



**HAL**  
open science

# N-Glycomic and Transcriptomic Changes Associated with CDX1 mRNA Expression in Colorectal Cancer Cell Lines

Stephanie Holst, Jennifer Wilding, Kamila Koprowska, Yoann Rombouts,  
Manfred Wuhrer

## ► To cite this version:

Stephanie Holst, Jennifer Wilding, Kamila Koprowska, Yoann Rombouts, Manfred Wuhrer. N-Glycomic and Transcriptomic Changes Associated with CDX1 mRNA Expression in Colorectal Cancer Cell Lines. *Cells*, 2019, 8 (3), pp.273. 10.3390/cells8030273 . hal-02347740

**HAL Id: hal-02347740**

**<https://hal.science/hal-02347740v1>**

Submitted on 23 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.




L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Article

# N-Glycomic and Transcriptomic Changes Associated with CDX1 mRNA Expression in Colorectal Cancer Cell Lines

Stephanie Holst <sup>1,\*</sup>, Jennifer L. Wilding <sup>2</sup>, Kamila Koprowska <sup>2</sup>, Yoann Rombouts <sup>1,3</sup>  
and Manfred Wuhrer <sup>1,\*</sup>

<sup>1</sup> Center for Proteomics and Metabolomics, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands

<sup>2</sup> Cancer and Immunogenetics Laboratory, Weatherall Institute of Molecular Medicine, Department of Oncology, University of Oxford, Oxford OX3 9DS, UK; wildingjl@gmail.com (J.L.W.); kamila.koprowska@gmail.com (K.K.)

<sup>3</sup> Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France; yoann.rombouts@ipbs.fr

\* Correspondence: stephanie.holst@gmx.net (S.H.); m.wuhrer@lumc.nl (M.W.); Tel.: +31-71-526-8744 (M.W.)

Received: 3 March 2019; Accepted: 18 March 2019; Published: 22 March 2019



**Abstract:** The caudal-related homeobox protein 1 (CDX1) is a transcription factor, which is important in the development, differentiation, and homeostasis of the gut. Although the involvement of CDX genes in the regulation of the expression levels of a few glycosyltransferases has been shown, associations between glycosylation phenotypes and CDX1 mRNA expression have hitherto not been well studied. Triggered by our previous study, we here characterized the *N*-glycomic phenotype of 16 colon cancer cell lines, selected for their differential CDX1 mRNA expression levels. We found that high CDX1 mRNA expression associated with a higher degree of multi-fucosylation on *N*-glycans, which is in line with our previous results and was supported by up-regulated gene expression of fucosyltransferases involved in antenna fucosylation. Interestingly, hepatocyte nuclear factors (HNF)4A and HNF1A were, among others, positively associated with high CDX1 mRNA expression and have been previously proven to regulate antenna fucosylation. Besides fucosylation, we found that high CDX1 mRNA expression in cancer cell lines also associated with low levels of sialylation and galactosylation and high levels of bisection on *N*-glycans. Altogether, our data highlight a possible role of CDX1 in altering the *N*-glycosylation of colorectal cancer cells, which is a hallmark of tumor development.

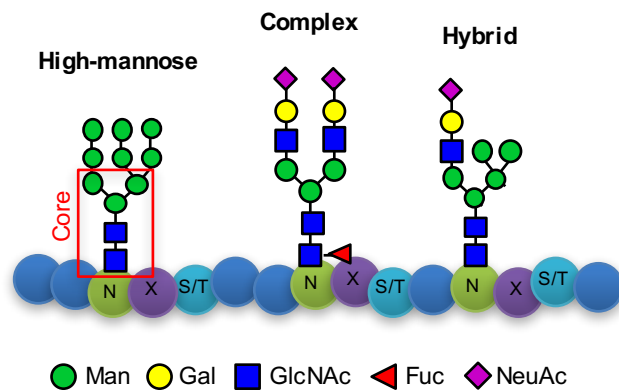
**Keywords:** colorectal cancer cell lines; *N*-glycosylation; fucosylation; caudal-related homeobox protein 1 (CDX1), differentiation; fucosyltransferase (FUT), hepatocyte nuclear factor (HNF)4A; HNF1A

## 1. Introduction

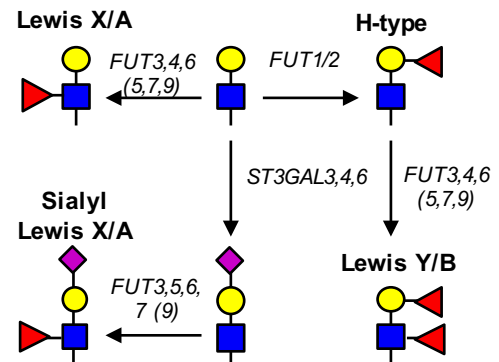
Glycans form an important part of the outer layer of the cell and are involved in major biological processes, including cell differentiation, adhesion, and interactions with other cells, pathogens, or the extracellular matrix, as well as cellular transformations such as cancer [1–3]. Glycans occur in many structural variants as modifications on proteins and lipids. One type of glycosylation on proteins is the *N*-glycosylation, in which oligosaccharides are attached via an *N*-glycosidic linkage to an asparagine (N) of the consensus sequence -N-X-S/T- (X is any amino acid except proline, S = serine, T is threonine) [4]. These so-called *N*-glycans share a common pentasaccharide core-structure and form, depending on the elongation, high-mannose type, complex-type, or hybrid-type *N*-glycans (Figure 1A) [4]. Changes of specific *N*-glycans and other glycans have been associated with several

diseases and cancers [5–7] and colorectal cancer associated glycan changes have been recently reviewed by us [8].

### A N-glycans



### B Lewis antigens



**Figure 1.** N-glycans and Lewis antigens. **(A)** N-glycans are attached to an asparagine (N) in the consensus sequence N-X-S/T with X any amino acid except proline, followed by serine (S) or threonine (T). Main monosaccharides involved in N-glycosylation are mannoses (Man), galactoses (Gal), N-acetylglucosamine (GlcNAc), fucoses (Fuc), and N-aceylneuraminic acid (NeuAc). N-glycans share a common core-structure, consisting of two GlcNAc and three Man. Depending on the elongation, three N-glycan types are differentiated, as follows: i) High-mannose type N-glycans, ii) complex-type N-glycans, and iii) hybrid-type N-glycans. The illustrated glycans represent examples. The number of monosaccharides added can vary and complex type glycans can exhibit more than two antennae. A detailed description of N-glycosylation is given by Stanley [4]. **(B)** Depiction of different Lewis-antigens and involved glycosyltransferase genes. Fucosyltransferase genes *FUT3,4,5,6,7,9* are involved in fucosylation  $\alpha$ 1,3- and  $\alpha$ 1,4-fucosylation of antennae GlcNAc, while *FUT1,2* attach  $\alpha$ 1,2-fucosylation to Gal-residues. Activity of several  $\alpha$ 2,3-sialyltransferase (*ST3GAL* genes) attach NeuAc residues to galactoses to form sialyl Lewis antigens.

By screening different colorectal cancer cell lines, we previously showed an association of caudal-related homeobox protein 1 (CDX1) mRNA expression with increased multi-fucosylation (more than one fucose), which is indicative for antenna fucosylation [9]. Antenna fucosylation is the result of the activity of fucosyltransferases encoded by the genes *FUT1, 2, 3, 4, 5, 6, 7, and 9* and leads to the expression of blood group related Lewis antigens (Figure 1B), which are glycan-epitopes on various glycoproteins and glycolipids [10,11]. While *FUT3,4,5,6,7, and 9* catalyze the addition of a fucose to the N-acetylglucosamine (GlcNAc) of the antenna in  $\alpha$ 1,3- and/or  $\alpha$ 1,4-linkage, *FUT1* and *FUT2* are responsible for the addition of a fucose to the galactose (Gal) in  $\alpha$ 1,2-linkage, forming the H-type epitope. *FUT2* is further called the secretor gene and polymorphisms leading to an inactive *FUT2* lead to the absence of blood type epitopes in saliva and various epithelial cell types, the so-called non-secretor phenotype [12,13]. *FUT8*, on the other hand, is the gene encoding for the enzyme mediating the attachment of a fucose-residue to the first GlcNAc of the N-glycan core, the so-called core-fucosylation. More detailed information on the biological role of fucosylation has been given elsewhere [14].

Fucosyltransferases are expressed in a cell type- and organ-dependent manner and altered expression of the enzymes, as well as the produced fucose-containing glycan epitopes, have been associated with various pathological conditions [15], including colorectal cancer [16,17]. However, literature remains controversial concerning the stage-dependent differences in expression of fucosylated glyco-epitopes in colorectal cancer. Conflicting data arises from various experimental conditions (lectin and antibody based vs. mass spectrometric detection, sample preparation,

and culture conditions) and sample sources (cell lines, mouse models, tissues vs. serum or plasma, biopsy location) as well as combining results from studies on various glycan classes (*N*-, *O*-, and glycosphingolipid glycans) vs. individual glycan classes. Moreover, core- and antenna-fucosylation are not always distinguished and only a few studies report on stage-dependent changes. While many groups reported a general increase of fucosylation in colon and other cancers [15], other recent studies have shown a stage-dependent decrease of fucosylation, especially Lewis-type antenna-fucosylation, with colon cancer metastasis [18,19]. Notably, it is still largely unclear to which extent differences in the glycosylation machinery, on the one hand, and differences in the expression levels of highly glycosylated proteins such as mucins (e.g., MUC2), on the other hand, contribute to these colorectal cancer glycosylation phenotypes.

The group around Miyoshi and Moriwaki hypothesize that deficiency of fucosylation via mutations of GDP-mannose-4,6-dehydratase (*GMDS*), an important player in the fucose biosynthesis pathway, helps cancer cells to escape from natural killer (NK) cell-mediated tumor immune surveillance [20,21]. Interestingly, mutations in the *GMDS* gene have been identified in metastatic lesions of some colon cancers (10%). Also the colon cancer cell line HCT116 bears a *GMDS* mutation, resulting in almost complete absence of fucosylation. Strikingly, the parental HCT116 cells with *GMDS* mutation revealed a more aggressive phenotype in tumor formation and metastasis in mice, as compared to the *GMDS*-rescued HCT116 cells [17]. The group speculated that the loss of fucosylation leads to an escape from NK cell-mediated tumor immune surveillance, while the *GMDS*-rescued HCT116 were more susceptible to TRAIL-induced apoptosis [21].

For the sialylated variants of the Lewis antigens (Figure 1B), sialyl Lewis X and sialyl Lewis A, the literature is clearer. The selectin ligands sialyl Lewis X and A have been associated with advanced stages and poor prognosis in colon cancer [22,23] which were attributed to selectin-mediated adhesion and extravasation leading to metastasis. Moreover, sialyl Lewis antigens are used as clinical markers for various cancers (e.g., sialyl Lewis A = CA19-9) [16].

CDX1 is a transcription factor that is involved in the modulation of a variety of processes including proliferation, apoptosis, cell-adhesion, and columnar morphology [24]. CDX1 is a primary controller of enterocyte differentiation and its expression is necessary for the transcriptional regulation of a large number of intestine-specific genes [25] required for the intestine development, differentiation, and maintenance of the intestinal phenotype [24,26]. Several markers of differentiation, including villin and cytokeratin 20, have been shown to be directly transcriptionally regulated by CDX1 [27,28]. Loss of differentiation is known to occur during cancer progression and, in colon cancer, the loss of villin has been associated with poor survival [29].

There is evidence of the loss or down-regulation of CDX1 expression in colon cancer tumors [30,31] and cell lines [32]. With respect to the underlying mechanism, Wong et al. have shown that the loss or reduction of CDX1 is often induced by promoter methylation [33]. Together, these observations indicate a potential role of CDX1 loss in tumor development.

Although glycosylation as well as CDX1 expression have both been shown to play a role in cell differentiation and cancer (suppression), hitherto only very little has been described with regard to CDX1-associated glycosylation profiles. Triggered by our previous findings [9], which suggested a relationship between increased multi-fucosylation (Lewis type glycan epitopes) and high CDX1 mRNA expression, here we characterized the *N*-glycosylation phenotype of a different set of colon cancer cell lines. These cells were specifically chosen for their high and low expression of CDX1 mRNA, with the aim of replicating our previously found association of fucosylation and CDX1 expression [9] with an independent set of cell lines. The glyco-phenotypic data were further correlated with differently expressed glyco-genes.

We confirmed a higher degree of multi-fucosylation in cell lines with high CDX1 gene expression, which was supported by higher mRNA expression of fucosyltransferases *FUT3* and *FUT6*, both involved in antenna fucosylation. Strikingly, other genes associated with fucosylation were likewise differentially expressed between the investigated cell lines with high vs. low CDX1

expression. A higher *GMDS* gene expression is positively associated with high CDX1 mRNA expression in the tested colorectal cancer cell lines and also transcription factors hepatocyte nuclear factor (*HNF1A* and *HNF4A*), which have been previously shown to regulate antenna fucosylation of human plasma proteins [34] and are associated with differentiation [35,36], showed higher expression in CDX1-high expressing cells. Furthermore, the *N*-glycomic characterization revealed a decrease in galactosylation and sialylation in cell lines with high CDX1 expression, compared to cells with low CDX1 gene expression.

## 2. Materials and Methods

### 2.1. Materials

8M guanidine hydrochloride (GuHCl), 1-hydroxybenzotriazole (HOBt) hydrate, 50% sodium hydroxide (NaOH), and super DHB matrix (2-hydroxy-5-methoxy-benzoic acid and 2,5-Dihydroxybenzoic acid, 1:9) were obtained from Sigma-Aldrich (Steinheim, Germany). HPLC SupraGradient acetonitrile (ACN) was obtained from Biosolve (Valkenswaard, The Netherlands). Dithiothreitol (DTT), ethanol, and sodium bicarbonate ( $\text{NaHCO}_3$ ) were from Merck (Darmstadt, Germany) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) was from Fluorochem (Hadfield, UK). Peptide *N*-Glycosidase F (PNGase F) was purchased from Roche Diagnostics (Mannheim, Germany). The peptide calibration standard was purchased from Bruker Daltonics (Bremen, Germany). MultiScreen<sup>®</sup> HTS 96 multiwell plates (pore size 0.45  $\mu\text{m}$ ) with high protein-binding membrane (hydrophobic Immobilon-P PVDF membrane) were purchased from Millipore (Amsterdam, The Netherlands), while the 96-well polypropylene 0.8 mL 96-deepwell plate and 96-well PCR plate polypropylene were from Greiner Bio (Alphen a/d Rijn, The Netherlands). All buffers were prepared using ultrapure water generated from an 18.2 M $\Omega$ -cm Purelab Ultra system from Elga (Ede, The Netherlands). Control Visucon-F plasma pool (citrate and 0.02 M HEPES buffered plasma pool from 20 healthy human donors) was obtained from Affinity Biologicals (Ancaster, Canada). Cell culture media were purchased from Gibco (Paisley, UK), fetal bovine serum (FBS) from Autogen Bioclear (Wiltshire, UK), and penicillin-streptomycin from Invitrogen (Paisley, UK). The RNeasy mini kit was purchased from Qiagen (Manchester, UK).

### 2.2. Cells and Cell Culture

Details of colorectal cancer cell line origin and characteristics can be found in Table 1 and more detailed in Supplementary Table S1. CC20, COLO678, GP2D, HCT116, ISRECO1, LIM1863, LS174T, PCJW, RCM1, and SW403 cell lines were cultured in the Dulbecco Modified Eagle Medium (DMEM). CAR1, HCA46, HDC8, OXCO1, and VACO429 were cultured in Iscove's Modified Dulbecco's Medium (IMDM). HCC56 cells were cultured in the RPMI-1640 medium. All of the above media were supplemented with 10% FBS and 1% penicillin-streptomycin. Cell lines were grown in a 10%  $\text{CO}_2$  incubator to 50% to 80% confluence before next passage or cell pellet preparation. Cell pellets were washed once with PBS, snap-frozen in dry ice, and stored at  $-80^\circ\text{C}$  until further analysis. Cell pellets were dissolved in water and cell counts were estimated using the Countess<sup>™</sup> Automated Cell Counter (Invitrogen), based on trypan blue staining.

**Table 1. Cell line characteristics.** CDX1 mRNA expression levels are given and cell lines below expression levels of 60 units were considered CDX1 low expressing cells; a.u., arbitrary units; n.a., data not available.

Cell Lines	CDX1 mRNA [a.u.]	Tumor	Duke's/grade	Differentiation	Lumen Formation
HCA46	3308.59	sigmoid colon adenocarcinoma	C	poorly	Lumen
HCC56	4115.35	colon adenocarcinoma/liver?	n.a.	moderately	Dense
GP2D	3058.48	colon adenocarcinoma	B	poorly	Intermediate
PCJW	2598.75	colon adenocarcinoma	C	poorly	Intermediate
LS174T	1147.36	colon adenocarcinoma	B/grade I	well	Intermediate
LIM1863	2862.38	colon adenocarcinoma	C/grade III	poorly	Lumen
SW403	2964.14	colon adenocarcinoma	C/grade III	n.a.	Intermediate
RCM1	1152.22	colon adenocarcinoma	n.a.	n.a.	Intermediate
ISRECO1	32.10	colon adenocarcinoma	n.a.	n.a.	n.a.
VACO429	28.69	colon adenocarcinoma	n.a.	n.a.	Dense
HCT116	59.94	colon adenocarcinoma	grade IV	poorly	Dense
CC20	59.51	sigmoid colon adenocarcinoma	B	well	Network/stellate
CAR1	49.35	colon adenocarcinoma	n.a.	n.a.	Dense
COLO678	16.21	colon adenocarcinoma, metastatic lymph node	n.a.	n.a.	Dense
HDC8	29.31	colon adenocarcinoma	C/grade III	n.a.	Intermediate
OXC01	49.69	colon adenocarcinoma	n.a.	n.a.	n.a.

### 2.3. N-glycan Release, Derivatization, and Purification

N-glycans were released in duplicate from three biological replicates per cell line using a PVDF-membrane based release protocol followed by linkage-specific sialic acid derivatization. Purification by cotton-HILIC-SPE and MALDI-TOF-MS analysis was performed as described previously [9]. Shortly, cell pellets were resuspended in water, disrupted by sonication, and proteins immobilized on 96-well PVDF-filter plates (~0.5 × 10E6 cells/well; 5 µL human plasma control; water as blanks) in the presence of chaotropic agents. After washing unbound materials, N-glycans were released overnight at 37 °C. Released glycans were derivatized using ethyl esterification allowing for discrimination of N-acetylneuraminic acid linkages ( $\alpha$ 2,3 vs.  $\alpha$ 2,6) [37] in a ratio of 20:100 (sample: derivatization reagent), and purified by cotton-thread HILIC-SPE [37].

### 2.4. MALDI-TOF-MS Analysis

Released, derivatized, and purified glycans (5 µL sample) were spotted on an anchor chip MALDI target plate (Bruker Daltonics) and co-crystallized with 0.5 µL of 5 mg/mL superDHB in 50% can, supplemented with 1 mM NaOH. Spectra were recorded in positive-ion reflector mode, after calibration with a Bruker peptide calibration kit, using a Bruker UltrafleXtreme™ mass spectrometer, controlled by FlexControl 3.4 software Build 119 (Bruker Daltonics). Mass spectra were obtained over an  $m/z$  range of 1000 to 5000 for a total of 10 000 shots (1000 Hz laser frequency, 200 shots per raster spot during complete random walk). Tandem mass spectrometry (MALDI-TOF-MS/MS) was performed for structural elucidation via fragmentation in gas-off TOF/TOF mode.

### 2.5. Data Processing and Analysis of MALDI-TOF-MS Spectra

Spectra were smoothed (Savitzky Golay algorithm, peak width:  $m/z$  0.06, 4 cycles), baseline corrected (Tophat algorithm), and exported to xy-files using FlexAnalysis 3.4 (Stable Build 76). Mean average spectra were generated per technical replicate, which were summed to one spectrum using the open-source software mMass (<http://www.mmass.org>; [38]). The spectrum was internally re-calibrated using glycan peaks of known composition (H5N2 at  $m/z$  1257.423, H6N2 at  $m/z$  1419.476, H7N2 at  $m/z$  1581.529, H8N2 at  $m/z$  1743.581, H5N4F1 at  $m/z$  1809.639, H5N4F2 at  $m/z$  1955.697, H5N4E1 at  $m/z$  1982.709, H10N2 at  $m/z$  2067.687, H6N5F1 at  $m/z$  2174.7715, H5N4L1E1 at  $m/z$  2255.793, H5N4E2 at  $m/z$  2301.835, H6N5E1 at  $m/z$  2347.8403, H7N6F1 at  $m/z$  2539.904, H6N5F4 at  $m/z$  2612.945, H6N5L1E1 at  $m/z$  2620.925, H7N6E1 at  $m/z$  2712.973, H7N6L2 at  $m/z$  2940.016, H9N8 at  $m/z$  3124.111, H7N6L4F1 at  $m/z$  3632.243) as calibrants (minimum five used), followed by peak picking in mMass, with cut-off signal-to-noise (S/N) 3. The peaklist was manually revised and analyzed in GlycoWorkbench 2.1 stable build 146 (<http://www.eurocarbdb.org/>) using the Glyco-Peakfinder tool

(<http://www.eurocarbdb.org/ms-tools/>) for generation of a glycan compositions list. Our novel in-house software, developed for automated data processing, MassyTools version 0.1.8.0 [39], was used for calibration using a 3<sup>rd</sup> degree function and targeted data extraction of the area under the curve for each individual mass spectrum. To prevent over-estimation of overlapping glycan species, only the first three isotopes were extracted and the area under the curve was corrected based on the theoretical isotopic pattern. The quality of the data was assessed using several quality parameters calculated within the software. Only good quality spectra (total intensity >  $1 \times 10^5$  and fraction of analyte area with  $S/N > 9$  is more than 50%) as well as analytes ( $S/N > 6$ , ppm < 20, quality score > 0.10) were included for analyses. Raw data after pre-processing is provided in the Supplementary Tables. Finally, the corrected area-under-the-curve values were rescaled to a total relative intensity of 100% for each spectrum. Selected glycan compositions were confirmed by MS/MS and a final peak list as well as MS/MS data is given in Supplementary Table S2. MS/MS spectra were manually interpreted and fragment ions annotated using GlycoWorkbench 2.1 according to the nomenclature of Domon and Costello [40]. Averages of direct traits per cell line were used to build a principle component analysis model in SIMCA Version 13.0 (Umetrics AB, Umea, Sweden), with seven random cross-validation (CV) groups.

For increased robustness, derived glycan traits such as galactosylation, fucosylation, sialylation, and others were calculated in SPSS Version 23 (IBM Corp, Armonk, NY). The formulas for calculation are given in Supplementary Table S4 and the average relative abundances are given in Supplementary Table S3. Due to non-normally distributed data, a two-tailed Mann–Whitney test was performed in Rstudio statistical software environment (Version 0.99.892, Kent, OH, USA, <http://www.r-project.org/>) with the significance level  $\alpha = 0.05$  to assess differences in *N*-glycosylation between CDX1 high and low expressing cells. Bonferroni correction was applied to *p*-values to adjust for multiple testing (Supplementary Table S3). Boxplots for visualization were generated in Rstudio and show the median with the interquartile range.

## 2.6. Gene Expression Microarrays and Data Analysis

Total RNA was extracted by using the RNeasy mini kit according to the manufacturer's instructions. Twenty micrograms of RNA of each sample were sent to the Molecular Biology Core Facility of the Paterson Institute for Cancer Research, Manchester, UK, for gene expression microarray analysis using the Human genome U133+2 chips, following the manufacturer's instructions (Affymetrix, High Wycombe, UK). Microarray data were analyzed using Partek Genomics Suite software. The data were log<sub>2</sub>-transformed and RMA-normalized (with GC correction) using quantile normalization with Median Polish for Probeset summarization as optional settings in the software.

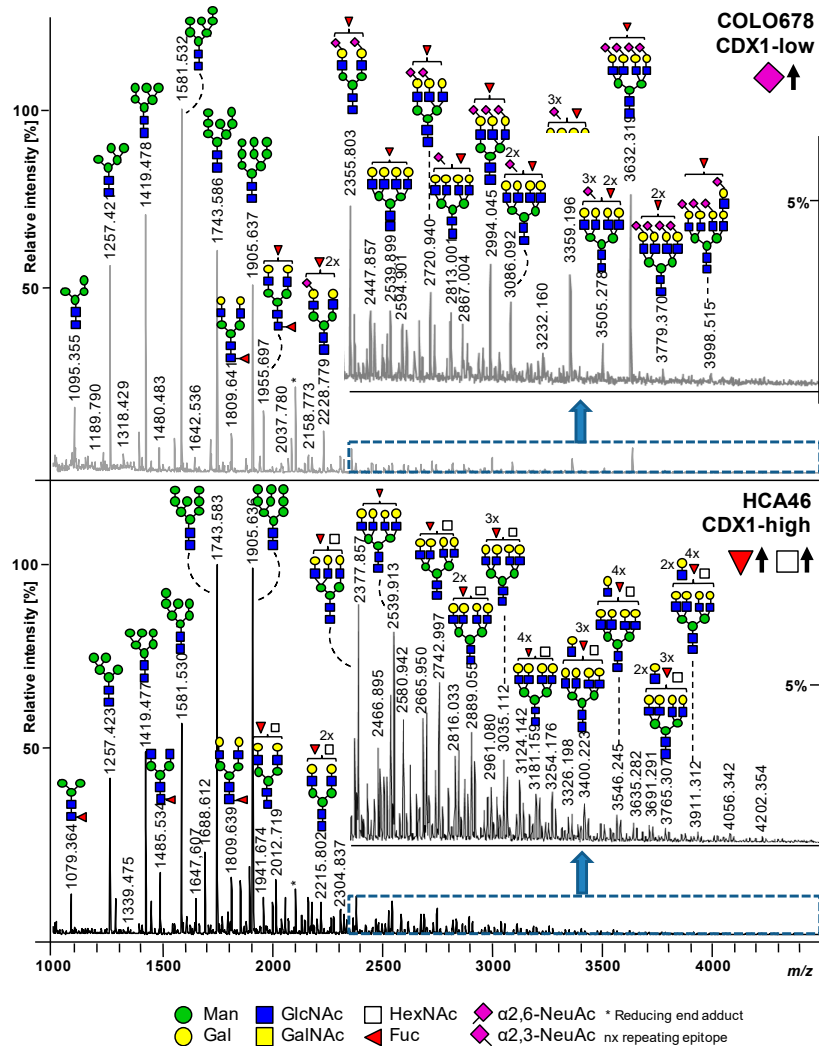
Glycosyltransferase genes (Supplementary Table S5A) as well as glycan-related genes (Supplementary Table S5B) were selected for analysis and differentially expressed genes were identified from a *t*-test comparing mean-expression levels of the 8 CDX1<sup>high</sup> vs. 8 CDX1<sup>low</sup> cell lines. The significant level was adjusted for multiple testing. Fold changes were calculated for CDX1<sup>high</sup> and CDX1<sup>low</sup> cell lines and for significantly different expressed glycosyltransferases, data was visualized as boxplots in Rstudio showing the median and interquartile range. To evaluate the correlation between relative abundances of *N*-glycans traits based on mass spectrometry data and glycosyltransferase gene expression in the 16 investigated cell lines, linear regression analysis was performed in GraphPad Prism Version 6 (GraphPad Software, Inc., La Jolla, CA).

### 3. Results

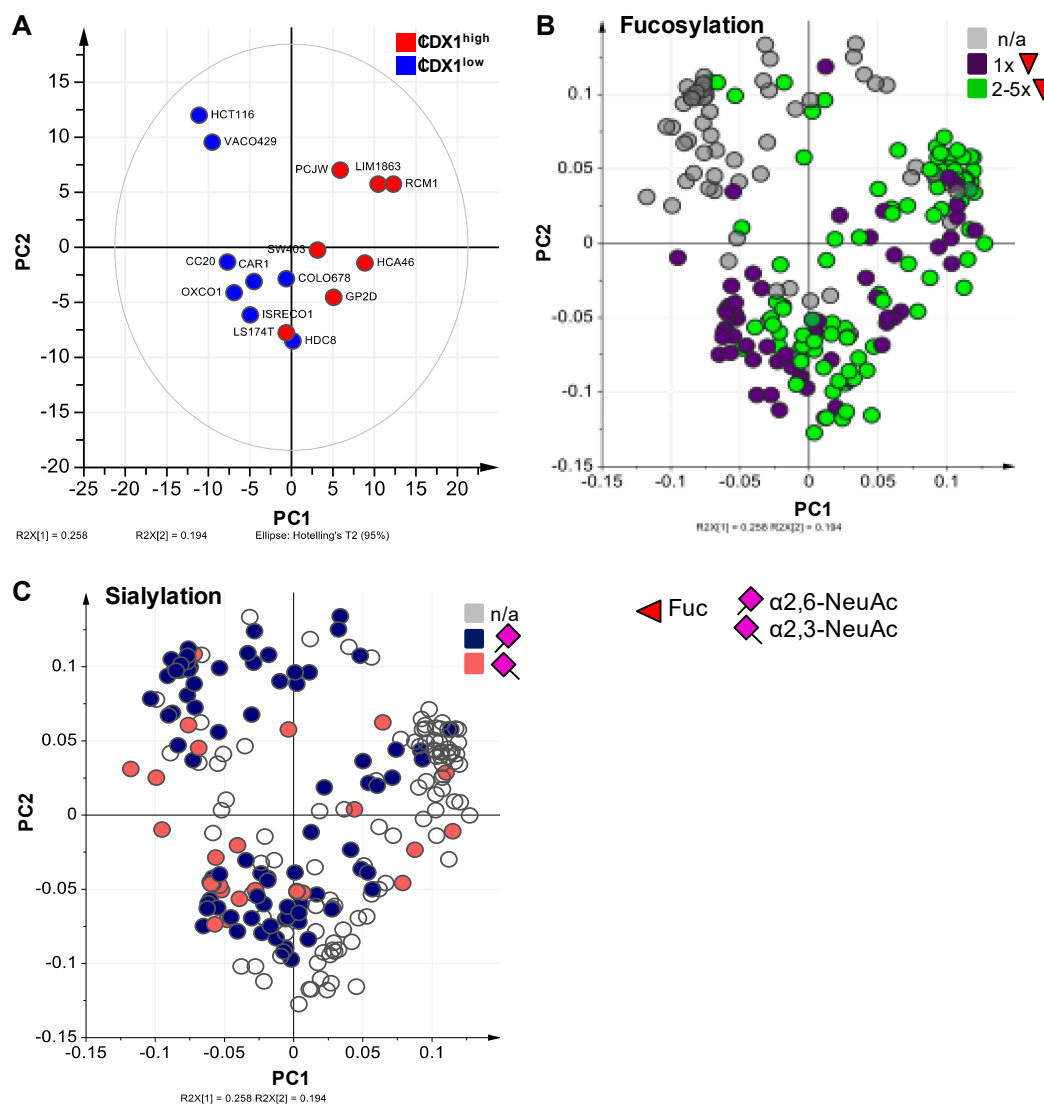
#### 3.1. CDX1<sup>high</sup> and CDX1<sup>low</sup> CRC Cells Exhibit Different N-Glycan Profiles

Our previous data suggested a positive association between the level of fucosylation on protein N-glycans and CDX1 mRNA expression in a set of colorectal cancer cell lines (15 cell lines CDX1/villin positive, 7 cell lines CDX1/villin low or negative) [9]. To validate our previous results and to further explore expression profiles of associated fucosyltransferases, we characterized the N-glycosylation of an independent set of colorectal cancer cell lines with high (8 cell lines; CDX1<sup>high</sup>) vs. low (8 cell lines; CDX1<sup>low</sup>) expression of CDX1 mRNA. The CDX1<sup>high</sup> cell lines investigated here have, on average, 65-fold higher CDX1 mRNA expression as compared to CDX1<sup>low</sup> cell lines (Table 1, Supplementary Table S1). Two exemplary mass spectra of the CDX1<sup>low</sup> cell line Colo678 and the CDX1<sup>high</sup> cell line HCA46 are shown in Figure 2. As observed in other cell line profiling studies, the N-glycomic profiles of all analyzed cell lines were dominated by high-mannose type N-glycans, but differed in the ratios of these glycans. With regard to complex type N-glycans, those derived from Colo678 exhibited more sialic acid residues (N-acetylneuraminic acid, NeuAc, purple diamond, angle indicates different linkages), especially in  $\alpha$ 2,3-linkage, while glycans from HCA46 were characterized by the presence of many fucoses (Fuc, red triangle), low sialylation level, and additional N-acetylhexosamines (HexNAc, white square). In total, 221 individual glycan species were identified across all cell lines, from which 81 could be characterized by tandem mass spectrometry (Supplementary Table S2). In order to evaluate whether CDX1<sup>high</sup> and CDX1<sup>low</sup> expressing cells can be distinguished based on their N-glycomic signature, a principal component analysis (PCA) was performed on the relative abundances of all individual N-glycans. This resulted in a model with four principle components (PC) explaining 68.9% of the variation in the data. The score plot of PC1 vs. PC2 showed clear separation of CDX1<sup>high</sup> and CDX1<sup>low</sup> cell lines along PC1 (Figure 3A) in this unsupervised model, demonstrating a different N-glycan profile between the two groups.





**Figure 2.** Exemplary MALDI-TOF-MS spectra. MALDI-TOF-MS spectrum of CDX1-low expressing cell line COLO678 (upper spectrum), and of CDX1-high expressing cell line HCA46 (lower spectrum). Spectra were recorded in positive ion reflectron mode on a Bruker UltrafleXtreme mass spectrometer. On the y-axis, the relative intensity is given with 100% corresponding to the highest peak in each spectrum. The spectrum range of  $m/z$  2350 to  $m/z$  4600 is enlarged in the inset. Main peaks are annotated with glycan cartoons, representing compositions, and the presence of additional structural isomers cannot be excluded. To simplify the cartoons, repeating epitopes are indicated as “ $nx$ ”. Green circle = mannose, Man; yellow circle = galactose, Gal; blue square = *N*-acetylglucosamine, GlcNAc; white square = *N*-acetylhexosamine, HexNAc; red triangle = fucose, Fuc; purple diamond = sialic acid, *N*-acetylneuraminic acid, NeuAc. Differences in *N*-acetylneuraminic acid linkages are indicated using different angles.

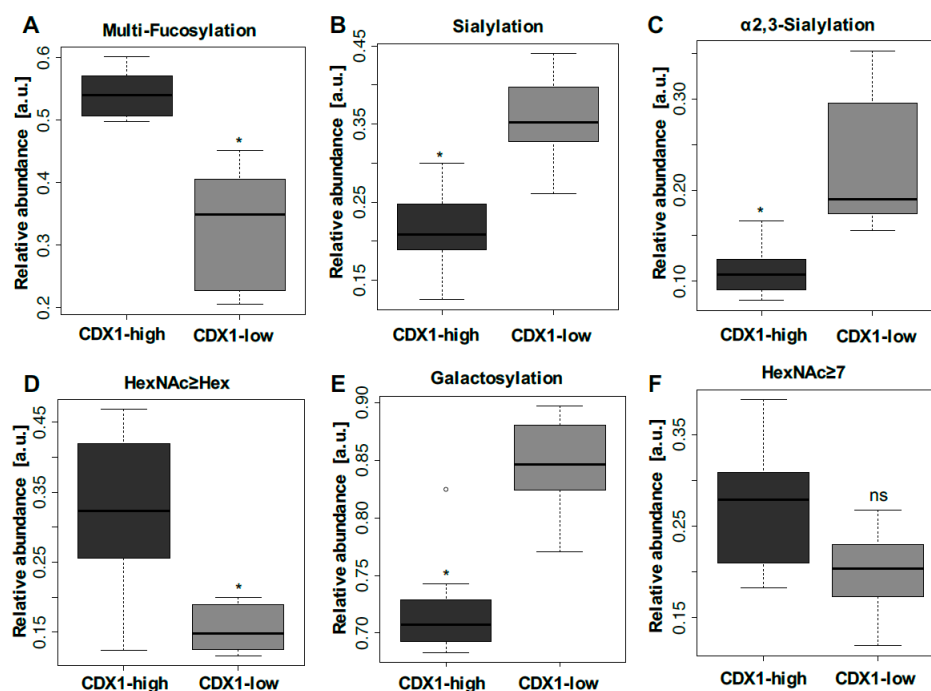


**Figure 3.** Principle component analysis of *N*-glycomic profiles. Principle component analysis (PCA) was performed to explore differences in *N*-glycosylation between CDX1-high and CDX1-low expressing cells. (A) Score plot of principal components (PC) 1 vs. 2 explaining 25.8% and 19.4% of the data, respectively. (B) Corresponding loading plot with color indications for multi- (green) and mono-fucosylated (purple) individual *N*-glycans, and (C) the same loading plot with color indications for  $\alpha$ 2,3-sialylated (rose) and  $\alpha$ 2,6-sialylated (blue) *N*-glycans. Red triangle = fucose, Fuc; purple diamond = sialic acid, *N*-acetylneuraminic acid, NeuAc. Differences in *N*-acetylneuraminic acid linkages are indicated using different angles.

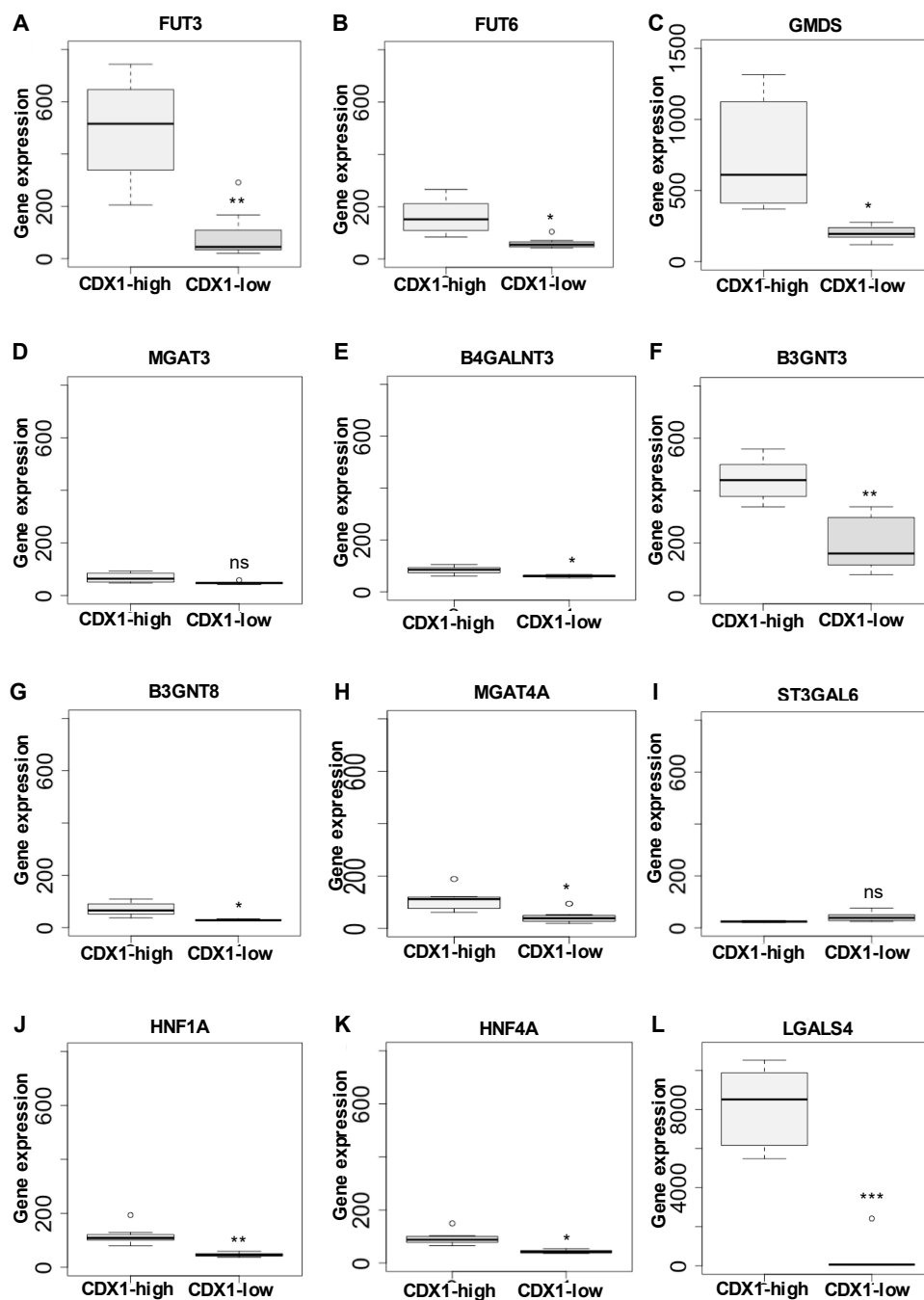
### 3.2. Higher Antenna Fucosylation on *N*-Glycans Characterized CDX1-high Expressing Colorectal Cancer Cell Lines

To assess which glycans drive the principal component separation of CDX1<sup>high</sup> and CDX1<sup>low</sup> colorectal cancer cells, derived glycan traits were calculated by grouping glycan species (direct traits) into classes according to glycosylation characteristics, such as sialylation and fucosylation. The relative abundances of derived traits as well as the calculations are given in Supplementary Tables S3 and S4. Derived traits gave insight into general structural differences and showed increased analytical robustness as compared to individual glycans. Observations from the example spectra were confirmed by comparing derived glycan traits of CDX1<sup>high</sup> and CDX1<sup>low</sup> expressing cell lines and evaluated using a Mann–Whitney test. The *p*-values were adjusted for multiple testing (Bonferroni). Labelling of the PC1 vs. PC2 score plots using glycan-derived traits showed that multi-fucosylated glycans (Figure 3B, colored in green) associated with the location of CDX1<sup>high</sup> cell lines in the score plot, in line with

observations from the two exemplary mass spectra (Figure 2). Accordingly, CDX1<sup>high</sup> expressing cell lines exhibit significantly higher levels of multi-fucosylation (presence of more than one fucose on a glycan), indicative for antenna fucosylation (Lewis X/A, Y/B), as compared to CDX1<sup>low</sup> expressing cell lines ( $\varnothing$  54% vs.  $\varnothing$  33%;  $p$ -value = 0.011; Figure 4A, Supplementary Table S3) and confirmed the association found in our previous study [9]. In order to map the *N*-glycomic phenotype to transcriptomic data, a gene microarray was performed and glycosyltransferase gene expression data was extracted for the 16 investigated cell lines (Supplementary Table S5A). Further, a linear regression analysis was used to evaluate the correlation between MS-based *N*-glycan traits and glycosyltransferase gene expression data (Supplementary Table S6). The trait CFa, representing multi-fucosylation in complex type *N*-glycans, showed significant correlation with fucosyltransferase genes *FUT2*, *3*, *4*, *6*, and *7* which are involved in antenna fucosylation (Supplementary Table S6), thereby indicating that the trait CFa is a good representation for antenna fucosylation in this data. The glycosyltransferase gene expression was also tested for differential expression between CDX1<sup>high</sup> and CDX1<sup>low</sup> cell lines and, in accordance with mass spectrometry data, all fucosyltransferases involved in antenna-fucosylation showed higher expression in CDX1<sup>high</sup> cell lines, though after correction for multiple testing only *FUT3* and *6*, involved in Lewis X/A biosynthesis, showed significantly increased expression (2.7- to 5.7-fold) in the eight CDX1<sup>high</sup> cell lines, compared to eight CDX1<sup>low</sup> cells (Figure 5A,B, Supplementary Table S5A). In contrast, mono-fucosylation, indicative for core-fucosylation (CFc), was lower in CDX1<sup>high</sup> cells as compared to CDX1<sup>low</sup> cells and also *FUT8* gene expression showed a trend towards lower expression in CDX1<sup>high</sup> cells (Supplementary Tables S3 and S5A).



**Figure 4.** Differentially expressed *N*-glycan traits. Derived *N*-glycan traits were calculated to evaluate glycosylation characteristics associated with CDX1 mRNA expression. Differential expressions were observed for (A) multi-fucosylation (more than one fucose) indicative for antenna fucosylation, (B) overall sialylation (C),  $\alpha$ 2,3-sialylation, (D) *N*-glycans with more or equal *N*-acetylhexosamines than hexoses (HexNAc  $\geq$  Hex), (E) galactosylation per antenna, and (F) *N*-glycans with seven or more HexNAcs. Differences were evaluated by Mann–Whitney test for derived traits and significances ( $p$ -value < 0.05) after multiple testing corrections are indicated (\*), ns = non-significant. Boxplots consistently indicate the median and interquartile range.



**Figure 5.** Differentially expressed genes related to glycosylation. *N*-glycomic phenotypes of the eight CDX1-high vs. eight CDX1-low cell lines were further compared to transcriptomic data of the same cell lines obtained from a gene expression microarray analysis using the Human genome U133+2 Affimetrix chips. The Y axis shows the linear normalized fluorescent intensities of the respective probesets. Differential expressions were observed for fucosyltransferase genes (A) *FUT3*, (B) *FUT6*, and (C) *GMDS*, for (D) *MGAT3* (bisecting GlcNAc) and (E) *B4GALNT3* (LacdiNAc), as well as (F) *B3GNT3* (poly-LacNAc), (G) *B3GNT8* (poly-LacNAc), (H) *MGAT4A* ( $\alpha$ 1-3-branching), and (I)  $\alpha$ 2,3-sialyltransferase gene *ST3GAL6*. Furthermore, transcription factors (J) hepatocyte nuclear factor (HNF)1A, and (K) HNF4A, as well as (L) soluble galectin 4 (*LGALS4*) showed increased expression with high CDX1 expression in the set of 16 colorectal cancer cell lines. Differences were evaluated by a *t*-test. Boxplots indicate the median and interquartile range. Significances after multiple testing correction are indicated and correspond to \* *p*-value < 0.05, \*\* *p*-value < 0.01, \*\*\* *p*-value < 0.001; ns non-significant.

Furthermore, the GDP-L-fucose precursor GDP-mannose 4,6-dehydratase (*GMD5* gene) gene expression was 3.7-fold ( $p$ -value  $1.4 \times 10E-03$ ) elevated with high CDX1 expression (Figure 5C; Supplementary Table S5B). The corresponding enzyme GDP-mannose 4,6-dehydratase is involved in the fucosylation process, indicating that various enzymes involved in fucosylation are differentially regulated in CDX1<sup>high</sup> versus CDX1<sup>low</sup> cells. As the transcription factors *HNF1A* and *HNF4A* have previously been shown to regulate antenna fucosylation in plasma [34] and have previously been associated with CDX1 expression as well as intestinal development [41–43], here we tested the expression levels in the investigated CDX1<sup>high</sup> and CDX1<sup>low</sup> cell lines. Interestingly, gene microarray data revealed significantly increased expression of *HNF1A* and *HNF4A* with high CDX1 mRNA levels ( $p$ -value (*HNF1A*) = 0.002;  $p$ -value (*HNF4A*) = 0.003; Figure 5J–K, Supplementary Table S5B). Moreover, soluble galectin 4 (*LGALS4*), a target gene of *HNF4A* and a glycan-binding protein, was 23-fold up-regulated in CDX1<sup>high</sup> cells ( $p$ -value =  $2.4 \times 10E-06$ ; Figure 5L, Supplementary Table S5B). Galectin 4 is highly expressed in the alimentary tract during the development and is associated with differentiation [44].

### 3.3. CDX1<sup>high</sup> Expressing CRC Cell Lines Exhibit a Lower Level of Sialylated N-Glycans as Compared to CDX1-Low Expressing Cells

In our previous study, we observed a trend towards a negative association between overall N-glycan sialylation and CDX1 mRNA expression, though with no significant difference after correction for multiple testing. To further explore this association, we labelled the score plot of PC1 vs. PC2 with derived sialylation traits and found that sialylated glycans largely marked CDX1<sup>low</sup> cell lines (Figure 3C, colored in blue and red). In agreement, overall sialylation was significantly lower in CDX1<sup>high</sup> cell lines compared to CDX1<sup>low</sup> expressing cell lines ( $\emptyset$  36% vs.  $\emptyset$  21%;  $p$ -value = 0.046; Figure 4B, Supplementary Table S3). Results from the current study showed pronounced sialylation differences for  $\alpha$ 2,3-sialylated N-glycans with  $\emptyset$  11% in CDX1<sup>high</sup> vs.  $\emptyset$  23% in CDX1<sup>low</sup> cell lines ( $p$ -value = 0.023; Figure 4C; Supplementary Table S3). Different sialyltransferases are involved in this sialylation and mRNA levels of sialyltransferases *ST3GAL3* and *6* correlated with relative abundances of overall sialylation in complex type N-glycans (Supplementary Table S6). In line with results from MS data, *ST3GAL3,4* and, especially, *ST3GAL6*, all three involved in sialyl Lewis antigen biosynthesis, were decreased in CDX1<sup>high</sup> cell lines as compared to CDX1<sup>low</sup> cell lines, though not significantly (Figure 5I, Supplementary Table S5A). Notably, sialylation per galactose in the MS data was likewise decreased from  $\emptyset$  48% (CDX1<sup>low</sup>) to  $\emptyset$  31% in CDX1<sup>high</sup> cell lines ( $p$ -value = 0.023; Supplementary Table S3), suggesting the decrease in sialylation being an independent event and not a mere result of substrate limitation through decreased galactosylation (see below).

### 3.4. CDX1 Expression in CRC Cell Lines Associated with N-Glycans CARRYING Additional N-Acetylhexosamine

While differences in levels of fucosylation and sialylation mainly explain the separation of CDX1<sup>high</sup> and CDX1<sup>low</sup> CRC cells along PC1, we sought to determine whether other glycan-derived traits could account for the separation in other PC dimensions. Glycan structures with the number of HexNAc equaling or exceeding the number of hexoses (Hex; HexNAc $\geq$ Hex) represent glycans with bisecting N-acetylglucosamine (GlcNAc), non-galactosylated antennae, or the addition of N-acetylgalactosamine (GalNAc) and showed a significantly increased expression in CDX1-positive cells in our previous study [9]. This could be validated in the new set of cells, in which CDX1<sup>high</sup> cell lines showed a more than 2-fold increase in the relative abundance of glycans with the feature HexNAc $\geq$ Hex, compared to CDX1<sup>low</sup> cells ( $p$ -value = 0.023; Figure 4D; Supplementary Table S3). Corresponding gene expression of the glycosyltransferase involved in the formation of bisecting GlcNAc (*MGAT3*) was significantly correlated with MS data (Supplementary Table S6), suggesting the presence of bisection and refining the data obtained by mass spectrometry. However, *MGAT3* gene

expression showed only a trend towards increased expression in CDX1<sup>high</sup> cells compared to CDX1<sup>low</sup> cells (Figure 5D, Supplementary Table S5A).

Other glycan epitopes potentially contributing to the HexNAc $\geq$ Hex class are the LacdiNAc (GalNAc $\beta$ 1–4GlcNAc) epitope, the SdA antigen (NeuAc $\alpha$ 2–3(GalNAc $\beta$ 1–4)Gal $\beta$ 1–4GlcNAc; NeuAc=N-acetylneuraminic acid, Gal=galactose), and the blood group A epitope (GalNAc $\alpha$ 1–3(Fuc $\alpha$ 1–2)Gal $\beta$ 1–3/4GlcNAc). Accordingly, one of the glycosyltransferases adding a GalNAc-residue to a GlcNAc to form LacdiNAc structures on N-glycans, *B4GALNT3*, was significantly correlated with the MS-based trait HexNAc $\geq$ Hex (Supplementary Table S6) and was, additionally, significantly higher expressed in CDX1<sup>high</sup> versus CDX1<sup>low</sup> cells (Figure 5E, Supplementary Table S5A). Other genes involved in the expression of these epitopes were not significantly different (Supplementary Table S5A).

While glycans with additional HexNAcs were elevated, galactosylation per antenna was found to be significantly lower in CDX1<sup>high</sup> cells than in CDX1<sup>low</sup> cell lines ( $\emptyset$  85% vs.  $\emptyset$  72%; *p-value* = 0.046; Figure 4E, Supplementary Table S3). Corresponding galactosyltransferases showed a corresponding trend (Supplementary Tables S5A).

### 3.5. CDX1 Expression in CRC Cell Lines Associated with Higher Branched N-Glycan-Derived Traits

Finally, N-glycan structures with seven or more HexNAcs, indicative for branched structures or (poly-) LacNAc repeats (–Gal $\beta$ 1–4GlcNAc $\beta$ 1–3Gal $\beta$ 1–3/4–GlcNAc–), showed a trend towards higher expression with high CDX1 expression in our previous data and were higher expressed in CDX1<sup>high</sup> cell lines of the new set of cells as compared to CDX1<sup>low</sup> expressing cell lines, with  $\emptyset$  27% vs.  $\emptyset$  20% (Figure 4F, Supplementary Table S3), though not significantly after multiple testing correction. Since our MS data could not sufficiently differentiate between LacNAc-repeat and additional antenna, we next analyzed the glycosyltransferase expression data to see if this would give a more detailed insight. Genes encoding for beta-1,3-N-acetylglucosaminyltransferase 3 (*B3GNT3*) and *B3GNT8*, both involved in the synthesis of type-1 chains and poly-LacNAc repeats, were around 2-fold up-regulated with high CDX1 expression (Figure 5F+G, Supplementary Table S5A) and *B3GNT3* gene expression showed significant correlation with the MS trait HexNAc $\geq$ 7 (Supplementary Table S6), suggesting the presence of LacNAc-repeat structures, to a certain degree. Additionally, the glycosyltransferase encoded by gene *MGAT4A*, which is involved in the branching on the 1,3-arm of N-glycans to form tri- and tetra-antennary N-glycans, was correlated with the relative abundance of the MS N-glycan trait HexNAc $\geq$ 7 (Supplementary Table S6). *MGAT4A* was further significantly higher expressed in CDX1<sup>high</sup> cells as compared to CDX1<sup>low</sup> cells (Figure 5H, Supplementary Table S5A). Of note, N-glycans featuring HexNAc $\geq$ Hex also contribute to this trait of N-glycan structures with seven or more HexNAcs, which is therefore not solely indicative for branching and poly-LacNAc repeats. In line, *B3GNT3* and *B3GNT8* also showed correlation with the MS trait HexNAc $\geq$ Hex (Supplementary Table S6).

Overall, we could validate our previous results and identify specific N-glycan features associated with high CDX1 mRNA expression, which were characterized by high multi-fucosylation, elevated levels of N-glycans containing additional HexNAcs as well as low galactosylation and low sialylation levels, with particularly decreased levels of  $\alpha$ 2,3-sialylation.

## 4. Discussion

In two independent sets of colorectal cancer cell lines we observed increased multi-fucosylation in CDX1<sup>high</sup> expressing cell lines, next to increased terminal HexNAc epitopes as well as decreased galactosylation and sialylation. Changes in glycosylation have mainly been attributed to alterations in the corresponding glycosyltransferases, as also described in literature [45,46]. Although several glycosyltransferase gene expressions correlated well in the study presented here, predictions on the glycan phenotype using gene expression data remain challenging since several biological effectors can

influence not only the expression of glycan-initiating, -elongating, and -degrading enzymes, but also the enzyme activity [47].

Guo and Pierce, for example, recently reviewed the involvement of transcription factors in glycan expression and reported that they regulate the expression of glycosyltransferase genes in a tissue- and cell-specific manner [47]. Interestingly, we found CDX1, a colon-specific transcription factor involved in cell differentiation, associated with different glycosylation features. Cell lines expressing high mRNA levels of CDX1 showed an *N*-glycan phenotype with increased multi-fucosylation, indicative for antenna-fucosylation. Supporting this association, fucosyltransferase genes *FUT3* and *6*, and *GMDS*, which is involved in GDP-L-fucose synthesis, were positively associated with high CDX1 expression. Differences were more pronounced for *FUT3*, suggesting an enhanced expression of the type-1 chain epitope Lewis A in CDX1<sup>high</sup> cells [48].

Our data suggests antenna fucosylation on *N*-glycans as a potential marker for epithelial differentiation in colorectal cancer cell lines. In line with our results here, a previous study using HCT116, which is characterized by a poor differentiation status and has a mutation in the *GMDS* gene, leading to low levels of fucosylation, showed that restoration of *GMDS* wild type expression enhanced fucosylation, suppressed tumor formation and reduced the metastatic potential, when injected into mice [20]. Additionally, HCT116 cells have shown higher levels of cancer stem cells (CSC) and loss of CDX1 expression, whereas induced expression of CDX1 led to reduced clonogenicity and restored the potential of cells for differentiation and lumen formation [49]. Supporting the hypothesis of fucosylation as a characteristic of epithelial cells, Breiman et al. identified fucosylated antigens in the mammary cancer cell line as markers of the epithelial state, which can contribute to cell adhesion through CLEC17A (Prolectin) [50]. During the epithelial-to-mesenchymal transition (EMT) of these cells, the expression of fucosylated antigens such as Lewis Y was decreased as a result of decreased expression of fucosyltransferases encoded by *FUT1* and *FUT3* genes [50].

The transcription factors *HNF1A* and *HNF4A* were also positively associated with high CDX1 mRNA expression in the examined cell lines. Lauc et al. identified in a genome-wide association study (GWAS) that fucosyltransferases *FUT3*, *5*, and *6*, all involved in antenna fucosylation, are positively regulated by *HNF1A* and its downstream factor *HNF4A*, while *FUT8*, initiating core-fucosylation, is inhibited [34]. The study showed that knock-down of *HNF1A* and *HNF4A* resulted in reduced expression of several fucosyltransferase genes involved in antenna fucosylation in HepG2 liver cells and, partly, in pancreatic Panc1 cells, while *FUT8*, responsible for *N*-glycan core-fucosylation, was upregulated. Furthermore, *GMDS* and *L-Fucokinase*, both involved in two different pathways of fucose synthesis, were drastically down-regulated upon *HNF1A* or *HNF4A* knock-down. Publicly available ChIP-seq data of ENCODE show that *HNF4A* and *HNF4G* transcription factors bind to *FUT2*, *3*, and *6* genes in cell lines of non-intestinal origin and, also, Lauc et al. confirmed the binding of *HNF1A* and *HNF4A* to multiple *FUT* genes and/or their promoters by ChiP analysis [34]. At the same time, several groups reported on the interaction between CDX1 as well as the highly related, partially redundant, CDX2 with other transcription factors, including *HNF1A* and *HNF4A*. Boyd et al. identified CDX2 binding sites on CDX1 and *HNF1B*, both being positively regulated by CDX2, while *HNF1B* is needed for activation of *HNF4A*, *HNF1A*, and *HNF3G* [43]. Direct binding of CDX2 to *HNF4A* was also shown to positively influence transcriptional activity [43]. *HNF4A* was identified as a transcriptional activator for intestinal differentiation and gene expression of *HNF4A* and its target genes were upregulated in colorectal cancer cell lines with a more epithelial phenotype, as compared to cells with a mesenchymal phenotype [51]. One target gene of *HNF4A* is galectin 4, which was described as tumor-suppressor in colon cancer [52,53], but also pancreatic duct adenocarcinoma [54], and was accordingly positively associated with high CDX1 mRNA expression in the current study.

Taking our findings and the reports from literature, one may speculate that antenna fucosylation of *N*-glycans is a marker for the epithelial state and that CDX1 is involved in the transcriptional

regulation of fucosyltransferases involved in (antenna-)fucosylation, leading to a more differentiated and less invasive phenotype in colorectal cancer cell lines.

Interestingly, in our investigated cell lines, CDX2 mRNA expression correlated with CDX1 mRNA expression (both high or both low), with the exception of the cell line RCM1. Consequently, similar observations as for CDX1-associated glycosylation were made for CDX2, but correlations with the *N*-glycan phenotype and glycosyltransferase expressions were less pronounced for CDX2 as compared to CDX1 (data not shown). In contrast, a direct involvement of CDX2 in the regulation of *FUT2*, which revealed potential binding sites for CDX1 and CDX2 and is involved in the generation of LeY/B antigens, was shown in colon cancer cell lines HT29 and DLD-1 [55]. Expression of CDX2, and thereby *FUT2*, could be reduced through treatment with epidermal growth factor (EGF)/bFGF [55]. On the other hand, the sialyl Lewis antigen promoting glycosyltransferases *ST3GAL1/3/4* and *FUT3* were transcriptionally up-regulated by c-Myc [55]. While sialyl Lewis types antigens are commonly associated with (colorectal) cancers [8], the CDX1<sup>high</sup> cell lines show a very low expression of  $\alpha$ 2,3-sialylation, thereby limiting the substrate for specific fucosyltransferases involved in the expression of sialyl Lewis epitopes (*FUT3,5,6,7*). The combined expression of  $\alpha$ 2,3-sialylation and antenna fucosylation, as reflected in the trait CLFa, may be considered a proxy for sialyl Lewis epitope expression. Both groups of cell lines in the present study only showed low levels of CLFa (~4%), indicating low levels of sialyl Lewis epitopes. In the case of CDX1<sup>low</sup> expressing cell lines, the combination of  $\alpha$ 2,3-sialylation with overall fucosylation increased (minimum of one fucose which may be core or antenna attached;  $\varnothing$  16%), which may likewise reflect increased sialyl Lewis epitope levels. Strikingly, in CDX1<sup>high</sup> expressing cell lines the expression of multi-fucosylation (indicative for antenna fucosylation) and  $\alpha$ 2,3-sialylation of *N*-glycans even showed opposite trends. Though *FUT6* seems to have a preference for the sialylated substrate, here, the results point towards the expression of mainly non-sialylated Lewis antigens via action of fucosyltransferases *FUT3* (Lewis A) and *FUT4* or *FUT6* (Lewis X) [48] in the CDX1<sup>high</sup> colorectal cancer cell lines analyzed here.

Especially in the context of sialyl Lewis epitopes, the importance to distinguish  $\alpha$ 2,3- and  $\alpha$ 2,6-sialylation becomes evident. The derivatization applied here differentially modifies the sialic acids in different linkages resulting in a detectable mass shift allowing the distinction between  $\alpha$ 2,3- and  $\alpha$ 2,6-sialylation [37]. Similar to  $\alpha$ 2,3-linked sialic acid,  $\alpha$ 2,8-linked sialic acid will be lactonized under the conditions of the derivatization step [56]. We therefore cannot differentiate between  $\alpha$ 2,3-linked sialic acid and  $\alpha$ 2,8-linked sialic on the basis of the observed masses alone. However, we performed in-depth MS/MS characterization of *N*-glycans from different colorectal cancer cells and could not find evidence for a fragment corresponding to an antenna with two or more sialic acids [9]. Although *N*-glycans with terminal  $\alpha$ 2,8-linked sialic acids are expressed in cancer cells, these structures might be expressed at low levels as compared to *N*-glycans with terminal  $\alpha$ 2,6- and  $\alpha$ 2,3-linked sialic acids. Moreover, gene expression of *ST8SIA1-5* was not significantly different between the here studied CDX1 high and CDX1 low expressing colorectal cell lines. Sialic acids can further be modified by O-acetyl-groups modulating the ligand function. O-acetylated glycans have been reported to be mainly present in the lower part of the intestinal tract [57]. Furthermore, a decrease of O-acetylation has been observed with colon cancer progression [58]. In line, we previously detected O-acetylated sialic acids of glycosphingolipid glycans, which were decreased in colorectal cancer tissues as compared to control tissues [59], but not on *N*-glycans of the same tissues [1]. The derivatization method applied in this study to characterize the *N*-glycans of colorectal cancer cell lines has been shown to preserve O-acetylation [60]. However, we could not find evidence of O-acetylated *N*-glycans in the investigated cell lines.

We further observed a decrease in galactosylation with high CDX1 mRNA expression. Previous reports showed that *CDX1*, *CDX2*, *HNF1A*, and *HNF1B* are involved in the transcriptional regulation of *B3GALT5*, the gene encoding for the enzyme responsible for type-1 chain (-Gal $\beta$ 1-3GlcNAc $\beta$ -) expression on glycolipids and glycoproteins, with preferences for *O*-glycans and



glycosphingolipid-glycans [61]. *B3GALT5* was found to be down-regulated in colon cancer cell lines, but up-regulated upon CaCo2 cell differentiation [61]. However, our mass spectrometric approach does not allow distinction between type-1 and type-2 chains (-Gal $\beta$ 1-4GlcNAc-), but enzyme levels showed an up-regulation of type-1 chain glycosyltransferases *B3GALTs* as well as *B3GNTs* according to observations by Isshiki et al. [61], while type-2 chain enzymes were decreased with high CDX1 expressions. Furthermore, many glycosyltransferases involved in glycan elongations act on *N*-glycans, *O*-glycans, glycosphingolipid-glycans, and/or others and changes on different glycan classes can be differently regulated. Therefore, it is striking that the *N*-glycomic data obtained by mass spectrometry are well in accordance with the transcriptomic data, suggesting that observed changes in, for example, fucosylation and sialylation are mainly attributed to *N*-glycans or are globally altered, but further investigations on *O*-glycans and glycosphingolipid-glycans are needed.

Our results further revealed the presence of terminal HexNAc residues to be increased in CDX1<sup>high</sup> expressing cells. Glycan motifs that may contribute to this increase include LacdiNAc structures as well as the Sda antigen; the latter was shown to be expressed in normal colon tissues and decreased during colon cancer progression [62]. LacdiNAc, on the other hand, was associated with differentiation of mammary epithelial cells and tumor suppression in neuroblastoma, whereas its expression was increased in human prostate, ovarian, and pancreatic cancers [63]. Furthermore, bisecting GlcNAc containing glycans contribute to this group and corresponding *MGAT3* gene expression was up-regulated with CDX1<sup>high</sup> cells. Several reports describe reduced bisection in several tumors [64] and it was shown that it suppresses *N*-glycan branching by *MGAT5* [65], as well as  $\alpha$ 2,3-sialylation [66], the latter being reduced in CDX1<sup>high</sup> cells.

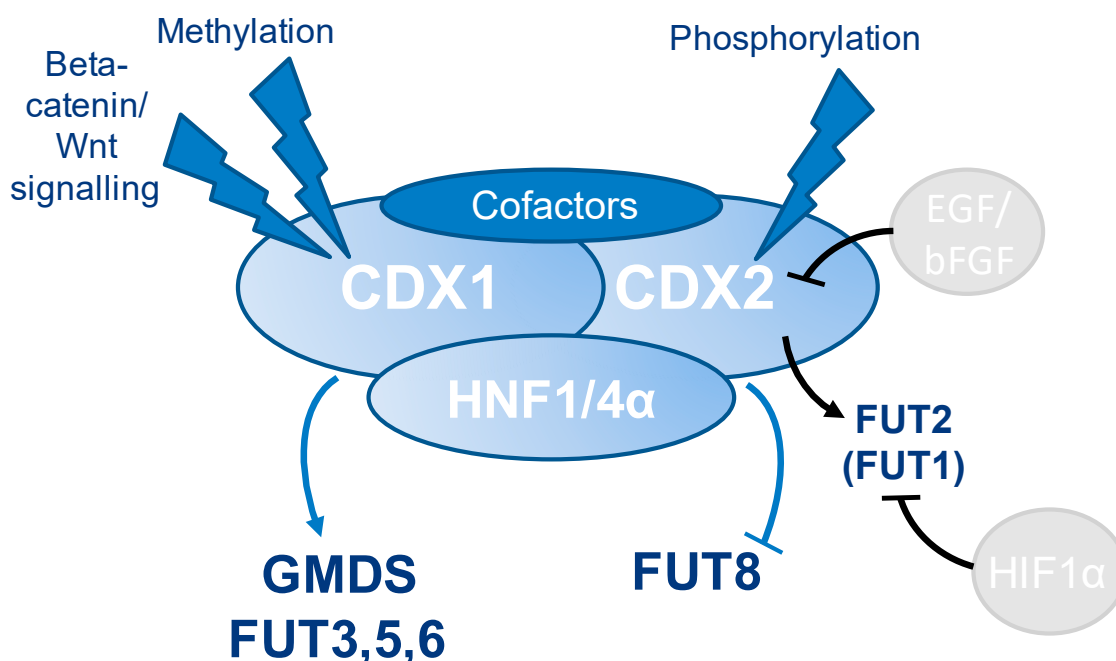
Strikingly, observed glycan phenotypes for CDX1<sup>low</sup> cells matched largely those described for cells undergoing EMT, which involves the loss of epithelial markers such as E-cadherin and gain of mesenchymal markers such as vimentin. EMT-associated glycan changes have previously been described and include the aforementioned loss of antenna-fucosylation, but also decrease in antennarity of *N*-glycans and bisection, whereas enhanced levels of high-mannose type glycans, core-fucosylation, and corresponding *FUT8* expression, as well as increased  $\alpha$ 2,6-sialylation and *ST6GAL1* gene expression, were observed [67,68]. Inhibition of *FUT3* and *FUT6* has been shown to affect TGF- $\beta$  receptor glycosylation, resulting in decreased fucosylation as well as *FUT3/6*-associated sialyl Lewis antigens and altered TGF- $\beta$ -mediated EMT and invasion in colorectal cancer cells [69]. Furthermore, loss of CDX 1 and/or CDX2 was shown to impact TGF beta signaling and tumor invasion in murine APC mutant colon cancer models. Upon loss of CDX1/2, cells were poorly differentiated, invasive, and developed a villous morphology, which was accompanied by the loss of the epithelial marker E-cadherin, whereas expression of vimentin, *Twist1*, *Zeb1*, and *Zeb2* was induced [70]. Finally, *HNF1A* and *HNF4A* have been described to prevent EMT in liver cancer [71–73]. The role of glycosylation in EMT and a potential involvement of CDX1 clearly needs further investigation.

While the involvement of transcription factors was shown for fucosyltransferases, reports on sialyltransferases and galactosyltransferases involved in the elongation of *N*-glycans are still lacking and more research on the integrated regulation, as well as competition of glycosyltransferases, is needed. Also, the role of LacdiNAc structures and other terminal HexNAc epitopes in differentiation and colorectal cancer needs further investigation.

## 5. Conclusions

Our data, in combination with reports from literature, suggest that CDX1 (and CDX2) are involved in the regulation of multiple glycosyltransferases, especially fucosyltransferases, likely via interactions with other transcription factors, such as *HNF4A* and *HNF1A*. However, CDX genes may (additionally) influence the expression of fucose-carrying glycoproteins themselves, thereby leading to an *N*-glycan phenotype with enhanced multi-fucosylation. Taking into account the interaction of the two CDX-proteins with *HNF1A*, *HNF4A*, and *HNF1B* and together with the proven role of *HNF1A* and *HNF4A* in the regulation of fucosylation, we hypothesize a cooperation of these transcription factors

being involved in the expression of *FUT* genes and thereby increasing fucosylation of glycoproteins with high CDX1/CDX2 expression (summarized in Figure 6). Certainly, more mechanistic studies are needed to elucidate the role of CDX1 in glycosyltransferase regulations.



**Figure 6.** Model for regulation of fucosylation. Based on our findings as well as reports from literature (Lauc, G. et al. (2010) *PLoS Genet* **6**; Guo, R. J., Suh, E. R. & Lynch, J. P. (2004) *Cancer Biol Ther* **3**, 593-601) we hypothesize that antenna-fucosylation is regulated by an interplay of several transcription factors, especially CDX1, CDX2, HNF1A, and HNF4A, which can be affected by several signaling pathways, promoter methylation, as well as treatment with epidermal growth factor (EGF)/bFGF (Sakuma, K., Aoki, M. & Kannagi, R. (2012) *Proceedings of the National Academy of Sciences of the United States of America* **109**, 7776-7781) or under hypoxia (Belo, A. I. et al. (2015) *FEBS Lett* **589**, 2359-2366).

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4409/8/3/273/s1>: Supplementary Tables S1–S6.

**Author Contributions:** Conceptualization, S.H., J.L.W., Y.R., and M.W.; Methodology, S.H. and K.K.; Software, S.H. and J.L.W.; Validation, S.H. and J.L.W.; Formal Analysis, S.H. and J.L.W.; Investigation, S.H. and J.L.W.; Resources, S.H. and J.L.W.; Data Curation, S.H. and Y.W.; Writing—Original Draft Preparation, S.H.; Writing—Review & Editing, J.L.W., K.K., Y.R., and M.W.; Visualization, S.H.; Supervision, Y.R. and M.W.; Funding Acquisition, M.W.

**Funding:** This project was supported by the European Union (Seventh Framework Programme HighGlycan project, grant number: 278535).

**Acknowledgments:** We thank Sir Walter Bodmer (Oxford University) for his support and for providing the cell lines used in this study, Gabi W. van Pelt (LUMC) for support with cell counting, Bas C. Jansen (LUMC) for help with data analysis tools, and Oleg Mayboroda (LUMC) with visualizations using R studio.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Schnaar, R.L. Glycobiology simplified: Diverse roles of glycan recognition in inflammation. *J. Leukoc. Biol.* **2016**, *99*, 825–838. [[CrossRef](#)]
2. Cummings, R.D.; Pierce, J.M. The Challenge and Promise of Glycomics. *Chem. Biol.* **2014**, *21*, 1–15. [[CrossRef](#)] [[PubMed](#)]

3. Taniguchi, N.; Kizuka, Y. Glycans and cancer: Role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. *Adv. Cancer Res.* **2015**, *126*, 11–51.
4. Stanley, P.; Schachter, H.; Taniguchi, N. N-Glycans. In *Essentials of Glycobiology*, 2nd ed.; Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., Etzler, M.E., Eds.; CSHL Press: Cold Spring Harbor, NY, USA, 2009.
5. Pinho, S.S.; Reis, C.A. Glycosylation in cancer: Mechanisms and clinical implications. *Nat. Rev. Cancer* **2015**, *15*, 540–555. [[CrossRef](#)] [[PubMed](#)]
6. de-Freitas-Junior, J.C.; Morgado-Diaz, J.A. The role of N-glycans in colorectal cancer progression: Potential biomarkers and therapeutic applications. *Oncotarget* **2015**, *12*, 19395–19413.
7. Glavey, S.V.; Huynh, D.; Reagan, M.R.; Manier, S.; Moschetta, M.; Kawano, Y.; Roccaro, A.M.; Ghobrial, I.M.; Joshi, L.; O'Dwyer, M.E. The cancer glycome: Carbohydrates as mediators of metastasis. *Blood Rev.* **2015**, *29*, 269–279. [[CrossRef](#)] [[PubMed](#)]
8. Holst, S.; Wuhrer, M.; Rombouts, Y. Chapter Six—Glycosylation Characteristics of Colorectal Cancer. In *Advances in Cancer Research*; Richard, R.D., Lauren, E.B., Eds.; Academic Press: London, UK, 2015; Volume 126, pp. 203–256.
9. Holst, S.; Deuss, A.J.; van Pelt, G.W.; van Vliet, S.J.; Garcia-Vallejo, J.J.; Koeleman, C.A.; Deelder, A.M.; Mesker, W.E.; Tollenaar, R.A.; Rombouts, Y.; et al. N-glycosylation Profiling of Colorectal Cancer Cell Lines Reveals Association of Fucosylation with Differentiation and Caudal Type Homeobox 1 (CDX1)/Villin mRNA Expression. *Mol. Cell. Proteomics* **2016**, *15*, 124–140. [[CrossRef](#)] [[PubMed](#)]
10. Dotz, V.; Wuhrer, M. Histo-blood group glycans in the context of personalized medicine. *Biochim. Biophys. Acta* **2016**, *1860*, 1596–1607. [[CrossRef](#)] [[PubMed](#)]
11. Le Pendu, J.; Marionneau, S.; Cailleau-Thomas, A.; Rocher, J.; Le Moullac-Vaidye, B.; Clement, M. ABH and Lewis histo-blood group antigens in cancer. *APMIS* **2001**, *109*, 9–31. [[CrossRef](#)] [[PubMed](#)]
12. Mollicone, R.; Cailleau, A.; Oriol, R. Molecular genetics of H, Se, Lewis and other fucosyltransferase genes. *Transfus. Clin. Biol.* **1995**, *2*, 235–242. [[CrossRef](#)]
13. Le Pendu, J. Histo-blood group antigen and human milk oligosaccharides: Genetic polymorphism and risk of infectious diseases. *Adv. Exp. Med. Biol.* **2004**, *554*, 135–143. [[PubMed](#)]
14. Becker, D.J.; Lowe, J.B. Fucose: Biosynthesis and biological function in mammals. *Glycobiology* **2003**, *13*, 41R–53R. [[CrossRef](#)] [[PubMed](#)]
15. Vajaria, B.N.; Patel, P.S. Glycosylation: A hallmark of cancer? *Glycoconj. J.* **2017**, *34*, 147–156. [[CrossRef](#)] [[PubMed](#)]
16. Miyoshi, E.; Moriwaki, K.; Nakagawa, T. Biological Function of Fucosylation in Cancer Biology. *J. Biochem.* **2008**, *143*, 725–729. [[CrossRef](#)] [[PubMed](#)]
17. Moriwaki, K.; Miyoshi, E. Fucosylation and gastrointestinal cancer. *World J. Hepatol.* **2010**, *2*, 151–161. [[CrossRef](#)] [[PubMed](#)]
18. Sethi, M.K.; Hancock, W.S.; Fanayan, S. Identifying N-Glycan Biomarkers in Colorectal Cancer by Mass Spectrometry. *Acc. Chem. Res.* **2016**, *49*, 2099–2106. [[CrossRef](#)] [[PubMed](#)]
19. Kaprio, T.; Satomaa, T.; Heiskanen, A.; Hokke, C.H.; Deelder, A.M.; Mustonen, H.; Hagstrom, J.; Carpen, O.; Saarinen, J.; Haglund, C. N-glycomic profiling as a tool to separate rectal adenomas from carcinomas. *Mol. Cell. Proteomics* **2015**, *14*, 277–288. [[CrossRef](#)] [[PubMed](#)]
20. Miyoshi, E.; Moriwaki, K.; Terao, N.; Tan, C.C.; Terao, M.; Nakagawa, T.; Matsumoto, H.; Shinzaki, S.; Kamada, Y. Fucosylation is a promising target for cancer diagnosis and therapy. *Biomolecules* **2012**, *2*, 34–45. [[CrossRef](#)] [[PubMed](#)]
21. Moriwaki, K.; Noda, K.; Furukawa, Y.; Ohshima, K.; Uchiyama, A.; Nakagawa, T.; Taniguchi, N.; Daigo, Y.; Nakamura, Y.; Hayashi, N.; et al. Deficiency of GMDS leads to escape from NK cell-mediated tumor surveillance through modulation of TRAIL signaling. *Gastroenterology* **2009**, *137*, 188–198. [[CrossRef](#)] [[PubMed](#)]
22. Nakagoe, T.; Fukushima, K.; Nanashima, A.; Sawai, T.; Tsuji, T.; Jibiki, M.; Yamaguchi, H.; Yasutake, T.; Ayabe, H.; Matuo, T.; et al. Expression of Lewis<sup>a</sup>, sialyl Lewis<sup>a</sup>, Lewis<sup>x</sup>, sialyl Lewis<sup>x</sup>, antigens as prognostic factors in patients with colorectal cancer. *Can. J. Gastroenterol. Hepatol.* **2000**, *14*, 753–760. [[CrossRef](#)]
23. Konno, A.; Hoshino, Y.; Terashima, S.; Motoki, R.; Kawaguchi, T. Carbohydrate expression profile of colorectal cancer cells is relevant to metastatic pattern and prognosis. *Clin. Exp. Metastasis* **2002**, *19*, 61–70. [[CrossRef](#)] [[PubMed](#)]

24. Guo, R.J.; Suh, E.R.; Lynch, J.P. The role of Cdx proteins in intestinal development and cancer. *Cancer Biol. Ther.* **2004**, *3*, 593–601. [[CrossRef](#)] [[PubMed](#)]
25. Freund, J.N.; Domon-Dell, C.; Keding, M.; Duluc, I. The Cdx-1 and Cdx-2 homeobox genes in the intestine. *Biochem. Cell Biol.* **1998**, *76*, 957–969. [[CrossRef](#)]
26. Domon-Dell, C.; Schneider, A.; Moucadel, V.; Guerin, E.; Guenet, D.; Aguilon, S.; Duluc, I.; Martin, E.; Iovanna, J.; Launay, J.F.; et al. Cdx1 homeobox gene during human colon cancer progression. *Oncogene* **2003**, *22*, 7913–7921. [[CrossRef](#)] [[PubMed](#)]
27. Arango, D.; Al-Obaidi, S.; Williams, D.S.; Dopeso, H.; Mazzolini, R.; Corner, G.; Byun, D.S.; Carr, A.A.; Murone, C.; Togel, L.; et al. Villin expression is frequently lost in poorly differentiated colon cancer. *Am. J. Pathol.* **2012**, *180*, 1509–1521. [[CrossRef](#)] [[PubMed](#)]
28. Chan, C.W.; Wong, N.A.; Liu, Y.; Bicknell, D.; Turley, H.; Hollins, L.; Miller, C.J.; Wilding, J.L.; Bodmer, W.F. Gastrointestinal differentiation marker Cytokeratin 20 is regulated by homeobox gene CDX1. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 1936–1941. [[CrossRef](#)] [[PubMed](#)]
29. Al-Maghrabi, J.; Gomaa, W.; Buhmeida, A.; Al-Qahtani, M.; Al-Ahwal, M. Loss of villin immunorexpression in colorectal carcinoma is associated with poor differentiation and survival. *ISRN Gastroenterol.* **2013**, *2013*, 679724. [[CrossRef](#)] [[PubMed](#)]
30. Silberg, D.G.; Furth, E.E.; Taylor, J.K.; Schuck, T.; Chiou, T.; Traber, P.G. CDX1 protein expression in normal, metaplastic, and neoplastic human alimentary tract epithelium. *Gastroenterology* **1997**, *113*, 478–486. [[CrossRef](#)] [[PubMed](#)]
31. Vider, B.Z.; Zimber, A.; Hirsch, D.; Estlein, D.; Chastre, E.; Prevot, S.; Gespach, C.; Yaniv, A.; Gazit, A. Human colorectal carcinogenesis is associated with deregulation of homeobox gene expression. *Biochem. Biophys. Res. Commun.* **1997**, *232*, 742–748. [[CrossRef](#)] [[PubMed](#)]
32. Mallo, G.V.; Soubeyran, P.; Lissitzky, J.C.; Andre, F.; Farnarier, C.; Marvaldi, J.; Dagorn, J.C.; Iovanna, J.L. Expression of the Cdx1 and Cdx2 homeotic genes leads to reduced malignancy in colon cancer-derived cells. *J. Biol. Chem.* **1998**, *273*, 14030–14036. [[CrossRef](#)] [[PubMed](#)]
33. Wong, N.A.; Britton, M.P.; Choi, G.S.; Stanton, T.K.; Bicknell, D.C.; Wilding, J.L.; Bodmer, W.F. Loss of CDX1 expression in colorectal carcinoma: Promoter methylation, mutation, and loss of heterozygosity analyses of 37 cell lines. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 574–579. [[CrossRef](#)] [[PubMed](#)]
34. Lauc, G.; Essafi, A.; Huffman, J.E.; Hayward, C.; Knezevic, A.; Kattla, J.J.; Polasek, O.; Gornik, O.; Vitart, V.; Abrahams, J.L.; et al. Genomics meets glycomics—the first GWAS study of human N-Glycome identifies HNF1alpha as a master regulator of plasma protein fucosylation. *PLoS Genet.* **2010**, *6*, e1001256. [[CrossRef](#)] [[PubMed](#)]
35. Cantarino, N.; Fernandez-Figueras, M.T.; Valero, V.; Musulen, E.; Malinverni, R.; Granada, I.; Goldie, S.J.; Martin-Caballero, J.; Douet, J.; Forcales, S.V.; et al. A cellular model reflecting the phenotypic heterogeneity of mutant HRAS driven squamous cell carcinoma. *Int. J. Cancer.* **2016**, *139*, 1106–1116. [[CrossRef](#)] [[PubMed](#)]
36. Lussier, C.R.; Babeu, J.P.; Auclair, B.A.; Perreault, N.; Boudreau, F. Hepatocyte nuclear factor-4alpha promotes differentiation of intestinal epithelial cells in a coculture system. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2008**, *294*, G418–G428. [[CrossRef](#)] [[PubMed](#)]
37. Reiding, K.R.; Blank, D.; Kuijper, D.M.; Deelder, A.M.; Wuhrer, M. High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification. *Anal. Chem.* **2014**, *86*, 5784–5793. [[CrossRef](#)] [[PubMed](#)]
38. Strohal, M.; Hassman, M.; Kosata, B.; Kodicek, M. mMass data miner: An open source alternative for mass spectrometric data analysis. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 905–908. [[CrossRef](#)] [[PubMed](#)]
39. Jansen, B.C.; Reiding, K.R.; Bondt, A.; Hipgrave Ederveen, A.L.; Palmblad, M.; Falck, D.; Wuhrer, M. MassyTools: A High-Throughput Targeted Data Processing Tool for Relative Quantitation and Quality Control Developed for Glycomic and Glycoproteomic MALDI-MS. *J. Proteome Res.* **2015**, *14*, 5088–5098. [[CrossRef](#)]
40. Domon, B.; Costello, C.E. A Systematic Nomenclature for Carbohydrate Fragmentations in Fab-MS Ms Spectra of Glycoconjugates. *Glycoconj. J.* **1988**, *5*, 397–409. [[CrossRef](#)]
41. Boudreau, F.; Rings, E.H.; van Wering, H.M.; Kim, R.K.; Swain, G.P.; Krasinski, S.D.; Moffett, J.; Grand, R.J.; Suh, E.R.; Traber, P.G. 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.* **2002**, *277*, 31909–31917. [[CrossRef](#)]

42. Krasinski, S.D.; Van Wering, H.M.; Tannemaat, M.R.; Grand, R.J. Differential activation of intestinal gene promoters: Functional interactions between GATA-5 and HNF-1 alpha. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2001**, *281*, G69–G84. [[CrossRef](#)]
43. Boyd, M.; Hansen, M.; Jensen, T.G.; Pearnau, A.; Olsen, A.K.; Bram, L.L.; Bak, M.; Tommerup, N.; Olsen, J.; Troelsen, J.T. Genome-wide analysis of CDX2 binding in intestinal epithelial cells (Caco-2). *J. Biol. Chem.* **2010**, *285*, 25115–25125. [[CrossRef](#)] [[PubMed](#)]
44. Huflejt, M.E.; Leffler, H. Galectin-4 in normal tissues and cancer. *Glycoconj. J.* **2004**, *20*, 247–255. [[CrossRef](#)] [[PubMed](#)]
45. Zhao, Y.Y.; Takahashi, M.; Gu, J.G.; Miyoshi, E.; Matsumoto, A.; Kitazume, S.; Taniguchi, N. Functional roles of N-glycans in cell signaling and cell adhesion in cancer. *Cancer Sci.* **2008**, *99*, 1304–1310. [[CrossRef](#)] [[PubMed](#)]
46. Christiansen, M.N.; Chik, J.; Lee, L.; Anugraham, M.; Abrahams, J.L.; Packer, N.H. Cell surface protein glycosylation in cancer. *Proteomics* **2014**, *14*, 525–546. [[CrossRef](#)] [[PubMed](#)]
47. Guo, H.; Pierce, J.M. Transcriptional Regulation of Glycan Expression. In *Glycoscience: Biology and Medicine*; Endo, T., Seeberger, H.P., Hart, W.G., Wong, C.-H., Taniguchi, N., Eds.; Springer: Tokyo, Japan, 2014; pp. 1–7.
48. Trinchera, M.; Aronica, A.; Dall’Olio, F. Selectin Ligands Sialyl-Lewis a and Sialyl-Lewis x in Gastrointestinal Cancers. *Biology* **2017**, *6*, 16. [[CrossRef](#)] [[PubMed](#)]
49. Yeung, T.M.; Gandhi, S.C.; Wilding, J.L.; Muschel, R.; Bodmer, W.F. Cancer stem cells from colorectal cancer-derived cell lines. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 3722–3727. [[CrossRef](#)]
50. Breiman, A.; Lopez Robles, M.D.; de Carne Trecesson, S.; Echasserieau, K.; Bernardeau, K.; Drickamer, K.; Imberty, A.; Barille-Nion, S.; Altare, F.; Le Pendu, J. Carcinoma-associated fucosylated antigens are markers of the epithelial state and can contribute to cell adhesion through CLEC17A (Prolectin). *Oncotarget* **2016**, *7*, 14064–14082. [[CrossRef](#)]
51. Vellinga, T.T.; den Uil, S.; Rinkes, I.H.; Marvin, D.; Ponsioen, B.; Alvarez-Varela, A.; Fatrai, S.; Scheele, C.; Zwijnenburg, D.A.; Snippert, H.; et al. Collagen-rich stroma in aggressive colon tumors induces mesenchymal gene expression and tumor cell invasion. *Oncogene* **2016**, *35*, 5263–5271. [[CrossRef](#)]
52. Kim, S.W.; Park, K.C.; Jeon, S.M.; Ohn, T.B.; Kim, T.I.; Kim, W.H.; Cheon, J.H. Abrogation of galectin-4 expression promotes tumorigenesis in colorectal cancer. *Cell. Oncol.* **2013**, *36*, 169–178. [[CrossRef](#)] [[PubMed](#)]
53. Satelli, A.; Rao, P.S.; Thirumala, S.; Rao, U.S. Galectin-4 functions as a tumor suppressor of human colorectal cancer. *Int. J. Cancer.* **2011**, *129*, 799–809. [[CrossRef](#)]
54. Belo, A.I.; van der Sar, A.M.; Tefsen, B.; van Die, I. Galectin-4 Reduces Migration and Metastasis Formation of Pancreatic Cancer Cells. *PLoS ONE* **2013**, *8*, e65957. [[CrossRef](#)] [[PubMed](#)]
55. 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. USA* **2012**, *109*, 7776–7781. [[CrossRef](#)] [[PubMed](#)]
56. Hanamatsu, H.; Nishikaze, T.; Miura, N.; Piao, J.; Okada, K.; Sekiya, S.; Iwamoto, S.; Sakamoto, N.; Tanaka, K.; Furukawa, J.I. Sialic Acid Linkage Specific Derivatization of Glycosphingolipid Glycans by Ring-Opening Aminolysis of Lactones. *Anal. Chem.* **2018**, *90*, 13193–13199. [[CrossRef](#)] [[PubMed](#)]
57. Varki, A. Sialic acids in human health and disease. *Trends Mol. Med.* **2008**, *14*, 351–360. [[CrossRef](#)] [[PubMed](#)]
58. Corfield, A.P.; Myerscough, N.; Warren, B.F.; Durdey, P.; Paraskeva, C.; Schauer, R. Reduction of sialic acid O-acetylation in human colonic mucins in the adenoma-carcinoma sequence. *Glycoconj. J.* **1999**, *16*, 307–317. [[CrossRef](#)] [[PubMed](#)]
59. Holst, S.; Stavenhagen, K.; Balog, C.I.; Koeleman, C.A.; McDonnell, L.M.; Mayboroda, O.A.; Verhoeven, A.; Mesker, W.E.; Tollenaar, R.A.; Deelder, A.M.; et al. Investigations on aberrant glycosylation of glycosphingolipids in colorectal cancer tissues using liquid chromatography and matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF-MS). *Mol. Cell. Proteomics* **2013**, *12*, 3081–3093. [[CrossRef](#)] [[PubMed](#)]
60. Falck, D.; Habegger, M.; Plomp, R.; Hook, M.; Bulau, P.; Wuhler, M.; Reusch, D. Affinity purification of erythropoietin from cell culture supernatant combined with MALDI-TOF-MS analysis of erythropoietin N-glycosylation. *Sci. Rep.* **2017**, *7*, 5324. [[CrossRef](#)] [[PubMed](#)]
61. Isshiki, S.; Kudo, T.; Nishihara, S.; Ikehara, Y.; Togayachi, A.; Furuya, A.; Shitara, K.; Kubota, T.; Watanabe, M.; Kitajima, M.; et al. Lewis type 1 antigen synthase (beta3Gal-T5) is transcriptionally regulated by homeoproteins. *J. Biol. Chem.* **2003**, *278*, 36611–36620. [[CrossRef](#)] [[PubMed](#)]

62. Groux-Degroote, S.; Wavelet, C.; Krzewinski-Recchi, M.A.; Portier, L.; Mortuaire, M.; Mihalache, A.; Trinchera, M.; Delannoy, P.; Malagolini, N.; Chiricolo, M.; et al. B4GALNT2 gene expression controls the biosynthesis of Sda and sialyl Lewis X antigens in healthy and cancer human gastrointestinal tract. *Int. J. Biochem. Cell Biol.* **2014**, *53*, 442–449. [[CrossRef](#)] [[PubMed](#)]
63. Hirano, K.; Matsuda, A.; Shirai, T.; Furukawa, K. Expression of LacdiNAc groups on N-glycans among human tumors is complex. *Biomed. Res. Int.* **2014**, *2014*, 981627. [[CrossRef](#)] [[PubMed](#)]
64. Gu, J.; Sato, Y.; Kariya, Y.; Isaji, T.; Taniguchi, N.; Fukuda, T. A mutual regulation between cell-cell adhesion and N-glycosylation: Implication of the bisecting GlcNAc for biological functions. *J. Proteome Res.* **2009**, *8*, 431–435. [[CrossRef](#)] [[PubMed](#)]
65. Yoshimura, M.; Ihara, Y.; Matsuzawa, Y.; Taniguchi, N. Aberrant glycosylation of E-cadherin enhances cell-cell binding to suppress metastasis. *J. Biol. Chem.* **1996**, *271*, 13811–13815. [[CrossRef](#)] [[PubMed](#)]
66. Lu, J.; Isaji, T.; Im, S.; Fukuda, T.; Kameyama, A.; Gu, J. Expression of N-acetylglucosaminyltransferase III suppresses alpha2,3 sialylation and its distinctive functions in cell migration are attributed to alpha2,6 sialylation levels. *J. Biol. Chem.* **2016**, *291*, 5708. [[CrossRef](#)] [[PubMed](#)]
67. Li, X.; Wang, X.; Tan, Z.; Chen, S.; Guan, F. Role of Glycans in Cancer Cells Undergoing Epithelial-Mesenchymal Transition. *Front. Oncol.* **2016**, *6*, 33. [[CrossRef](#)] [[PubMed](#)]
68. Tan, Z.; Lu, W.; Li, X.; Yang, G.; Guo, J.; Yu, H.; Li, Z.; Guan, F. Altered N-Glycan expression profile in epithelial-to-mesenchymal transition of NMuMG cells revealed by an integrated strategy using mass spectrometry and glycogene and lectin microarray analysis. *J. Proteome Res.* **2014**, *13*, 2783–2795. [[CrossRef](#)] [[PubMed](#)]
69. Hirakawa, M.; Takimoto, R.; Tamura, F.; Yoshida, M.; Ono, M.; Murase, K.; Sato, Y.; Osuga, T.; Sato, T.; Iyama, S.; et al. Fucosylated TGF- $\beta$  receptors transduces a signal for epithelial-mesenchymal transition in colorectal cancer cells. *Br. J. Cancer* **2014**, *110*, 156–163. [[CrossRef](#)] [[PubMed](#)]
70. Hryniuk, A.; Grainger, S.; Savory, J.G.; Lohnes, D. Cdx1 and Cdx2 function as tumor suppressors. *J. Biol. Chem.* **2014**, *289*, 33343–33354. [[CrossRef](#)] [[PubMed](#)]
71. Ishikawa, F.; Nose, K.; Shibamura, M. Downregulation of hepatocyte nuclear factor-4alpha and its role in regulation of gene expression by TGF-beta in mammary epithelial cells. *Exp. Cell Res.* **2008**, *314*, 2131–2140. [[CrossRef](#)]
72. Pelletier, L.; Rebouissou, S.; Vignjevic, D.; Bioulac-Sage, P.; Zucman-Rossi, J. HNF1alpha inhibition triggers epithelial-mesenchymal transition in human liver cancer cell lines. *BMC Cancer* **2011**, *11*, 427. [[CrossRef](#)] [[PubMed](#)]
73. Yang, M.; Li, S.N.; Anjum, K.M.; Gui, L.X.; Zhu, S.S.; Liu, J.; Chen, J.K.; Liu, Q.F.; Ye, G.D.; Wang, W.J.; et al. A double-negative feedback loop between Wnt-beta-catenin signaling and HNF4alpha regulates epithelial-mesenchymal transition in hepatocellular carcinoma. *J. Cell Sci.* **2013**, *126*, 5692–5703. [[CrossRef](#)]

