Transfusion Medicine and Hemotherapy

Review Article

Transfus Med Hemother 2024;51:225–235
DOI: 10.1159/000538863

Received: November 27, 2023
Accepted: April 11, 2024
Published online: June 19, 2024

Regulation of the Lewis Blood Group Antigen Expression: A Literature Review Supplemented with Computational Analysis

Martin Wipplinger[a] Sylvia Mink[b, c] Maike Bublitz[a] Christoph Gassner[a]

[a]Institute of Translational Medicine, Private University in the Principality of Liechtenstein, Triesen, Liechtenstein; [b]Central Medical Laboratories, Feldkirch, Austria; [c]Medical-Scientific Faculty, Private University of the Principality of Liechtenstein, Triesen, Liechtenstein

Keywords

Lewis blood group · Transcription · Transcription factors · Gene regulation · Translation

Abstract

Background: The Lewis (Le) blood group system, unlike most other blood groups, is not defined by antigens produced internally to the erythrocytes and their precursors but rather by glycan antigens adsorbed on to the erythrocyte membrane from the plasma. These oligosaccharides are synthesized by the two fucosyltransferases *FUT2* and *FUT3* mainly in epithelial cells of the digestive tract and transferred to the plasma. At their place of synthesis, some Lewis blood group carbohydrate antigen variants also seem to be involved in various gastrointestinal malignancies. However, relatively little is known about the transcriptional regulation of *FUT2* and *FUT3*. ***Summary:*** To address this question, we screened existing literature and additionally used in silico prediction tools to identify novel candidate regulators for *FUT2* and *FUT3* and combine these findings with already known data on their regulation. With this approach, we were able to describe a variety of transcription factors, RNA binding proteins and microRNAs, which increase *FUT2* and *FUT3* transcription and translation upon interaction. ***Key Messages:*** Understanding the regulation of *FUT2* and *FUT3* is crucial to fully understand the blood group system Lewis (ISBT 007 LE) phenotypes, to shed light on the role of the different Lewis antigens in various pathologies, and to identify potential new diagnostic targets for these diseases.

© 2024 The Author(s).
Published by S. Karger AG, Basel

Plain Language Summary

The Lewis (Le) blood group system, in contrast to the majority of blood groups, is not able to synthesize its antigens itself. It depends on the attachment of different oligosaccharides to the erythrocyte membrane, which are adsorbed from the plasma. These glycans are modified by the fucosyltransferases 2 and 3 enzymes (FUT2/3). Beside their role in defining the Lewis blood group, FUT2 and FUT3 are also known to be involved in the susceptibility and progression of various gastrointestinal pathologies, like inflammatory bowel diseases (IBD) or colorectal cancer (CRC). Even though different expression levels of *FUT2* and *FUT3* have been described in these malignancies, relatively little is known about the mechanisms behind their transcriptional regulation. In this review, we aim to shed light on transcription factors (TFs) responsible for *FUT2* and *FUT3* expression as well as on post-transcriptional regulators by the means of RNA binding proteins (RBPs) and microRNAs (miRNAs). To achieve our goal, we combined previous knowledge on *FUT2* and *FUT3* expression regulation with a computational analysis to predict additional novel regulators. On this way, we are able to broaden our knowledge on *FUT2* and *FUT3* expression regulation and consequently might be able to transfer our findings into diagnostics or therapeutics in the future.

© 2024 The Author(s).
Published by S. Karger AG, Basel

karger@karger.com
www.karger.com/tmh

© 2024 The Author(s).
Published by S. Karger AG, Basel

This article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC) (http://www.karger.com/Services/OpenAccessLicense). Usage and distribution for commercial purposes requires written permission.

Karger
OPEN ACCESS

Correspondence to:
Martin Wipplinger, martin.wipplinger@ufl.li

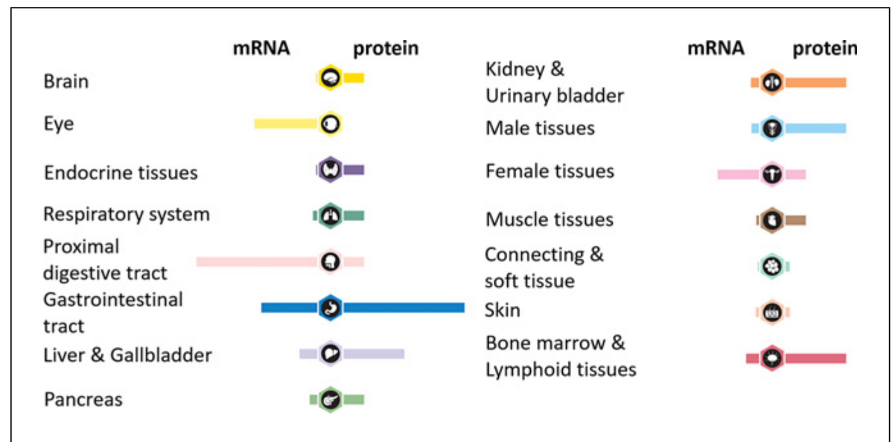

Fig. 1. Expression of FUT3 mRNA and protein in different human tissues. Graphic used from the human protein atlas database (https://www.proteinatlas.org/ENSG00000171124-FUT3/tissue).

Biochemical Structure of Lewis Antigens and Their Synthesis

The Lewis (Le) blood group system, also listed as ISBT 007 by the International Society of Blood Transfusion (ISBT) Working Party for Red Cell Immunogenetics and Blood Group Terminology, differs from most other blood group systems because the glycan antigens of the Lewis blood group system are not synthesized by the red blood cells themselves but adsorbed onto the erythrocyte membrane from the plasma. The antigen source is epithelial cells in digestive, bronchial, mammary, seminal, urinary, or orbital tissue, which secrete glycan antigens either to the outside via exocrine secretion or into the interstitial space from where they can reach the plasma [1].The epithelial cells of the digestive tract seem to be the main origin of the Lewis antigens [2]. This observation is consistent with data from the human protein atlas (https://www.proteinatlas.org/), reporting the highest expression levels of the Le-antigen synthesizing enzyme FUT3 in the gastrointestinal tract (shown in Fig. 1).

The biosynthesis of the Lewis blood group antigens is governed by FUT3, which encodes for the α(1,3/4)-fucosyltransferase Lewis (Le) [3]. The exact antigen product is however depending on the precursor fucosylated by Le. Le^a is produced if individuals are missing FUT2, which encodes for the α(1,2)-fucosyltransferase secretor (Se) [4]. Therefore, a fucose moiety is added at the α1-4 end of the *N*-acetylglucosamine residue. In the presence of both FUT2/Se and FUT3/Le, the final antigen is determined by the active ABO glycosyltransferases. In blood group O individuals, Le^b is synthesized, while ALe^b and BLe^b are made in blood groups A and B, respectively (shown in Fig. 2a). Of note, trace amounts of Le^a will be present in most humans but are difficult to detect phenotypically. There are also rare cases of individuals with simultaneous Le^a and Le^b expression, resulting from a functionally impaired FUT2/Se enzyme being enzymatically less efficient [5].

Above-described Lewis antigens are all synthesized from type 1 precursors. However, also type 2 precursor chains are targeted by FUT2 and FUT3, resulting in the synthesis of Le^x and Le^y, isomeric forms of Le^a and Le^b, respectively. However, there are only minimal amounts of Le^x and Le^y present on the red blood cell surface and are therefore not considered as antigens of the Lewis blood group system [7]. Le^a or Le^x may additionally occur in their sialylated forms (shown in Fig. 2b), which in the case of sialyl-Le^x act as ligands for E-selectin in neutrophils [3]. Of note, processing of type 2 precursors can involve the additional fucosyltransferases FUT4/5/6/7 [4].

Linkage of Lewis Antigen Polymorphism with Disease

The functionality of the FUT2 and FUT3 enzymes and the absence or presence of their antigen products can have severe effects on the onset of diseases, especially in the gastrointestinal tract. The carbohydrate-based Lewis antigens have been shown to have a correlation with the gut microbiome and are considered as influential on the development of inflammatory bowel diseases (IBDs) [5]. Among IBDs, Morbus Crohn's disease (CD) and ulcerative colitis (UC), both characterized by a chronic inflammatory state in the gastric entity, have been linked with loss-of-function mutations in both *FUT2* [5] and *FUT3* [8]. In a CD study from Forni et al., a correlation between an *FUT2* polymorphism and the susceptibility to CD in a Belgian study cohort has been shown. However, this has not been confirmed in an Italian study [9]. There are various known SNPs in *FUT2* that result in an enzymatically inactive "non-secretor" phenotype, which has been shown to increase CD susceptibility [10]. The known non-secretor-related SNP in *FUT2* (*W143X*) has been reported to have differences

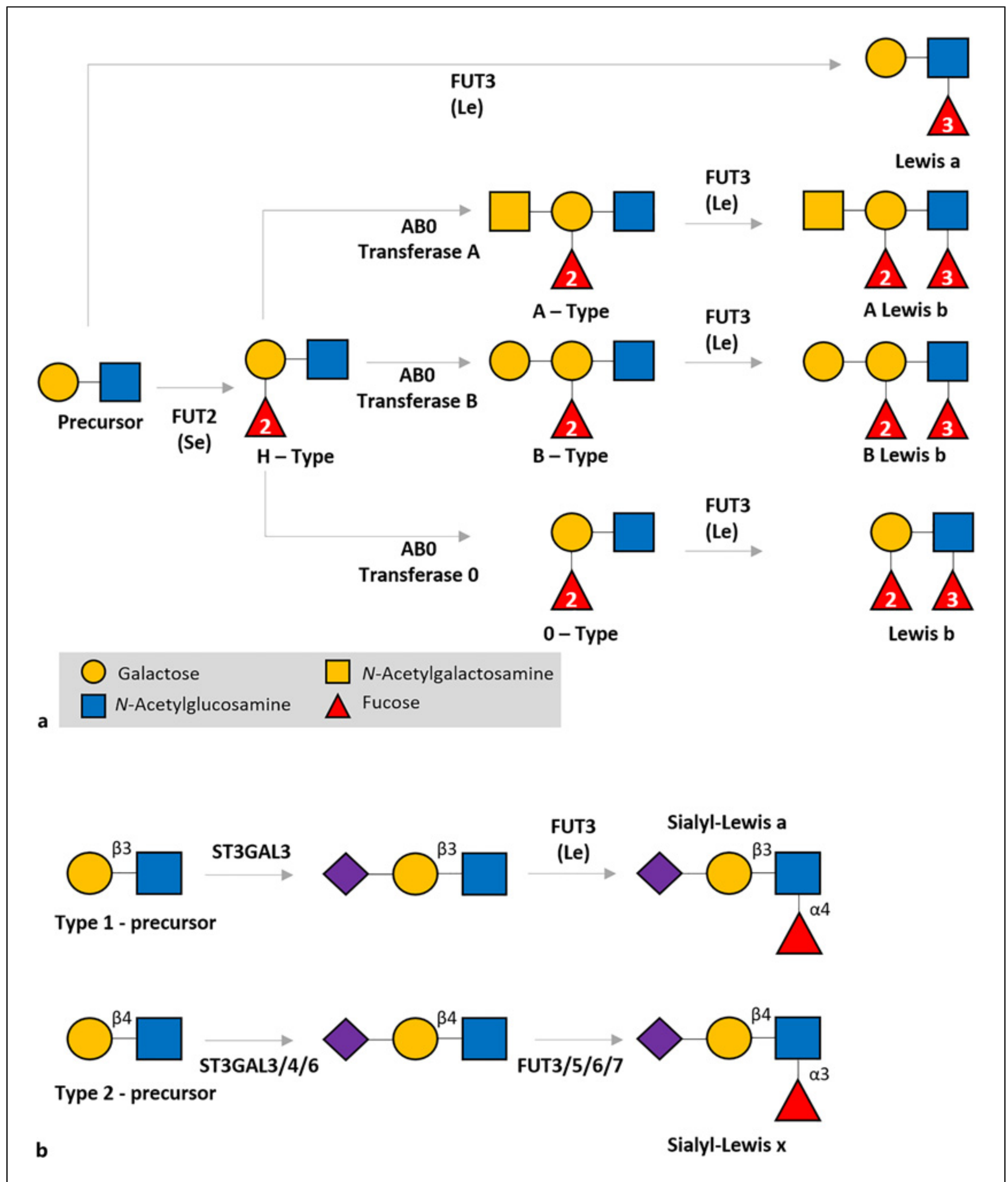

Fig. 2. Lewis antigen synthesis. **a** Biosynthesis pathway of Lewis antigens catalyzed by FUT2, FUT3, and other glycosyltransferases. Figure adapted from [6]. Numbers in fucose indicate if added by FUT2 (2) or FUT3 (3). **b** Synthesis of sialyl-Lewis variants. sLe^a is formed from type 1 precursors by ST3GAL3 and FUT3. Contrary sLex is synthesized from type 2 precursors by ST3GAL3/4/6 and FUT3/5/6/7.

in the host microbiome community compared to healthy individuals. These alterations in microbial composition and diversity are further linked to disease-by genotype influence in CD patients [11]. Regarding *FUT3*, a recent meta-analysis identified the *rs3745635* gene polymorphism being associated with IBD susceptibility, at the same time ruling out other causal polymorphisms [8]. Another study in a Chinese UC population revealed that SNPs in *FUT3* are correlated with vulnerability to UC [12].

Besides IBD, *FUT2* polymorphisms defining the secretor/non-secretor phenotype have been linked with either susceptibility or resistance to gastroenteritis-causing noroviruses and rotaviruses, respectively. The viruses use the Lewis antigens to attach to host cells, consequently infecting them [13].

Scientific reports on the regulation of *FUT2* and *FUT3* transcription in healthy individuals are scarce. In this review, current knowledge of transcriptional regulation of

Table 1. Lewis phenotypes with their corresponding genotypes and frequencies in three different ethnic populations [3]

Antigens			Genotype		Frequencies		
anti-Le^a	anti-Le^b	phenotype	Lewis	secretor	European	African	Asian
+	–	Le(a+b–)	Le	sese	19–22%	19–23%	<1%
–	+	Le(a–b+)	Le	SE	70–72%	52–55%	62%
–	–	Le(a–b–)	lele	sese	4–11%	22–29%	11%
+	+	Le(a+b+)	Le	Se^w	<1%	<1%	27%

According to the Blood Group Allele Tables of the ISBT, SE alleles are represented by *FUT2*01*, se by *FUT2*01N*, and Se^W by *FUT2*01W* alleles in the H blood group system table (ISBT 018), respectively. In analogy, Le stands for ISBT alleles of the *FUT3*01* series and le for ISBT alleles of the *FUT3*01N* series (listed in the table for the Lewis blood group system ISBT 007) (https://www.isbtweb.org/isbt-working-parties/rcibgt/blood-group-allele-tables.html).

FUT2 and *FUT3* expression is supplemented by bioinformatics analyses, aiming to predict new candidate transcription factors (TFs) and micro-RNAs (miRNAs) that interact with the two genes, in order to improve our understanding of transcriptional regulation of the Lewis and secretor fucosyltransferases.

Epigenetic Control of *FUT3* Expression

Regulation of gene expression is essential for cells in order to be able to react to intra- and extracellular signals, differentiate during development, and to mediate cellular activity [14]. Regulation of transcription can occur at various stages of the process, beginning with chromatin remodeling mediating accessibility to the DNA. This is mediated by the methylation status of target gene promotors, which are typically rich in cytosine-guanine dinucleotide (CpG) repeats, so-called CpG islands. A heavily methylated promoter (hypermethylated) is associated with low transcription activity, while low methylation status (hypomethylated) allows transcription initiation [15]. *FUT3* promotor methylation in pigs has been shown to interfere with the accessibility for the two TFs HIF1A and Sp1, consequently influencing produced mRNA levels [16, 17]. The production of Le^a in human gastric cancer cells has been shown to be linked to *FUT3* promotor methylation status [18]. Treatment with demethylation agents resulted in an increase of Lewis enzyme and Le^a levels. These results were confirmed by using a luciferase reporter gene of the *FUT3* promotor in methylated and unmethylated forms, showing that high *FUT3* expression in gastric cells correlates with promoter hypomethylation [18]. This observation was followed up by a recent machine learning approach to identify different methylation patterns as prediction factors regarding risk habits, nodal status, tumor staging, prognosis, and HPV infection in cancer, in which also FUT3 was found to be involved [19] (Table 1).

In silico Analysis of the Transcriptional Regulation of *FUT2* and *FUT3*

An important step of regulation takes place during the actual transcription process via the interaction of target genes with DNA-binding TFs. TFs are able to recognize specific sequence motifs located in enhancer elements and recruit either RNA polymerase II or additional co-factors upon binding, enhancing, or repressing transcription [20].

Since there are relatively little published data on the regulation of *FUT2* and *FUT3* transcription, we decided to complement the existing knowledge with the help of IT-based algorithms. Getting a better understanding of the regulation could be used to predict susceptibility to diseases, such as virus infections or cancer, caused by altered expression of *FUT2* and *FUT3*. To identify possible TFs regulating *FUT2* and *FUT3* expression we used the PROMO TF prediction algorithm (https://alggen.lsi.upc.es). This online Webtool uses the data of the TRANSFAC collection of TFs [21] as source for target sequence information and predicts potential TF binding sites by positional weight matrices [22, 23]. For both genes *FUT2* and *FUT3*, a 2000 bp long sequence upstream of exon 1, obtained from the Ensembl genome browser (https://www.ensembl.org/index.html), was analyzed [24] (online suppl. Tables 1, 2; for all online suppl. material, see https://doi.org/10.1159/000538863).

PROMO identified 43 and 48 potential TFs with multiple binding motifs for FUT2 and FUT3, respectively. Comparing the predicted TFs, we found that 30 of them were predicted to interact with both FUT2 and FUT3. Additionally, FUT2 contains unique sites for 13 TFs while we found 18 TFs exclusively for FUT3 (shown in Fig. 3a; Table 2).

We further filtered our predicted TFs in order to reduce the number of possible false positives. We only considered factors with PROMO's dissimilarity metric

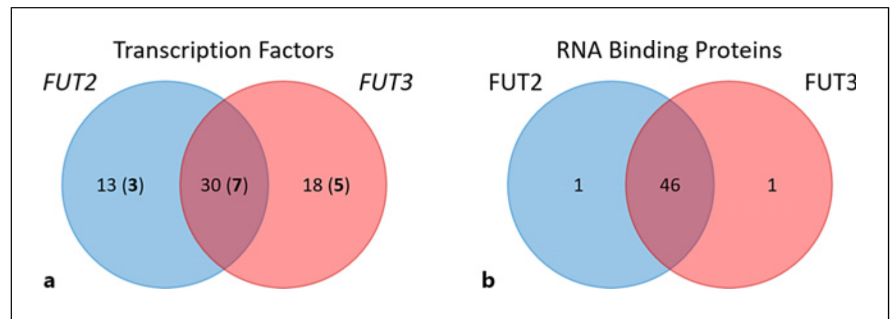

Fig. 3. TFs binding the genes (**a**) and RBPs binding the mRNAs (**b**) of *FUT2* and *FUT3* individually and their overlap. **a** TFs were predicted with PROMO (https://alggen.lsi.upc.es) [23]. Bold numbers in brackets correspond to TFs with E <1 and dissimilarity <0.3. **b** Only RBP predictions from both RBPmap (http://rbpmap.technion.ac.il/) [25] and RBPsuite database (http://www.csbio.sjtu.edu.cn/bioinf/RBPsuite/) [26] are considered. Venn diagrams were created using https://bioinformatics.psb.ugent.be/webtools/Venn/.

Table 2. List of individual and shared TFs and RBPs of FUT2 and FUT3

Target Gene	Transcription Factors	RNA Binding Proteins
Predicted to bind exclusively to *FUT2*	AP-1; AR; c-Fos; Elk-1; FOXO4; Ik-1; IRF-1; MAZ; NF-1; NF-AT1; PITX2; RBP-Jkappa; TBP	TRA2A
Predicted to bind exclusively to *FUT3*	c-Myb; **c-Myc**; CREB; GATA-2; GCF; HIF-1; **HNF-1A**; HOXD9; HOXD10; **IRF-2**; LEF-1; NF-Y; PEA3; RelA; SRY; TCF-4; **TCF-4E**; USF1	FXR2
Binding simultaneously to FUT2 & FUT3	**AP-2alphaA**; **c-Jun**; **C/EBPalpha**; C/EBPbeta; **c-Ets-1**; c-Ets-2; ER-alpha; FOXP3; **GATA-1**; **GR**; GR-alpha; GR-beta; **HNF-3alpha**; NF-AT1; **p53**; RAR-beta; Pax-5; **RXR-alpha**; PXR-1:RXR-alpha; PR A; PR B; Sp1; STAT4; T3R-beta1; TFII-I; TFIID; **USF2**;VDR; **XBP-1**; YY1	CPEB4; EIF4G2; EWSR1; FMR1; FUBP3; FUS; FXR1; HNRNPA1; HNRNPA2B1; HNRNPC; HNRNPD; HNRNPF; HNRNPH1; HNRNPK; HNRNPM; HNRNPL; HNRNPU; IGF2BP1; IGF2BP2; IGF2BP3; KHDRBS1; KHSRP; LIN28A; MATR3; PABPC4; PABPN1; PCBP2; PCBP1; PUM1; PUM2; QKI; RBFOX1; RBFOX2; RBM5; RBM15B; RBM22; RBM47; SFPQ; SRSF1; SRSF7; SRSF9; SRSF10; TAF15; TARDBP; TIA1; U2AF2;

Listed TFs were predicted with PROMO (*https://alggen.lsi.upc.es*). Bold TF has a lower false-positive probability (E <1) and a high similarity (dissimilarity value <0.3) to FUT2 (red), FUT3 (blue), or both (black). Underlined TFs are known to be involved in hematopoiesis and have been described to regulate expression of other blood group genes as well. Only RBP predictions from both RBPmap (http://rbpmap.technion.ac.il/) [27] and RBPsuite database (http://www.csbio.sjtu.edu.cn/bioinf/RBPsuite/) [28] are considered.

of <0.3, which indicates a high similarity to the TF's canonical target sequence. We further set the threshold for the E value to <1, to eliminate hits with a high chance of random occurrence of the cognate sequence. The resulting 15 TFs are blotted in Figure 4, showing possible overlaps of motifs and their exact location in a 2,000-bp sequence upstream of the transcription start. With this, we are able to identify TF motif clusters of around 200 bp length, which would correspond to enhancer regions.

By taking a closer look at the predicted regulators, we also found TFs that are currently well known for regulating blood group antigen expression. Among these, GATA-1 and GATA-2, which belong to the GATA family of TFs, have been described to play major roles during hematopoiesis: GATA-1 is responsible for driving the differentiation of

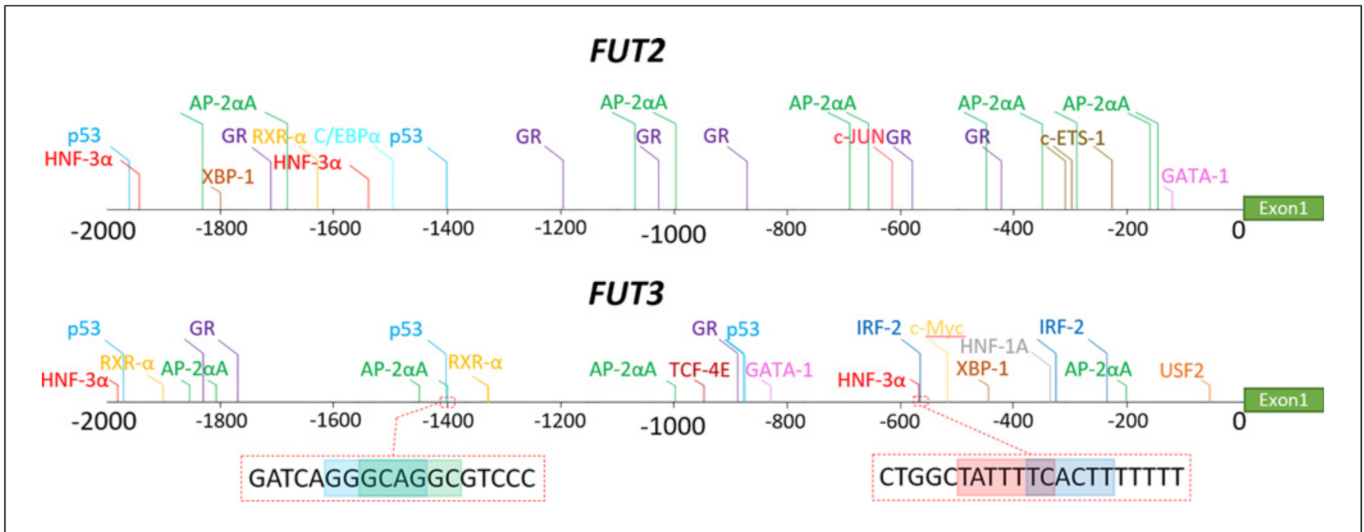

Fig. 4. Location of predicted TFs in *FUT2* and *FUT3*. Only TFs passing a stringency filter (dissimilarity <0.3 and E value <1) are shown. In *FUT3*, overlapping binding sites are observed for p53 and AP-2αA and for HNF-3α and IRF-2. The exact DNA binding site sequences are displayed in dashed boxes.

Table 3. List of TFs described to regulate FUT2 or FUT3 transcription in malignancies

TF	Transcription	Associated disease	Reference
c-Myc	FUT3 ↑	CRC	[33]
NF-κB	FUT3 ↑	HSV-1 infection	[35]
CDX1	FUT2/3 ↑	CRC	[36] [37]
CDX2	FUT2 ↑	CRC	[33]
ETS2	FUT3 ↑	CRC	[37]
HNF1A	FUT3 ↑	CRC	[37] [38]
HNF4A	FUT3 ↑	CRC	[37] [38]
MECOM	FUT3 ↑	CRC	[37]
MYB	FUT3 ↑	CRC	[37]
Snail	FUT3 ↓	Breast cancer	[39]
ZEB1/2	FUT3 ↓	Breast cancer	[39]

progenitor cells into erythrocytes and GATA-2 and 3 are crucial in maintaining constant self-renewal of hematopoietic stem cells [29]. More recently, GATA-1 has also been described to be involved in the transcriptional regulation of 33 blood group genes, also including sites for *FUT2* [30]. Consistent with this, PROMO predicted GATA-1 binding sites upstream of *FUT2*. For the PROMO-predicted GATA-1 site in *FUT3*, experimental evidence is still lacking.

Another highly interesting TF is Sp1, a zinc finger protein that is known to regulate expression via binding GC-rich elements in target genes, as mentioned above. It is considered essential during early development as Sp1 deficiency in mice is lethal at a very early developmental stage [31]. A recent study also linked Sp1 to the hematopoietic system as $Sp1^{-/-}$ mice lack the ability to progress beyond the progenitor phase of the red blood cell development and therefore are not able to build fully differentiated red blood cells [32]. Similar to GATA-1, PROMO predicts Sp1 binding sites in the promoters of both *FUT2* and *FUT3*, together suggesting a link between fucosyltransferase expression and healthy erythrocyte development.

Expression of *FUT2* and *FUT3* in Carcinogenesis, Tumor Growth, and Viral Infection

Although relatively little is known about the regulation of *FUT2* and *FUT3* expression in healthy individuals, a large body of literature exists on the relationship between *FUT2/FUT3* expression and cancer (shown in Table 3). Tumors exploit increased expression of *FUT2* and *FUT3* as it is linked to increased epithelial-mesenchymal transition (EMT) levels [33]. In colorectal cancer (CRC) diagnostics, high expression levels of sialyl-Le^a glycans are used as tumor markers, not only to predict overall survival but also metastatic potential [34]. Still, there are limitations as SLe^x is not exclusively synthesized by FUT 3 but rather by several alpha1,3-fucosyltransferases (see Fig. 2b). Therefore, additional detailed determination of the Lewis antigens on the RBC, potentially including molecular analysis, would be necessary for the use of SLe^x as a biomarker. However, detailing the fully comprehensive diagnostics workflow for full serological and molecular workup of such phenotypes is out of the scope of this review.

Table 4. miRNAs predicted to bind FUT2 and FUT3 mRNA

FUT2		
5′UTR	CDS	3′UTR
hsa-miR-6807-5p	hsa-miR-183-5p	hsa-miR-3663-5p
hsa-miR-6842-3p	hsa-miR-892a	hsa-miR-4717-3p
	hsa-miR-4433a-3p	hsa-miR-4729
	hsa-miR-4510	**hsa-miR-6127**
	hsa-miR-5187-5p	**hsa-miR-6133**
	hsa-miR-6127	**hsa-miR-6807-5p**
	hsa-miR-6133	hsa-miR-6829-3p
	hsa-miR-6842-3p	hsa-miR-7109-5p
	hsa-miR-7151-3p	**hsa-miR-7151-3p**
FUT3		
5′UTR	CDS	3′UTR
hsa-miR-502-5p	hsa-miR-4710	**hsa-miR-6764-5p**
hsa-miR-4539	**hsa-miR-502-5p**	hsa-miR-1294
hsa-miR-6764-5p	hsa-miR-9986	**hsa-miR-502-5p**
	hsa-miR-4690-5p	hsa-miR-1303
		hsa-miR-4690-5p
		hsa-miR-499a-3p
		hsa-miR-1303
		hsa-miR-2355-5p
		hsa-miR-3619-5p
		hsa-miR-214-3p
		hsa-miR-588
		hsa-miR-4726-3p

Only miRNAs also found in the miRBD database [50] are considered. Bold miRNAs are found in two different regions of the FUT2/3 RNA.

In colon cancer, the caudal-type homeobox 2 (CDX2) TF is described to bind to the *FUT2* promotor and upregulate its transcription [33]. CDX2 is exclusively expressed in the intestine and responsible for intestinal tissue differentiation [40]. Sakuma et al. [33] showed that during the EMT, colon cancer cells repress CDX2 and consequently the expression of *FUT2,* inducing a shift in favor of the sialyl Lewis x/a glycans ($sLe^{x/a}$). Additional binding site motifs for CDX1, HNF1A, and HNF4A were identified in the *FUT2* 5' regulatory region; however, these have not yet been further investigated [33, 36].

One of the TFs involved in *FUT3* expression is the well-known c-Myc, an oncogene sometimes referred as a "master regulator" due to its involvement in cellular signaling cascades regulating cellular growth and metabolism [41]. Upon the exposure to EGF or bFGF, c-Myc phosphorylation status changes, by elevating the levels of phospho-c-$Myc^{Ser62/Thr58}$. Especially the hyperphosphorylation of Ser62 of c-Myc is linked to its increased recruitment, binding to its target promoter and consequently causing an increase in *FUT3* expression in colon cancer cells [33].

Another TF that has been shown to mediate FUT3 transcription is nuclear factor-κB (NF-κB) [35]. There are five known members of the NF-κB family, NF-κB1 (p50), NF-κB2 (p52), RelA (p65), RelB, and c-Rel, which mediate transcription by interacting with specific DNA motifs as various hetero- or homo-dimers [42]. When human diploid fibroblasts infected with herpes simplex virus type 1 (HSV-1) were simultaneously treated with NF-κB inhibitors, they showed a significant decrease in *FUT3* expression. Short interfering RNA-mediated knockdown experiments identified RelA as the main activating TF as its absence resulted in a strong decrease of *FUT3* transcription. This is in line with our prediction of RelA binding to the distal activating element upstream of *FUT3*. However, it is worth noting that NF-κB alone was not able to induce elevated transcription levels and that a combination with additional factors seems to be necessary for driving transcription of *FUT3* [35].

In a recent study, Wang et al. [37] performed an in-depth analysis of 22 CRC cell lines investigating the role of different Lewis antigen levels as well as revealing various TFs mediating *FUT3* transcription. By comparing the glycosylation status of different Lewis antigens with the expression levels of known TFs, they were able to find a positive correlation of the TFs CDX1, ETS2, HNF1A, HNF4A, MECOM, and MYB with higher levels of Lewis-mediated fucosylation. The connection between CDX-1, highly fucosylated *N*-glycans, and elevated *FUT3* levels were validated by Holst et al. [36] who found higher *FUT3* gene expression levels and a higher abundance of multi-fucosylation in CDX-1 high colon cancer cell lines compared to CDX-1 low cells. Using a combination approach of high-throughput glycan analysis with a genome-wide association study, HNF1A and its downstream target HNF4A were shown to act as key regulators for the fucosyltransferase and fucose biosynthesis gene expression. Knockdown of both TFs led to a significant decrease of *FUT3* gene expression levels [38].

In contrast to the activating TFs discussed above, Snail and ZEB1/2 act as transcription repressors for *FUT3* [39]. Snail1 belongs to the Snail family of zinc-finger TFs, which are involved in embryonal development and have been shown to be major drivers of EMT [43]. ZEB1/2 are part of the ZEB family and are implicated in cellular plasticity, dissemination, and EMT as well [44]. Both Snail and ZEB1/2 have been shown to harbor various binding sites in the *FUT3* gene [39]. An analysis of 51 breast cancer cell lines also showed a negative correlation between Snail and ZEB1/2 and *FUT3* [39].

Notably, not all TFs described in this chapter have been found in our in silico prediction. This can be explained by the lack of recognition motifs for CDX-1, HNF4A, MECOM, Snail, and ZEB1/2 in the PROMO

database. The missing prediction of CDX2 binding to the *FUT2* promotor, however, is surprising, given its presence in the database.

Post-Transcriptional Regulation: An in silico Analysis

RNA Binding Proteins

Post-transcriptional gene regulation describes the process following mRNA synthesis and includes mRNA maturation, nuclear-cytosolic transport, mRNA stability, and translation. During all these processes, RNA-binding proteins (RBPs) influence the biogenesis, fate, and function of their target RNAs [45]. So far, 1,914 different RBPs have been characterized based on their ability to interact with RNA through a specific RNA-binding domain, corresponding to 7.5% of all protein-coding genes [27]. Their recognition motifs are frequently found in the 5′ and 3′ untranslated regions (UTRs) of the mRNAs [28].

To identify possible RBPs interacting with either *FUT2* or *FUT3* mRNA, we used two Webtools: RBPmap (http://rbpmap.technion.ac.il/) [46] and RBPsuite (http://www.csbio.sjtu.edu.cn/bioinf/RBPsuite/) [47]. Both are able to identify experimentally verified motifs on a target RNA sequence and predict the exact binding positions of RBPs (see online suppl. Tables 3, 4). This way, we identified 305 RBPs from both Webtools that are predicted to interact with *FUT2* mRNA. Further analysis resulted in 47 RBPs that are listed by both searches and are thus considered to more likely to bind *FUT2* mRNA. Similarly, for *FUT3*, we found 302 different RBPs that are predicted to interact with its mRNA. Among these, there were 47 *FUT3* hits overlapping for both database searches. Interestingly, *FUT2* and *FUT3* share 46 identical RBPs of each of their 47 hits and only show one single unique predicted RBP each: TRA2A and FXR2, respectively (shown in Fig. 3b; Table 2). Still, as for many other data reported here, experimental data are necessary to validate our observed in silico hits and prove the post-transcriptional processes that are affected by these RBPs.

One RBP known to interact with *FUT3* mRNA is DDX39B, a member of the DEAD box RNA helicase family, which is characterized by the presence of a preserved DEAD motif (Asp-Glu-Ala-Asp). DEAD box RNA helicases are known to be involved in nearly all cellular RNA metabolic processes, like splicing, ribosome biogenesis, nuclear export, translation, and degradation [25]. In human CRC, DDX39B mediates *FUT3* mRNA splicing and its nuclear export, resulting in increased Lewis expression levels [26]. Interestingly, in the CRC context, elevated *FUT3* expression correlates with increased fucosylation of TGFβR-1, which promotes TGFβ downstream signaling and consequently initiates the EMT program and promotes tumor progression [26]. Of note, DDX39B was not included in either of the two databases used for our in silico analysis.

Micro-RNAs

miRNAs are small, non-coding RNA molecules that regulate gene expression by hybridizing with their specific target mRNA, either inducing its degradation or inhibiting translation. We used the miRWalk Webtool (http://mirwalk.umm.uni-heidelberg.de/) developed by the University of Heidelberg to search for possible miRNA binding sites in the *FUT2* and *FUT3* mRNAs [48]. To further corroborate our search results, we only considered miRNAs that we could confirm using the additional database miRBD [49]. With this, we identified 15 miRNAs that are predicted to target *FUT2* and 15 different miRNAs that harbor one or more binding sites on *FUT3* (shown in Table 4). Mapping the exact positions of the predicted miRNA binding sites, we located 9 sites in both, the coding sequence and 3′UTR of FUT2, while in FUT3 the majority of 12 sites is located in the 3′UTR.

Interestingly, one of our candidates, hsa-miR.502-5p, is predicted to have binding motifs in the 5′UTR, coding sequence, and 3′UTR of *FUT3* mRNA. This miRNA has been described as negative regulator for the TF SP-1 and was shown to inhibit proliferation, migration, and invasion in gastric cancer cells, labeling it as a tumor suppressor [51, 50].

An additional predicted regulator miRNA is miR-17-5p, which has been shown to interact with *FUT2* mRNA and act as negative regulator for its expression [52]. In the setting of osteoarthritis, this downregulation of *FUT2* is abolished by the non-coding RNA HOTAIR, which is able to bind miR-17-5p and in this way indirectly restores the expression of *FUT2* [52].

Summary, Conclusions, and Outlook

There is accumulating evidence that regulatory control of *FUT2* and *FUT3* expression and resulting changes in levels of Le^a and Le^b fucosylation or shift to $sLex/^a$ might have important implications in several pathologies such as IBD and cancer. In order to shed light on roles of the different Lewis antigens in these pathologies, it is essential to understand which regulators are involved, at what stage they take effect, and how they affect the expression of their target. With our in silico approach to identify novel candidate TFs for both *FUT2* and *FUT3*, we predicted several interesting new factors that have not been described yet. TFs like GATA-1, known already for regulating other blood group antigens and being of general importance in hematopoiesis [30], or Sp1, described to be essential during hematopoietic differentiation [32], are noteworthy hits that warrant further investigation and experimental validation. Combining

these in silico data with regulators already validated in experimental studies, such as c-Myc, CDX1, or HNF1A, will improve our understanding of the complex transcriptional regulation of *FUT2* and *FUT3* genes.

When it comes to identifying factors that influence the expression of FUT2 and FUT3, there are pronounced gaps between our in silico findings and the experimental in vitro reports by others. This may be due to the fact that binding and action of TFs in FUT2 and FUT3 expression have as yet been poorly studied in experiments. It might also be due to the lack of known binding motifs in the PROMO databases or the nucleotide range used for the analysis. We chose the range of 2000 bp upstream of the transcription start site, according to previous studies indicating the regulatory role of this area [18, 53]. Increasing the area further upstream could possibly show additional TF sites, and looking for downstream regulatory elements could be of great interest for future studies. Another reason for a low overlap of experimentally identified and electronically predicted TFs involved in the expression of FUT2 and FUT3 might be the tendency of different TFs to interact with each other. They can either co-operate in supporting each other to bind [54] or be in competition with each other [55], this way adding another layer of transcriptional regulation. In addition, it must be assumed that the competition of different TFs with identical or overlapping binding motifs (e.g., for p53 and AP-2αA and for HNF-3α and IRF-2, see Fig. 4) will be governed by their respective binding affinities to the target genes [56]. This is of particular importance as most of the TF examples studied in vitro were described in cancer cells, which are generally known to have drastically altered gene expression profiles in comparison to healthy cells, hence might present with altered TF concentrations, shifting binding equilibria in competitive scenarios.

Additionally, with RBPs and miRNAs inciting more and more interest in their regulatory role in recent years, we were able to identify novel targets that are likely to interact with *FUT2* or *FUT3* and regulate their expression on a post-transcriptional level. Taken together, these points will improve our understanding of the regulation of *FUT2* and *FUT3* and consequently of the expression of the Lewis blood group antigens. This could also be transferred into diagnostic use as some of the TFs mentioned in this review could be used as prediction markers for susceptibility to IBDs or to predict overall survival and metastatic potential in colon cancer.

Conflict of Interest Statement

C.G. acts as a consultant to Inno-Train GmbH, Kronberg im Taunus, Germany, a provider of genotyping kits for molecular blood group diagnostics since 1998 and holds the European and US patents P3545102 and US20190316189 on the "determination of the genotype underlying the S-s-U-phenotype of the MNSs blood group system." All other authors state no conflict of interest.

Funding Sources

The Institute of Translational Medicine of the Private University in the Principality of Liechtenstein is financially supported by the Hans Groeber Stiftung, Vaduz, Fürstentum Liechtenstein, and the Tarom Foundation, Schaan, Fürstentum Liechtenstein.

Author Contributions

M.W. and C.G. designed the study, collected and analyzed data, and wrote the paper. S.M. and M.B. collected data and reviewed the paper.

References

1 Oriol R, Le Pendu J, Sparkes RS, Sparkes MC, Crist M, Gale RP, et al. Insights into the expression of ABH and Lewis antigens through human bone marrow transplantation. Am J Hum Genet. 1981;33(4):551–60.
2 Hauser R. Le(a) and Le(b) tissue glycosphingolipids. Transfusion. 1995;35(7): 577–81. https://doi.org/10.1046/j.1537-2995.1995.35795357880.x
3 Trinchera M, Aronica A, Dall'Olio F. Selectin ligands sialyl-Lewis a and sialyl-Lewis x in gastrointestinal cancers. Biology. 2017;6(1):16. https://doi.org/10.3390/biology6010016
4 de Vries T, Knegtel RM, Holmes EH, Macher BA. Fucosyltransferases: structure/function studies. Glycobiology. 2001;11(10):119R–128R. https://doi.org/10.1093/glycob/11.10.119r
5 Hu M, Zhang X, Li J, Chen L, He X, Sui T. Fucosyltransferase 2: a genetic risk factor for intestinal diseases. Front Microbiol. 2022;13: 940196. https://doi.org/10.3389/fmicb.2022.940196
6 Barbé L, Le Moullac-Vaidye B, Echasserieau K, Bernardeau K, Carton T, Bovin N, et al. Histo-blood group antigen-binding specificities of human rotaviruses are associated with gastroenteritis but not with in vitro infection. Sci Rep. 2018;8(1):12961. https://doi.org/10.1038/s41598-018-31005-4
7 Cartron JP, Molecular basis of human blood group antigens. Boston, MA: Springer US; Imprint; Springer; 1995.
8 Zheng J, Zhu T. Polymorphism of fucosyltransferase 3 gene is associated with inflammatory bowel disease: a systematic review. Asian Biomed. 2023;17(2):45–54. https://doi.org/10.2478/abm-2023-0044
9 Forni D, Cleynen I, Ferrante M, Cassinotti A, Cagliani R, Ardizzone S, et al. ABO histo-blood group might modulate predisposition to Crohn's disease and affect disease behavior. J Crohns Colitis. 2014;8(6):489–494. https://doi.org/10.1016/j.crohns.2013
10 McGovern DPB, Jones MR, Taylor KD, Marciante K, Yan X, Dubinsky M, et al. Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. Hum Mol Genet. 2010;19(17):3468–76. https://doi.org/10.1093/hmg/ddq248
11 Rausch P, Rehman A, Künzel S, Häsler R, Ott SJ, Schreiber S, et al. Colonic mucosa-associated microbiota is influenced by an interaction of Crohn disease and FUT2 (Secretor) genotype. Proc Natl Acad Sci U S A. 2011;108(47):19030–5. https://doi.org/10.1073/pnas.1106408108
12 Hu D, Zhang D, Zheng S, Guo M, Lin X, Jiang Y. Association of ulcerative colitis with FUT2 and FUT3 polymorphisms in patients from southeast China. PLoS One. 2016;11(1): e0146557. https://doi.org/10.1371/journal.pone.0146557

13 Godefroy E, Barbé L, Le Moullac-Vaidye B, Rocher J, Breiman A, Leuillet S, et al. Microbiota-induced regulatory T cells associate with FUT2-dependent susceptibility to rotavirus gastroenteritis. Front Microbiol. 2023;14:1123803. https://doi.org/10.3389/fmicb.2023.1123803
14 Casamassimi A, Ciccodicola A. Transcriptional regulation: molecules, involved mechanisms, and misregulation. Int J Mol Sci. 2019;20(6):1281. https://doi.org/10.3390/ijms20061281
15 Dhar GA, Saha S, Mitra P, Nag Chaudhuri R. DNA methylation and regulation of gene expression: guardian of our health. Nucleus. 2021;64(3):259–70. https://doi.org/10.1007/s13237-021-00367-y
16 Wu Z, Shi D, Jin J, Fan H, Bao W, Wu S. DNA methylation of pig FUT3 promoter alters mRNA expression to regulate E. coli F18 susceptibility. Genes. 2021;12(10):1586. https://doi.org/10.3390/genes12101586
17 Yang L, Liu L, Cheng J, Wu Z, Bao W, Wu S. Association analysis of DNA methylation and the tissue/developmental expression of the FUT3 gene in Meishan pigs. Gene. 2023;851:147016. https://doi.org/10.1016/j.gene.2022.147016
18 Serpa J, Mesquita P, Mendes N, Oliveira C, Almeida R, Santos-Silva F, et al. Expression of Lea in gastric cancer cell lines depends on FUT3 expression regulated by promoter methylation. Cancer Lett. 2006;242(2):191–7. https://doi.org/10.1016/j.canlet.2005.11.009
19 Krishnan NM, Dhas K, Nair J, Palve V, Bagwan J, Siddappa G, et al. A minimal DNA methylation signature in oral tongue squamous cell carcinoma links altered methylation with tumor attributes. Mol Cancer Res. 2016;14(9):805–19. https://doi.org/10.1158/1541-7786.MCR-15-0395
20 Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. Cell. 2013;152(6):1237–51. https://doi.org/10.1016/j.cell.2013.02.014
21 Matys V, Fricke E, Geffers R, Gössling E, Haubrock M, Hehl R, et al. TRANSFAC: transcriptional regulation, from patterns to profiles. Nucleic Acids Res. 2003;31(1):374–8. https://doi.org/10.1093/nar/gkg108
22 Farré D, Roset R, Huerta M, Adsuara JE, Roselló L, Albà MM, et al. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. Nucleic Acids Res. 2003;31(13):3651–3. https://doi.org/10.1093/nar/gkg605
23 Messeguer X, Escudero R, Farré D, Núñez O, Martínez J, Albà MM. PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinforma. 2002;18(2):333–4. https://doi.org/10.1093/bioinformatics/18.2.333
24 Martin FJ, Amode MR, Aneja A, Austine-Orimoloye O, Azov AG, Barnes I, et al. Ensembl 2023. Nucleic Acids Res. 2023;51(D1):D933–41. https://doi.org/10.1093/nar/gkac958
25 Linder P, Jankowsky E. From unwinding to clamping: the DEAD box RNA helicase family. Nat Rev Mol Cell Biol. 2011;12(8):505–16. https://doi.org/10.1038/nrm3154
26 He C, Li A, Lai Q, Ding J, Yan Q, Liu S, et al. The DDX39B/FUT3/TGFβR-I axis promotes tumor metastasis and EMT in colorectal cancer. Cell Death Dis. 2021;12(1):74. https://doi.org/10.1038/s41419-020-03360-6
27 Qin H, Ni H, Liu Y, Yuan Y, Xi T, Li X, et al. RNA-binding proteins in tumor progression. J Hematol Oncol. 2020;13(1):90. https://doi.org/10.1186/s13045-020-00927-w
28 Hodson DJ, Screen M, Turner M. RNA-binding proteins in hematopoiesis and hematological malignancy. Blood. 2019;133(22):2365–73. https://doi.org/10.1182/blood-2018-10-839985
29 Tremblay M, Sanchez-Ferras O, Bouchard M. GATA transcription factors in development and disease. Development. 2018;145(20):dev164384. https://doi.org/10.1242/dev.164384
30 Wu PC, Lee YQ, Möller M, Storry JR, Olsson ML. Elucidation of the low-expressing erythroid CR1 phenotype by bioinformatic mining of the GATA1-driven blood-group regulome. Nat Commun. 2023;14(1):5001. https://doi.org/10.1038/s41467-023-40708-w
31 Marin M, Karis A, Visser P, Grosveld F, Philipsen S. Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation. Cell. 1997;89(4):619–28. https://doi.org/10.1016/s0092-8674(00)80243-3
32 Gilmour J, Assi SA, Jaegle U, Kulu D, van de Werken H, Clarke D, et al. A crucial role for the ubiquitously expressed transcription factor Sp1 at early stages of hematopoietic specification. Development. 2014;141(12):2391–401. https://doi.org/10.1242/dev.106054
33 Sakuma K, Aoki M, Kannagi R. Transcription factors c-Myc and CDX2 mediate E-selectin ligand expression in colon cancer cells undergoing EGF/bFGF-induced epithelial-mesenchymal transition. Proc Natl Acad Sci U S A. 2012;109(20):7776–81. https://doi.org/10.1073/pnas.1111135109
34 Matsui T. Sialyl Lewisa expression as a predictor of the prognosis of colon carcinoma patients in a prospective randomized clinical trial. JPN J Clin Oncol. 2004;34(10):588–93. https://doi.org/10.1093/jjco/hyh110
35 Nordén R, Samuelsson E, Nyström K. NFκB-mediated activation of the cellular FUT3, 5 and 6 gene cluster by herpes simplex virus type 1. Glycobiology. 2017;27(11):999–1005. https://doi.org/10.1093/glycob/cwx079
36 Holst S, Wilding JL, Koprowska K, Rombouts Y, Wuhrer M. N-glycomic and transcriptomic changes associated with CDX1 mRNA expression in colorectal cancer cell lines. Cells. 2019:8.
37 Wang D, Madunić K, Zhang T, Mayboroda OA, Lageveen-Kammeijer GSM, Wuhrer M. High diversity of glycosphingolipid glycans of colorectal cancer cell lines reflects the cellular differentiation phenotype. Mol Cell Proteomics. 2022;21(6):100239. https://doi.org/10.1016/j.mcpro.2022.100239
38 Lauc G, Essafi A, Huffman JE, Hayward C, Knežević A, Kattla JJ, et al. Genomics meets glycomics-the first GWAS study of human N-Glycome identifies HNF1α as a master regulator of plasma protein fucosylation. PLoS Genet. 2010;6(12):e1001256. https://doi.org/10.1371/journal.pgen.1001256
39 Breiman A, López Robles MD, de Carné Trécesson S, Echasserieau K, Bernardeau K, Drickamer K, et al. Carcinoma-associated fucosylated antigens are markers of the epithelial state and can contribute to cell adhesion through CLEC17A (Prolectin). Oncotarget. 2016;7(12):14064–82. https://doi.org/10.18632/oncotarget.7476
40 Saad RS, Ghorab Z, Khalifa MA, Xu M. CDX2 as a marker for intestinal differentiation: its utility and limitations. World J Gastrointest Surg. 2011;3(11):159–66. https://doi.org/10.4240/wjgs.v3.i11.159
41 Miller DM, Thomas SD, Islam A, Muench D, Sedoris K. c-Myc and cancer metabolism. Clin Cancer Res. 2012;18(20):5546–53. https://doi.org/10.1158/1078-0432.ccr-12-0977
42 Liu T, Zhang L, Joo D, Sun SC. NF-κB signaling in inflammation. Signal Transduct Target Ther. 2017;2:17023. https://doi.org/10.1038/sigtrans.2017.23
43 Wu Y, Zhou BP. Snail: more than EMT. Cell Adh Migr. 2010;4(2):199–203. https://doi.org/10.4161/cam.4.2.10943
44 Drápela S, Bouchal J, Jolly MK, Culig Z, Souček K. ZEB1: a critical regulator of cell plasticity, DNA damage response, and therapy resistance. Front Mol Biosci. 2020;7:36. https://doi.org/10.3389/fmolb.2020.00036
45 Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nat Rev Genet. 2014;15(12):829–45. https://doi.org/10.1038/nrg3813
46 Paz I, Kosti I, Ares M, Cline M, Mandel-Gutfreund Y. RBPmap: a web server for mapping binding sites of RNA-binding proteins. Nucleic Acids Res. 2014;42(Web Server issue):W361–7. https://doi.org/10.1093/nar/gku406
47 Pan X, Fang Y, Li X, Yang Y, Shen HB. RBPsuite: RNA-protein binding sites prediction suite based on deep learning. BMC genomics. 2020;21(1):884. https://doi.org/10.1186/s12864-020-07291-6
48 Sticht C, de La Torre C, Parveen A, Gretz N. miRWalk: an online resource for prediction of microRNA binding sites. PLoS One. 2018;13(10):e0206239. https://doi.org/10.1371/journal.pone.0206239
49 Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res. 2020;48(D1):D127–31. https://doi.org/10.1093/nar/gkz757
50 You W, Liu X, Yu Y, Chen C, Xiong Y, Liu Y, et al. miR-502-5p affects gastric cancer progression by targeting PD-L1. Cancer Cell Int. 2020;20:395. https://doi.org/10.1186/s12935-020-01479-2
51 Peng X, Wu M, Liu W, Guo C, Zhan L, Zhan X. miR-502-5p inhibits the proliferation, migration and invasion of gastric cancer cells by targeting SP1. Oncol Lett. 2020;20(3):2757–62. https://doi.org/10.3892/ol.2020.11808
52 Hu J, Wang Z, Shan Y, Pan Y, Ma J, Jia L. Long non-coding RNA HOTAIR promotes osteoarthritis progression via miR-17-5p/FUT2/β-catenin axis. Cell Death Dis. 2018;9(7):711. https://doi.org/10.1038/s41419-018-0746-z

53 Dabrowska A, Baczyńska D, Widerak K, Laskowska A, Ugorski M. Promoter analysis of the human alpha1,3/4-fucosyltransferase gene (FUT III). Biochim Biophys Acta. 2005;1731(1):66–73. https://doi.org/10.1016/j.bbaexp.2005.08.009
54 Morgunova E, Taipale J. Structural perspective of cooperative transcription factor binding. Curr Opin Struct Biol. 2017;47:1–8. https://doi.org/10.1016/j.sbi.2017.03.006
55 Zhang Y, Ho TD, Buchler NE, Gordân R. Competition for DNA binding between paralogous transcription factors determines their genomic occupancy and regulatory functions. Genome Res. 2021;31(7):1216–29. https://doi.org/10.1101/gr.275145.120
56 Kribelbauer JF, Rastogi C, Bussemaker HJ, Mann RS. Low-affinity binding sites and the transcription factor specificity paradox in eukaryotes. Annu Rev Cell Dev Biol. 2019;35:357–79. https://doi.org/10.1146/annurev-cellbio-100617-062719