltransferase antisense sequences impairs the proliferative

and tumorigenic ability of human colon carcinoma cells," *Mol Carcinog,* vol. 27, no. 4, pp. 280-8, Apr, 2000.

[44]   A. Opolski, A. Laskowska, J. Madej *et al.*, "Metastatic potential of human CX-1 colon adenocarcinoma cells is dependent on the expression of sialosyl Le(a) antigen," *Clin Exp Metastasis,* vol. 16, no. 8, pp. 673-81, Nov, 1998.

[45]   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.

[46]   X. Yin, K. Rana, V. Ponmudi *et al.*, "Knockdown of fucosyltransferase III disrupts the adhesion of circulating cancer cells to E-selectin without affecting hematopoietic cell adhesion," *Carbohydr Res,* vol. 345, no. 16, pp. 2334-42, Nov 2, 2010.

[47]   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.

[48]   O. Saitoh, W. C. Wang, R. Lotan *et al.*, "Differential glycosylation and cell surface expression of lysosomal membrane glycoproteins in sublines of a human colon cancer exhibiting distinct metastatic potentials," *J Biol Chem,* vol. 267, no. 8, pp. 5700-11, Mar 15, 1992.

[49]   N. Yamada, Y. S. Chung, S. Takatsuka *et al.*, "Increased sialyl Lewis A expression and fucosyltransferase activity with acquisition of a high metastatic capacity in a colon cancer cell line," *Br J Cancer,* vol. 76, no. 5, pp. 582-7, 1997.

[50]   M. Thurin, and T. Kieber-Emmons, "SA-Lea and tumor metastasis: the old prediction and recent findings," *Hybrid Hybridomics,* vol. 21, no. 2, pp. 111-6, Apr, 2002.

[51]   K. Tei, N. Kawakami-Kimura, O. Taguchi *et al.*, "Roles of cell adhesion molecules in tumor angiogenesis induced by cotransplantation of cancer and endothelial cells to nude rats," *Cancer Res,* vol. 62, no. 21, pp. 6289-96, Nov 1, 2002.

[52]   S. Mathieu, R. Gerolami, J. Luis *et al.*, "Introducing alpha(1,2)-linked fucose into hepatocarinoma cells inhibits vasculogenesis and tumor growth," *Int J Cancer,* vol. 121, no. 8, pp. 1680-9, Oct 15, 2007.

[53]   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.

[54]   C. Ohyama, S. Tsuboi, and M. Fukuda, "Dual roles of sialyl Lewis X oligosaccharides in tumor metastasis and rejection by natural killer cells," *EMBO J,* vol. 18, no. 6, pp. 1516-25, Mar 15, 1999.

[55]   C. Ohyama, S. Kanto, K. Kato *et al.*, "Natural killer cells attack tumor cells expressing high levels of sialyl Lewis x oligosaccharides," *Proc Natl Acad Sci U S A,* vol. 99, no. 21, pp. 13789-94, Oct 15, 2002.

[56]   F. Dall'Olio, N. Malagolini, M. Trinchera *et al.*, "Mechanisms of cancer-associated glycosylation changes," *Front Biosci (Landmark Ed),* vol. 17, pp. 670-99, 2012.

[57]   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.

[58]   P. V. Beum, J. Singh, M. Burdick *et al.*, "Expression of core 2 beta-1,6-N-acetylglucosaminyltransferase in a human pancreatic cancer cell line results in altered expression of MUC1 tumor-associated epitopes," *J Biol Chem,* vol. 274, no. 35, pp. 24641-8, Aug 27, 1999.

[59]   N. Matsuura, T. Narita, N. Hiraiwa *et al.*, "Gene expression of fucosyl- and sialyl-transferases which synthesize sialyl Lewisx, the carbohydrate ligands for E-selectin, in human breast cancer," *Int J Oncol,* vol. 12, no. 5, pp. 1157-64, May, 1998.

[60]   A. Togayachi, T. Kudo, Y. Ikehara *et al.*, "Up-regulation of Lewis enzyme (Fuc-TIII) and plasma-type alpha1,3fucosyltransferase (Fuc-TVI) expression determines the augmented expression of sialyl Lewis x antigen in non-small cell lung cancer," *Int J Cancer,* vol. 83, no. 1, pp. 70-9, Sep 24, 1999.

[61]   M. Perez-Garay, B. Arteta, E. Llop *et al.*, "alpha2,3-sialyltransferase ST3Gal IV promotes migration and metastasis in pancreatic adenocarcinoma cells and tends to be highly expressed in pancreatic adenocarcinoma tissues," *Int J Biochem Cell Biol,* vol. 45, no. 8, pp. 1748-57, Aug, 2013.

[62]   T. Dohi, M. Hashiguchi, S. Yamamoto *et al.*, "Fucosyltransferase-producing sialyl Le(a) and sialyl Le(x) carbohydrate antigen in benign and malignant gastrointestinal mucosa," *Cancer,* vol. 73, no. 6, pp. 1552-61, Mar 15, 1994.

[63]   H. Ito, N. Hiraiwa, M. Sawada-Kasugai *et al.*, "Altered mRNA expression of specific molecular species of fucosyl- and sialyl-transferases in human colorectal cancer tissues," *Int J Cancer,* vol. 71, no. 4, pp. 556-64, May 16, 1997.

[64]   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.

[65]   E. H. Holmes, S. Hakomori, and G. K. Ostrander, "Synthesis of type 1 and 2 lacto series glycolipid antigens in human colonic adenocarcinoma and derived cell lines is due to activation of a normally unexpressed beta 1----3N-acetylglucosaminyltransferase," *J Biol Chem,* vol. 262, no. 32, pp. 15649-58, Nov 15, 1987.

[66]   N. T. Marcos, A. Magalhaes, B. Ferreira *et al.*, "Helicobacter pylori induces beta3GnT5 in human gastric cell lines, modulating expression of the SabA ligand sialyl-Lewis x," *J Clin Invest,* vol. 118, no. 6, pp. 2325-36, Jun, 2008.

[67]   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.

[68]   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.

[69]   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.

[70]   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.

[71]   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.

[72]   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.

[73]   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.

[74]   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.

[75]   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.

[76]   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.

[77]   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.

[78]   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.

[79]   L. Bry, P. G. Falk, and J. L. Gordon, "Genetic engineering of carbohydrate biosynthetic pathways in transgenic mice demonstrates cell cycle-associated regulation of glycoconjugate production in small intestinal epithelial cells," *Proc Natl Acad Sci U S A,* vol. 93, no. 3, pp. 1161-6, Feb 6, 1996.

[80]   D. Zhou, T. R. Henion, F. B. Jungalwala *et al.*, "The beta 1,3-galactosyltransferase beta 3GalT-V is a stage-specific embryonic antigen-3

(SSEA-3) synthase," *J Biol Chem,* vol. 275, no. 30, pp. 22631-4, Jul 28, 2000.

[81]   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.

[82]   C. R. Beck, J. L. Garcia-Perez, R. M. Badge *et al.*, "LINE-1 elements in structural variation and disease," *Annu Rev Genomics Hum Genet,* vol. 12, pp. 187-215, Sep 22, 2011.

[83]   B. Chenais, "Transposable elements and human cancer: a causal relationship?," *Biochim Biophys Acta,* vol. 1835, no. 1, pp. 28-35, Jan, 2013.

[84]   E. J. Pritham, "Transposable elements and factors influencing their success in eukaryotes," *J Hered,* vol. 100, no. 5, pp. 648-55, Sep-Oct, 2009.

[85]   B. G. Thornburg, V. Gotea, and W. Makalowski, "Transposable elements as a significant source of transcription regulating signals," *Gene,* vol. 365, pp. 104-10, Jan 3, 2006.

[86]   P. Jern, and J. M. Coffin, "Effects of retroviruses on host genome function," *Annu Rev Genet,* vol. 42, pp. 709-32, 2008.

[87]   P. Medstrand, L. N. van de Lagemaat, C. A. Dunn *et al.*, "Impact of transposable elements on the evolution of mammalian gene regulation," *Cytogenet Genome Res,* vol. 110, no. 1-4, pp. 342-52, 2005.

[88]   W. Makalowski, "SINEs as a genomic scrap yard: an essay on genomic evolution," *The Impact of Short Interspersed Elements (SINEs) on the Hpst Genome*, R. J. Maraia, ed., 1995, pp. 81–104.

[89]   J. Brosius, and S. J. Gould, "On "genomenclature": a comprehensive (and respectful) taxonomy for pseudogenes and other "junk DNA"," *Proc Natl Acad Sci U S A,* vol. 89, no. 22, pp. 10706-10, Nov 15, 1992.

[90]   X. S. Wang, Z. Zhang, H. C. Wang *et al.*, "Rapid identification of UCA1 as a very sensitive and specific unique marker for human bladder carcinoma," *Clin Cancer Res,* vol. 12, no. 16, pp. 4851-8, Aug 15, 2006.

[91]   Y. Zhang, I. A. Maksakova, L. Gagnier *et al.*, "Genome-wide assessments reveal extremely high levels of polymorphism of two active families of mouse endogenous retroviral elements," *PLoS Genet,* vol. 4, no. 2, pp. e1000007, Feb, 2008.

[92]   E. D. Sverdlov, "Perpetually mobile footprints of ancient infections in human genome," *FEBS Lett,* vol. 428, no. 1-2, pp. 1-6, May 22, 1998.

[93]   S. G. Kurdyukov, Y. B. Lebedev, Artamonova, II *et al.*, "Full-sized HERV-K (HML-2) human endogenous retroviral LTR sequences on human chromosome 21: map locations and evolutionary history," *Gene,* vol. 273, no. 1, pp. 51-61, Jul 25, 2001.

[94]   H. L. Yu, Z. K. Zhao, and F. Zhu, "The role of human endogenous retroviral long terminal repeat sequences in human cancer (Review)," *Int J Mol Med,* vol. 32, no. 4, pp. 755-62, Oct, 2013.

[95]   A. Feuchter, and D. Mager, "Functional heterogeneity of a large family of human LTR-like promoters and enhancers," *Nucleic Acids Res,* vol. 18, no. 5, pp. 1261-70, Mar 11, 1990.

[96]   P. Medstrand, J. R. Landry, and D. L. Mager, "Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans," *J Biol Chem,* vol. 276, no. 3, pp. 1896-903, Jan 19, 2001.

[97]   D. L. Mager, D. G. Hunter, M. Schertzer *et al.*, "Endogenous retroviruses provide the primary polyadenylation signal for two new human genes (HHLA2 and HHLA3)," *Genomics,* vol. 59, no. 3, pp. 255-63, Aug 1, 1999.

[98]   C. Baust, W. Seifarth, H. Germaier *et al.*, "HERV-K-T47D-Related long terminal repeats mediate polyadenylation of cellular transcripts," *Genomics,* vol. 66, no. 1, pp. 98-103, May 15, 2000.

[99]   N. Kato, S. Pfeifer-Ohlsson, M. Kato *et al.*, "Tissue-specific expression of human provirus ERV3 mRNA in human placenta: two of the three ERV3 mRNAs contain human cellular sequences," *J Virol,* vol. 61, no. 7, pp. 2182-91, Jul, 1987.

[100]   C. J. Cohen, W. M. Lock, and D. L. Mager, "Endogenous retroviral LTRs as promoters for human genes: a critical assessment," *Gene,* vol. 448, no. 2, pp. 105-14, Dec 15, 2009.

[101]   G. Bourque, B. Leong, V. B. Vega *et al.*, "Evolution of the mammalian transcription factor binding repertoire via transposable elements," *Genome Res,* vol. 18, no. 11, pp. 1752-62, Nov, 2008.

[102]   C. Feschotte, "Transposable elements and the evolution of regulatory networks," *Nat Rev Genet,* vol. 9, no. 5, pp. 397-405, May, 2008.

[103]   W. Seifarth, O. Frank, U. Zeilfelder *et al.*, "Comprehensive analysis of human endogenous retrovirus transcriptional activity in human tissues with a retrovirus-specific microarray," *J Virol,* vol. 79, no. 1, pp. 341-52, Jan, 2005.

[104]   D. Taruscio, and A. Mantovani, "Factors regulating endogenous retroviral sequences in human and mouse," *Cytogenet Genome Res,* vol. 105, no. 2-4, pp. 351-62, 2004.

[105]   D. Baek, C. Davis, B. Ewing *et al.*, "Characterization and predictive discovery of evolutionarily conserved mammalian alternative promoters," *Genome Res,* vol. 17, no. 2, pp. 145-55, Feb, 2007.

[106]   J. Takeda, Y. Suzuki, M. Nakao *et al.*, "H-DBAS: alternative splicing database of completely sequenced and manually annotated full-length cDNAs based on H-Invitational," *Nucleic Acids Res,* vol. 35, no. Database issue, pp. D104-9, Jan, 2007.

[107]   S. J. Cooper, N. D. Trinklein, E. D. Anton *et al.*, "Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome," *Genome Res,* vol. 16, no. 1, pp. 1-10, Jan, 2006.

[108]   R. V. Davuluri, Y. Suzuki, S. Sugano *et al.*, "The functional consequences of alternative promoter use in mammalian genomes," *Trends Genet,* vol. 24, no. 4, pp. 167-77, Apr, 2008.

[109]   J. R. Landry, and D. L. Mager, "Functional analysis of the endogenous retroviral promoter of the human endothelin B receptor gene," *J Virol,* vol. 77, no. 13, pp. 7459-66, Jul, 2003.

[110]   K. Puomila, O. Simell, K. Huoponen *et al.*, "Two alternative promoters regulate the expression of lysinuric protein intolerance gene SLC7A7," *Mol Genet Metab,* vol. 90, no. 3, pp. 298-306, Mar, 2007.

[111]   C. Swan, S. A. Richards, N. P. Duroudier *et al.*, "Alternative promoter use and splice variation in the human histamine H1 receptor gene," *Am J Respir Cell Mol Biol,* vol. 35, no. 1, pp. 118-26, Jul, 2006.

[112]   M. V. Tzvetkov, C. Meineke, E. Oetjen *et al.*, "Tissue-specific alternative promoters of the serotonin receptor gene HTR3B in human brain and intestine," *Gene,* vol. 386, no. 1-2, pp. 52-62, Jan 15, 2007.

[113]   L. Ding, J. C. Mychaleckyj, and A. N. Hegde, "Full length cloning and expression analysis of splice variants of regulator of G-protein signaling RGS4 in human and murine brain," *Gene,* vol. 401, no. 1-2, pp. 46-60, Oct 15, 2007.

[114]   E. A. Shephard, P. Chandan, M. Stevanovic-Walker *et al.*, "Alternative promoters and repetitive DNA elements define the species-dependent tissue-specific expression of the FMO1 genes of human and mouse," *Biochem J,* vol. 406, no. 3, pp. 491-9, Sep 15, 2007.

[115]   A. Pozner, J. Lotem, C. Xiao *et al.*, "Developmentally regulated promoter-switch transcriptionally controls Runx1 function during embryonic hematopoiesis," *BMC Dev Biol,* vol. 7, pp. 84, 2007.

[116]   Y. F. Cheung, Z. Kan, P. Garrett-Engele *et al.*, "PDE4B5, a novel, super-short, brain-specific cAMP phosphodiesterase-4 variant whose isoform-specifying N-terminal region is identical to that of cAMP phosphodiesterase-4D6 (PDE4D6)," *J Pharmacol Exp Ther,* vol. 322, no. 2, pp. 600-9, Aug, 2007.

[117]   E. Logette, A. Wotawa, S. Solier *et al.*, "The human caspase-2 gene: alternative promoters, pre-mRNA splicing and AUG usage direct isoform-specific expression," *Oncogene,* vol. 22, no. 6, pp. 935-46, Feb 13, 2003.

[118]   E. Meshorer, D. Toiber, D. Zurel *et al.*, "Combinatorial complexity of 5' alternative acetylcholinesterase transcripts and protein products," *J Biol Chem,* vol. 279, no. 28, pp. 29740-51, Jul 9, 2004.

[119]   S. Nozell, and X. Chen, "p21B, a variant of p21(Waf1/Cip1), is induced by the p53 family," *Oncogene,* vol. 21, no. 8, pp. 1285-94, Feb 14, 2002.

[120]   D. E. Quelle, F. Zindy, R. A. Ashmun *et al.*, "Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest," *Cell,* vol. 83, no. 6, pp. 993-1000, Dec 15, 1995.

[121]   P. Tripal, M. Bauer, E. Naschberger *et al.*, "Unique features of different members of the human guanylate-binding protein family," *J Interferon Cytokine Res,* vol. 27, no. 1, pp. 44-52, Jan, 2007.

[122]   M. A. Olszewski, J. Gray, and D. J. Vestal, "In silico genomic analysis of the human and murine guanylate-binding protein (GBP) gene clusters," *J Interferon Cytokine Res,* vol. 26, no. 5, pp. 328-52, May, 2006.

[123]   W. J. Kent, C. W. Sugnet, T. S. Furey *et al.*, "The human genome browser at UCSC," *Genome Res,* vol. 12, no. 6, pp. 996-1006, Jun, 2002.

[124]   T. Kojima, M. Oh-eda, K. Hattori *et al.*, "Molecular cloning and expression of megakaryocyte potentiating factor cDNA," *J Biol Chem,* vol. 270, no. 37, pp. 21984-90, Sep 15, 1995.

[125]   D. Urwin, and R. A. Lake, "Structure of the Mesothelin/MPF gene and characterization of its promoter," *Mol Cell Biol Res Commun,* vol. 3, no. 1, pp. 26-32, Jan, 2000.

[126]   V. Blikstad, F. Benachenhou, G. O. Sperber *et al.*, "Evolution of human endogenous retroviral sequences: a conceptual account," *Cell Mol Life Sci,* vol. 65, no. 21, pp. 3348-65, Nov, 2008.

[127]   C. Fuke, M. Shimabukuro, A. Petronis *et al.*, "Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study," *Ann Hum Genet,* vol. 68, no. Pt 3, pp. 196-204, May, 2004.

[128]   D. Reiss, Y. Zhang, and D. L. Mager, "Widely variable endogenous retroviral methylation levels in human placenta," *Nucleic Acids Res,* vol. 35, no. 14, pp. 4743-54, 2007.

[129]   S. Noorali, I. C. Rotar, C. Lewis *et al.*, "Role of HERV-W syncytin-1 in placentation and maintenance of human pregnancy," *Appl Immunohistochem Mol Morphol,* vol. 17, no. 4, pp. 319-28, Jul, 2009.

[130]   R. Hassan, and M. Ho, "Mesothelin targeted cancer immunotherapy," *Eur J Cancer,* vol. 44, no. 1, pp. 46-53, Jan, 2008.

[131]   S. M. Rawn, and J. C. Cross, "The evolution, regulation, and function of placenta-specific genes," *Annu Rev Cell Dev Biol,* vol. 24, pp. 159-81, 2008.

[132]   I. Bieche, A. Laurent, I. Laurendeau *et al.*, "Placenta-specific INSL4 expression is mediated by a human endogenous retrovirus element," *Biol Reprod,* vol. 68, no. 4, pp. 1422-9, Apr, 2003.

[133]   R. S. Hansen, and S. M. Gartler, "5-Azacytidine-induced reactivation of the human X chromosome-linked PGK1 gene is associated with a large region of cytosine demethylation in the 5' CpG island," *Proc Natl Acad Sci U S A,* vol. 87, no. 11, pp. 4174-8, Jun, 1990.

[134]   A. Merlo, J. G. Herman, L. Mao *et al.*, "5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers," *Nat Med,* vol. 1, no. 7, pp. 686-92, Jul, 1995.

[135]   M. Esteller, "Epigenetics in cancer," *N Engl J Med,* vol. 358, no. 11, pp. 1148-59, Mar 13, 2008.

[136]   E. Ballestar, and A. P. Wolffe, "Methyl-CpG-binding proteins. Targeting specific gene repression," *Eur J Biochem,* vol. 268, no. 1, pp. 1-6, Jan, 2001.

[137]   A. Bird, "DNA methylation patterns and epigenetic memory," *Genes Dev,* vol. 16, no. 1, pp. 6-21, Jan 1, 2002.

[138]   D. N. Mancini, S. M. Singh, T. K. Archer *et al.*, "Site-specific DNA methylation in the neurofibromatosis (NF1) promoter interferes with binding of CREB and SP1 transcription factors," *Oncogene,* vol. 18, no. 28, pp. 4108-19, Jul 15, 1999.

[139]   A. T. Hark, C. J. Schoenherr, D. J. Katz *et al.*, "CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus," *Nature,* vol. 405, no. 6785, pp. 486-9, May 25, 2000.

[140]   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.

[141]   M. Ehrlich, "DNA methylation in cancer: too much, but also too little," *Oncogene,* vol. 21, no. 35, pp. 5400-13, Aug 12, 2002.

[142] M. J. Hoffmann, and W. A. Schulz, "Causes and consequences of DNA hypomethylation in human cancer," *Biochem Cell Biol,* vol. 83, no. 3, pp. 296-321, Jun, 2005.

[143] T. A. Rauch, X. Zhong, X. Wu *et al.*, "High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer," *Proc Natl Acad Sci U S A,* vol. 105, no. 1, pp. 252-7, Jan 8, 2008.

[144] S. E. Goelz, B. Vogelstein, S. R. Hamilton *et al.*, "Hypomethylation of DNA from benign and malignant human colon neoplasms," *Science,* vol. 228, no. 4696, pp. 187-90, Apr 12, 1985.

[145] C. Grunau, M. E. Brun, I. Rivals *et al.*, "BAGE hypomethylation, a new epigenetic biomarker for colon cancer detection," *Cancer Epidemiol Biomarkers Prev,* vol. 17, no. 6, pp. 1374-9, Jun, 2008.

[146] P. Novak, T. Jensen, M. M. Oshiro *et al.*, "Agglomerative epigenetic aberrations are a common event in human breast cancer," *Cancer Res,* vol. 68, no. 20, pp. 8616-25, Oct 15, 2008.

[147] 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.

[148] 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.

[149] A. Portela, and M. Esteller, "Epigenetic modifications and human disease," *Nat Biotechnol,* vol. 28, no. 10, pp. 1057-68, Oct, 2010.

[150] R. Lister, M. Pelizzola, R. H. Dowen *et al.*, "Human DNA methylomes at base resolution show widespread epigenomic differences," *Nature,* vol. 462, no. 7271, pp. 315-22, Nov 19, 2009.

[151] L. Laurent, E. Wong, G. Li *et al.*, "Dynamic changes in the human methylome during differentiation," *Genome Res,* vol. 20, no. 3, pp. 320-31, Mar, 2010.

[152] 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.

[153] 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.

[154] G. C. Hon, R. D. Hawkins, O. L. Caballero *et al.*, "Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer," *Genome Res,* vol. 22, no. 2, pp. 246-58, Feb, 2012.

[155] T. A. Rauch, X. Wu, X. Zhong *et al.*, "A human B cell methylome at 100-base pair resolution," *Proc Natl Acad Sci U S A,* vol. 106, no. 3, pp. 671-8, Jan 20, 2009.

[156] 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.

[157] M. Kulis, A. C. Queiros, R. Beekman *et al.*, "Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer," *Biochim Biophys Acta*, Aug 9, 2013.

[158] M. Kulis, S. Heath, M. Bibikova *et al.*, "Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia," *Nat Genet,* vol. 44, no. 11, pp. 1236-42, Nov, 2012.

[159] S. Deng, G. A. Calin, C. M. Croce *et al.*, "Mechanisms of microRNA deregulation in human cancer," *Cell Cycle,* vol. 7, no. 17, pp. 2643-6, Sep 1, 2008.

[160] M. V. Iorio, C. Piovan, and C. M. Croce, "Interplay between microRNAs and the epigenetic machinery: an intricate network," *Biochim Biophys Acta,* vol. 1799, no. 10-12, pp. 694-701, Oct-Dec, 2010.

[161] C. P. Ponting, P. L. Oliver, and W. Reik, "Evolution and functions of long noncoding RNAs," *Cell,* vol. 136, no. 4, pp. 629-41, Feb 20, 2009.

[162] S. J. Brown, P. Stoilov, and Y. Xing, "Chromatin and epigenetic regulation of pre-mRNA processing," *Hum Mol Genet,* vol. 21, no. R1, pp. R90-6, Oct 15, 2012.

[163] L. A. Chasin, "Searching for splicing motifs," *Adv Exp Med Biol,* vol. 623, pp. 85-106, 2007.

[164] S. Schwartz, E. Meshorer, and G. Ast, "Chromatin organization marks exon-intron structure," *Nat Struct Mol Biol,* vol. 16, no. 9, pp. 990-5, Sep, 2009.

[165] R. F. Luco, Q. Pan, K. Tominaga *et al.*, "Regulation of alternative splicing by histone modifications," *Science,* vol. 327, no. 5968, pp. 996-1000, Feb 19, 2010.

[166] E. A. Hoivik, T. E. Bjanesoy, O. Mai *et al.*, "DNA methylation of intronic enhancers directs tissue-specific expression of steroidogenic factor 1/adrenal 4 binding protein (SF-1/Ad4BP)," *Endocrinology,* vol. 152, no. 5, pp. 2100-12, May, 2011.

[167] D. Aran, S. Sabato, and A. Hellman, "DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes," *Genome Biol,* vol. 14, no. 3, pp. R21, Mar 12, 2013.

[168] T. K. Kim, M. Hemberg, J. M. Gray *et al.*, "Widespread transcription at neuronal activity-regulated enhancers," *Nature,* vol. 465, no. 7295, pp. 182-7, May 13, 2010.

[169] C. A. Melo, J. Drost, P. J. Wijchers *et al.*, "eRNAs are required for p53-dependent enhancer activity and gene transcription," *Mol Cell,* vol. 49, no. 3, pp. 524-35, Feb 7, 2013.

[170] D. B. Mendel, and G. R. Crabtree, "HNF-1, a member of a novel class of dimerizing homeodomain proteins," *J Biol Chem,* vol. 266, no. 2, pp. 677-80, Jan 15, 1991.

[171] F. Tronche, and M. Yaniv, "HNF1, a homeoprotein member of the hepatic transcription regulatory network," *Bioessays,* vol. 14, no. 9, pp. 579-87, Sep, 1992.

[172] M. Frain, G. Swart, P. Monaci *et al.*, "The liver-specific transcription factor LF-B1 contains a highly diverged homeobox DNA binding domain," *Cell,* vol. 59, no. 1, pp. 145-57, Oct 6, 1989.

[173]   I. Bach, and M. Yaniv, "More potent transcriptional activators or a transdominant inhibitor of the HNF1 homeoprotein family are generated by alternative RNA processing," *EMBO J,* vol. 12, no. 11, pp. 4229-42, Nov, 1993.

[174]   S. Baumhueter, D. B. Mendel, P. B. Conley *et al.*, "HNF-1 shares three sequence motifs with the POU domain proteins and is identical to LF-B1 and APF," *Genes Dev,* vol. 4, no. 3, pp. 372-9, Mar, 1990.

[175]   I. Bach, M. G. Mattei, S. Cereghini *et al.*, "Two members of an HNF1 homeoprotein family are expressed in human liver," *Nucleic Acids Res,* vol. 19, no. 13, pp. 3553-9, Jul 11, 1991.

[176]   M. Pontoglio, "Hepatocyte nuclear factor 1, a transcription factor at the crossroads of glucose homeostasis," *J Am Soc Nephrol,* vol. 11 Suppl 16, pp. S140-3, Nov, 2000.

[177]   K. Yamagata, N. Oda, P. J. Kaisaki *et al.*, "Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3)," *Nature,* vol. 384, no. 6608, pp. 455-8, Dec 5, 1996.

[178]   G. D. Wu, L. Chen, K. Forslund *et al.*, "Hepatocyte nuclear factor-1 alpha (HNF-1 alpha) and HNF-1 beta regulate transcription via two elements in an intestine-specific promoter," *J Biol Chem,* vol. 269, no. 25, pp. 17080-5, Jun 24, 1994.

[179]   P. G. Traber, L. Yu, G. D. Wu *et al.*, "Sucrase-isomaltase gene expression along crypt-villus axis of human small intestine is regulated at level of mRNA abundance," *Am J Physiol,* vol. 262, no. 1 Pt 1, pp. G123-30, Jan, 1992.

[180]   V. Lorenzsonn, H. Korsmo, and W. A. Olsen, "Localization of sucrase-isomaltase in the rat enterocyte," *Gastroenterology,* vol. 92, no. 1, pp. 98-105, Jan, 1987.

[181]   E. Suh, L. Chen, J. Taylor *et al.*, "A homeodomain protein related to caudal regulates intestine-specific gene transcription," *Mol Cell Biol,* vol. 14, no. 11, pp. 7340-51, Nov, 1994.

[182]   F. Boudreau, E. H. Rings, H. M. van Wering *et al.*, "Hepatocyte nuclear factor-1 alpha, GATA-4, and caudal related homeodomain protein Cdx2 interact functionally to modulate intestinal gene transcription. Implication for the developmental regulation of the sucrase-isomaltase gene," *J Biol Chem,* vol. 277, no. 35, pp. 31909-17, Aug 30, 2002.

[183]   M. S. Keller, T. Ezaki, R. J. Guo *et al.*, "Cdx1 or Cdx2 expression activates E-cadherin-mediated cell-cell adhesion and compaction in human COLO 205 cells," *Am J Physiol Gastrointest Liver Physiol,* vol. 287, no. 1, pp. G104-14, Jul, 2004.

[184]   M. Trinchera, N. Malagolini, M. Chiricolo *et al.*, "The biosynthesis of the selectin-ligand sialyl Lewis x in colorectal cancer tissues is regulated by fucosyltransferase VI and can be inhibited by an RNA interference-based approach," *Int J Biochem Cell Biol,* vol. 43, no. 1, pp. 130-9, Jan, 2011.

[185]   E. Pilozzi, M. R. Onelli, V. Ziparo *et al.*, "CDX1 expression is reduced in colorectal carcinoma and is associated with promoter hypermethylation," *J Pathol,* vol. 204, no. 3, pp. 289-95, Nov, 2004.

[186] C. Wicking, L. A. Simms, T. Evans *et al.*, "CDX2, a human homologue of Drosophila caudal, is mutated in both alleles in a replication error positive colorectal cancer," *Oncogene,* vol. 17, no. 5, pp. 657-9, Aug 6, 1998.

[187] 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.

[188] G. Gilliland, S. Perrin, K. Blanchard *et al.*, "Analysis of cytokine mRNA and DNA: detection and quantitation by competitive polymerase chain reaction," *Proc Natl Acad Sci U S A,* vol. 87, no. 7, pp. 2725-9, Apr, 1990.

[189] C. Mitchelmore, J. T. Troelsen, N. Spodsberg *et al.*, "Interaction between the homeodomain proteins Cdx2 and HNF1alpha mediates expression of the lactase-phlorizin hydrolase gene," *Biochem J,* vol. 346 Pt 2, pp. 529-35, Mar 1, 2000.

[190] T. Sakaguchi, X. Gu, H. M. Golden *et al.*, "Cloning of the human claudin-2 5'-flanking region revealed a TATA-less promoter with conserved binding sites in mouse and human for caudal-related homeodomain proteins and hepatocyte nuclear factor-1alpha," *J Biol Chem,* vol. 277, no. 24, pp. 21361-70, Jun 14, 2002.

[191] A. D'Angelo, O. Bluteau, M. A. Garcia-Gonzalez *et al.*, "Hepatocyte nuclear factor 1alpha and beta control terminal differentiation and cell fate commitment in the gut epithelium," *Development,* vol. 137, no. 9, pp. 1573-82, May, 2010.

[192] M. A. Gama-Sosa, V. A. Slagel, R. W. Trewyn *et al.*, "The 5-methylcytosine content of DNA from human tumors," *Nucleic Acids Res,* vol. 11, no. 19, pp. 6883-94, Oct 11, 1983.

[193] M. Ehrlich, "DNA hypomethylation in cancer cells," *Epigenomics,* vol. 1, no. 2, pp. 239-59, Dec, 2009.

[194] R. Kikuchi, S. Yagi, H. Kusuhara *et al.*, "Genome-wide analysis of epigenetic signatures for kidney-specific transporters," *Kidney Int,* vol. 78, no. 6, pp. 569-77, Sep, 2010.

[195] P. Kheradpour, J. Ernst, A. Melnikov *et al.*, "Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay," *Genome Res,* vol. 23, no. 5, pp. 800-11, May, 2013.

[196] K. D. Robertson, "DNA methylation, methyltransferases, and cancer," *Oncogene,* vol. 20, no. 24, pp. 3139-55, May 28, 2001.

[197] R. Rebollo, M. T. Romanish, and D. L. Mager, "Transposable elements: an abundant and natural source of regulatory sequences for host genes," *Annu Rev Genet,* vol. 46, pp. 21-42, 2012.

SCIENTIFIC PRODUCTS

Publications on the thesis topic:

✓ Transcriptional control of the *B3GALT5* gene by a retroviral promoter and methylation of distant regulatory elements. Zulueta A, Caretti A, Signorelli P, Dall'olio F, Trinchera M. FASEB J. 2013 Oct 15. [Epub ahead of print]

✓ DNA methylation and histone modifications modulate the β1,3 galactosyltransferase β3Gal-T5 native promoter in cancer cells. Caretti A, Sirchia SM, Tabano S, Zulueta A, Dall'Olio F, Trinchera M. Int J Biochem Cell Biol. 2012 Jan;44(1):84-90.

Presentations on international events:

✓ 2nd Latin-American Congress of Glicobiology, Mexico, July 31 - August 2, 2013. Oral presentation: Transcriptional regulation of β1,3 galactosyltransferase β3Gal-T5: Role of hepatocyte nuclear factor HNF1.

ACKNOWLEDGMENTS

Un dottorato per chiunque è il culmine di un lungo percorso segnato da diversi sacrifici ed sforzi. Decidersi a farlo a 33 anni, avendo già un figlio e venire in Italia ad inseguire questo sogno fa sembrare il sogno una follia. Una strada che decisi di intraprendere 3 anni fa, piena di dubbi e paure. Arrivai in una giornata d'inverno fredda pero soleggiata, in questo lab pieno di gente che dal primo momento mi fecce sentire come a casa. La mia gratitudine va a tutto il personale del lab per il calore con che sono stata accolta. Ed è per loro che mi sono permessa di scrivere queste righe nel mio patetico e povero italiano.

Un ringraziamento particolare va ai miei supervisori, al mio tutore il Prof. Riccardo Ghidoni, e al Dr. Marco Trinchera. Al Prof. Ghidoni, che nonostante i multipli impegni, ha sempre trovato il tempo per ascoltarmi e consigliarmi, grazie per il supporto, e la saggezza che hanno un valore inestimabile per me. A Marco, per indirizzare il mio progetto di tesi, per il prezioso insegnamento in tutte le fasi di questa tesi, per il suo ottimismo e la sua acutezza.

Ringrazio inoltre ad Anna, per il suo contributo scientifico fondamentale pero soprattutto, per esserci sempre e trovare il consiglio giusto ed equilibrato. Vorrei ringraziare tanto Nadia, per la sua serenità che è da imparare, Paola B. per la preoccupazione costante e le parole di incoraggiamento, Paola S., Rita e il Prof. Samaja per il supporto.

Infine, ringrazio di cuore quelle che hanno vissuto più di vicino le mie paranoie e disavventure, le compagne di tavolo e di viaggio, quelle che non si lamentano mai di sentirmi tutti i giorni con un racconto diverso di Cuba o di Diego, si mostrano persino interessate e restano lì ad ascoltarmi senza lamentarsi del mio italiano disastroso o della mia irrefrenabile logorrea post pranzo…. Bimbe, Grazie infinite!! E-scusate, eh.

Un pensamiento va a mi tierra Cuba, a la que espero un día volver. Finalmente, quisiera agradecer a mi tio y el resto de la familia, amigos y ex-colegas que son parte de este logro y porque aún desde la distancia logran siempre hacerme llegar su calor. Primero que todo, a mis abuelos por el amor infinito y la preocupación constante por hacer de mi una mejor persona. A mi padre Rodo por todo el cariño y la complicidad. Agradezco muchísimo a Rafael por su amor, paciencia y apoyo incondicional. Tengo muchas razones por las que agradecer a mi madre pero la principal es el ejemplo de genialidad y tenacidad que es para mi, no digo mas porque sería trivial.

Y ahora sí finalmente, un GRACIAS inmenso a mi príncipe enano por su amor ilimitado, ternura y alegría contagiosa en los días más difíciles. Por dejarse arrastrar en este sueño-locura... Lo sé que ha sido duro para ti, creeme que lo sé.

Alla fine,.... se la rifarei, questa esperienza? Ad occhi chiusi!

GRAZIE!!

..You may say I'm a dreamer,
but I'm not the only one...