



Article

Stool Glycoproteomics Signatures of Pre-Cancerous Lesions and Colorectal Cancer

Janine Soares^{1,2,3,†}, Mariana Eiras^{1,†}, Dylan Ferreira^{1,2,4,5,6}, Daniela A. R. Santos^{1,2,7},
Marta Relvas-Santos^{1,2,5,6,8}, Beatriz Santos¹, Martina Gonçalves¹, Eduardo Ferreira¹, Renata Vieira⁹,
Luís Pedro Afonso^{1,9}, Lúcio Lara Santos^{1,2,10,11,12}, Mário Dinis-Ribeiro^{7,13,14}, Luís Lima^{1,*,‡}
and José Alexandre Ferreira^{1,2,11,*,‡}

¹ Experimental Pathology and Therapeutics Group, Research Center of IPO Porto (CI-IPOP), RISE@CI-IPOP (Health Research Network), Portuguese Oncology Institute of Porto (IPO Porto), Porto Comprehensive Cancer Center Raquel Seruca (Porto.CCC Raquel Seruca), 4200-072 Porto, Portugal; janinepaivasoares@gmail.com (J.S.); mariana.eiras@ipoporto.min-saude.pt (M.E.); i38795@ipoporto.min-saude.pt (D.F.); daniela.r.santos@ipoporto.min-saude.pt (D.A.R.S.); marta.relvas.santos@ipoporto.min-saude.pt (M.R.-S.); beatriz.santos@ipoporto.min-saude.pt (B.S.); matina.goncalves@ipoporto.min-saude.pt (M.G.); i37310@ipoporto.min-saude.pt (E.F.); luis.afonso@ipoporto.min-saude.pt (L.P.A.); lucio.santos@ipoporto.min-saude.pt (L.L.S.)

² Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto, 4050-313 Porto, Portugal

³ REQUIMTE-LAQV, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal

⁴ Center for Applied Medical Research, University of Navarra, 31008 Pamplona, Spain

⁵ i3S—Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135 Porto, Portugal

⁶ INEB-Instituto Nacional de Engenharia Biomédica, Universidade do Porto, 4200-135 Porto, Portugal

⁷ Faculty of Medicine (FMUP), University of Porto, 4200-072 Porto, Portugal;

mario.ribeiro@ipoporto.min-saude.pt

⁸ REQUIMTE-LAQV, Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal

⁹ Department of Pathology, Portuguese Oncology Institute of Porto, 4200-072 Porto, Portugal;

renata.vieira@ipoporto.min-saude.pt

¹⁰ FF-I3ID, University Fernando Pessoa, 4249-004 Porto, Portugal

¹¹ GlycoMatters Biotech, 4500-162 Espinho, Portugal

¹² Department of Surgical Oncology, Portuguese Oncology Institute of Porto (IPO-Porto), 4200-072 Porto, Portugal

¹³ Precancerous Lesions and Early Cancer Management Group, Research Center of IPO Porto (CI-IPOP), Rise@CI-IPOP (Health Research Group), Portuguese Institute of Oncology of Porto (IPO Porto), Porto Comprehensive Cancer Center Raquel Seruca (Porto.CCC Raquel Seruca), 4200-072 Porto, Portugal

¹⁴ Department of Gastroenterology, Portuguese Oncology Institute of Porto, 4200-072 Porto, Portugal

* Correspondence: luis.carlos.lima@ipoporto.min-saude.pt (L.L.); jose.a.ferreira@ipoporto.min-saude.pt (J.A.F.)

† These authors contributed equally to this work.

‡ These authors also contributed equally to this work.



Citation: Soares, J.; Eiras, M.; Ferreira, D.; Santos, D.A.R.; Relvas-Santos, M.; Santos, B.; Gonçalves, M.; Ferreira, E.; Vieira, R.; Afonso, L.P.; et al. Stool Glycoproteomics Signatures of Pre-Cancerous Lesions and Colorectal Cancer. *Int. J. Mol. Sci.* **2024**, *25*, 3722. <https://doi.org/10.3390/ijms25073722>

Academic Editor: Peter J. K. Kuppen

Received: 1 March 2024

Revised: 22 March 2024

Accepted: 25 March 2024

Published: 27 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Colorectal cancer (CRC) screening relies primarily on stool analysis to identify occult blood. However, its sensitivity for detecting precancerous lesions is limited, requiring the development of new tools to improve CRC screening. Carcinogenesis involves significant alterations in mucosal epithelium glycoalkalx that decisively contribute to disease progression. Building on this knowledge, we examined patient series comprehending premalignant lesions, colorectal tumors, and healthy controls for the T-antigen—a short-chain O-glycosylation of proteins considered a surrogate marker of malignancy in multiple solid cancers. We found the T-antigen in the secretions of dysplastic lesions as well as in cancer. In CRC, T-antigen expression was associated with the presence of distant metastases. In parallel, we analyzed a broad number of stools from individuals who underwent colonoscopy, which showed high T expressions in high-grade dysplasia and carcinomas. Employing mass spectrometry-based lectin-affinity enrichment, we identified a total of 262 proteins, 67% of which potentially exhibited altered glycosylation patterns associated with cancer and advanced pre-cancerous lesions. Also, we found that the stool (glyco)proteome of pre-cancerous lesions is enriched for protein species involved in key biological processes linked to humoral and innate immune responses. This study offers a thorough analysis of the stool glycoproteome, laying the groundwork for harnessing glycosylation alterations to improve non-invasive cancer detection.

Keywords: cancer glycosylation; stool glycoproteins; glycoproteomics; pre-cancerous lesions; colorectal cancer

1. Introduction

Colorectal cancer (CRC) is the third most common neoplasia and the second leading cause of cancer-related death globally, posing a significant global health concern [1–3]. Early detection has been demonstrated to prolong patients' lives by more than five years, with approximately 90% of patients experiencing positive outcomes from timely interventions [2,4]. However, this percentage drops to 10% when CRC is diagnosed in advanced stages, particularly with distant metastases, underscoring the major role of screening programs [2]. Among the various screening modalities, stool analysis has emerged as a cornerstone approach due to its non-invasive nature and the possibility of detecting occult blood, a potential indicator of CRC or precancerous lesions [5,6]. Stool-based molecular analysis holds the potential for the identification of individuals bearing colon lesions and enables clinicians to intervene promptly, thereby improving patient outcomes. Furthermore, it may serve as a valuable tool for monitoring CRC patients undergoing treatment, aiding in response assessment and disease progression monitoring [2,7]. The most widely adopted method for CRC screening through stool analysis is the Fecal Immunochemical Test (FIT), which detects the presence of occult blood using antibodies specifically targeting human hemoglobin [6]. FIT demonstrates good sensitivity in detecting cancerous lesions (approximately 70%), namely adenocarcinomas. However, its sensitivity for detecting precancerous lesions decreases to 30%, limiting the overall effectiveness of early screening. To detect advanced lesions at higher risk of progressing to cancer remains a major limitation [6,8]. Furthermore, approximately 30% of FIT-positive cases involve individuals without lesions in the colon and rectum, resulting in unnecessary colonoscopies [6]. Notably, colonoscopy is widely used as the gold standard procedure for colorectal lesion identification due to its high sensitivity for detecting and allowing the removal of pre- and cancerous lesions [2,5]. Nevertheless, colonoscopy is an invasive procedure with inherent limitations such as low patient adherence and high medical and economic burden [9,10]. These limitations underscore the need to develop non-invasive methods based on new molecular biomarkers, such as (glyco)proteins, which have the potential to enhance CRC screening [11].

The pathogenesis of CRC involves intricate molecular alterations, including profound changes in the mucosal epithelium glycocalyx, which play a decisive role in disease progression. Particularly, protein *N*- and *O*-glycosylations have been extensively investigated through in-depth mass spectrometry-based glycomics over the past years, revealing significant alterations linked to disease [12]. Specifically, a decrease in the abundance of core 3 *O*-glycans has been observed in inflamed colon and cancer compared to normal colon tissues [12,13]. Consequently, tumors tend to accumulate core 2 *O*-glycans as well as abnormally shorter *O*-glycan signatures such as the simple disaccharide antigen Thomsen–Friedenreich antigen (TF or T-antigen; Gal(β 1–3)GalNAc(α 1–O–Ser/Thr), also termed core 1 [13,14]. The T-antigen is mostly absent from normal cells but has been frequently observed in tumors, even though its pathological and clinical consequences remain to be fully elucidated [15]. Moreover, the T-antigen has been detected in liver metastases, the most common hematogenic dissemination in CRC [16]. Furthermore, it has been observed in other solid tumors, such as gastric, lung, prostate, and melanoma, which are linked to aggressiveness [17–20]. However, apart from certain types of mucins (MUC1, MUC4), the identity of the glycoproteins carrying these glycans remains largely unknown [21,22]. This hampers our comprehension of disease contexts involving the T-antigen and our capacity to effectively explore its biomarker potential.

Building on these observations, we hypothesize that glycoproteins carrying this immature *O*-glycan may be shed or secreted into stools and potentially serve as surrogate biomarkers of malignancy. Therefore, the present study aims to characterize the expression

of the T-antigen in premalignant lesions and CRC tumors compared to healthy tissues, providing grounds for further investigations in stools. Our investigation extends to characterizing the T-antigen expression in stool samples from CRC screening programs, providing valuable insights into disease-associated changes in the stool (glyco)proteome. This is expected to not only provide insights into the potential of these glycans in the context of non-invasive CRC screening but also pave the way for more in-depth characterization of the yet unexplored potential of the glycocalyx in this context.

2. Results

In this work, we hypothesized that the expression of the immature O-glycan T-antigen could hold potential cancer signatures, given its rare expression in healthy human tissues [15]. Furthermore, we explored the presence of these glycoproteins in stools as potential non-invasive signatures of disease.

2.1. T-Antigen in Colorectal Tissues

We started by screening the T-antigen using peanut agglutinin (PNA) lectin, which has been shown to specifically target the T-antigen [23,24]. The T-antigen was evaluated in histologically normal colon and rectum tissues, low-grade (n = 5) and high-grade (n = 3) pre-cancerous lesions, and a series of tumor samples representing all stages of the disease. The T-antigen was not detected in the histologically normal colorectal mucosa but was found in 50% of pre-cancerous lesions (60% low-grade and 33% high-grade dysplasia; Figure 1b), mainly low-grade dysplasia. It was also found in approximately 45% of the cancer cases, predominantly in extracellular secretions in gland lumens and/or mucin pools (Figure 1a). Few cases present scattered expression in the cytoplasm of epithelial cells, most likely belonging to precursors of more extended glycosidic chains in the O-glycan biosynthesis pathways (Figure 1a). Despite the high expression of the T-antigen in CRC, we did not find associations between positivity and expression level (intensity and/or extension) with the degree of invasion, clinical stage, tumor location, or survival (Table 1). However, we found that a higher percentage of T-positive tumors were metastatic at the time of presentation (Table 1). In summary, we have shown the disease-specific nature of the T-antigen, providing evidence that it could be expressed in both pre-cancerous lesions as well as cancer, being associated with the presence of metastases. The almost exclusive expression of the T-antigen on secretions and very rare detection in epithelial cells suggest possible remodeling of the mucus layer glycocalyx by the microbiome with a potential impact on disease, as suggested by different reports [25,26].

Table 1. Association between T-antigen expression and individuals' clinicopathological characteristics.

		T-Antigen Positive	T-Antigen Negative	p-Value
		n (%)	n (%)	
Stage	I	1 (5.55)	0 (0)	0.085
	II	0 (0)	4 (18.18)	
	III	3 (16.67)	7 (31.82)	
	IV	14 (77.78)	11 (50.00)	
Tumor (T)	T1	0 (0)	1 (4.76)	0.558
	T2	1 (5.55)	0 (0)	
	T3	14 (77.78)	17 (80.95)	
	T4	3 (16.67)	3 (14.29)	
	Missing information		1 (4.76)	

Table 1. Cont.

	T-Antigen Positive	T-Antigen Negative	<i>p</i> -Value
	<i>n</i> (%)	<i>n</i> (%)	
Lymph node metastases (N)			
N0	5 (27.78)	6 (27.27)	0.679
N1	7 (38.89)	5 (22.73)	
N2	5 (27.78)	9 (40.91)	
N3	1 (5.55)	2 (9.09)	
Distant metastases (M)			
M0	4 (22.22)	12 (54.5)	0.039
M1	14 (77.78)	10 (45.5)	
Tumor location			
Right colon	3 (16.67)	4 (18.19)	0.152
Left colon	13 (72.22)	10 (45.45)	
Rectum	2 (11.11)	8 (36.36)	

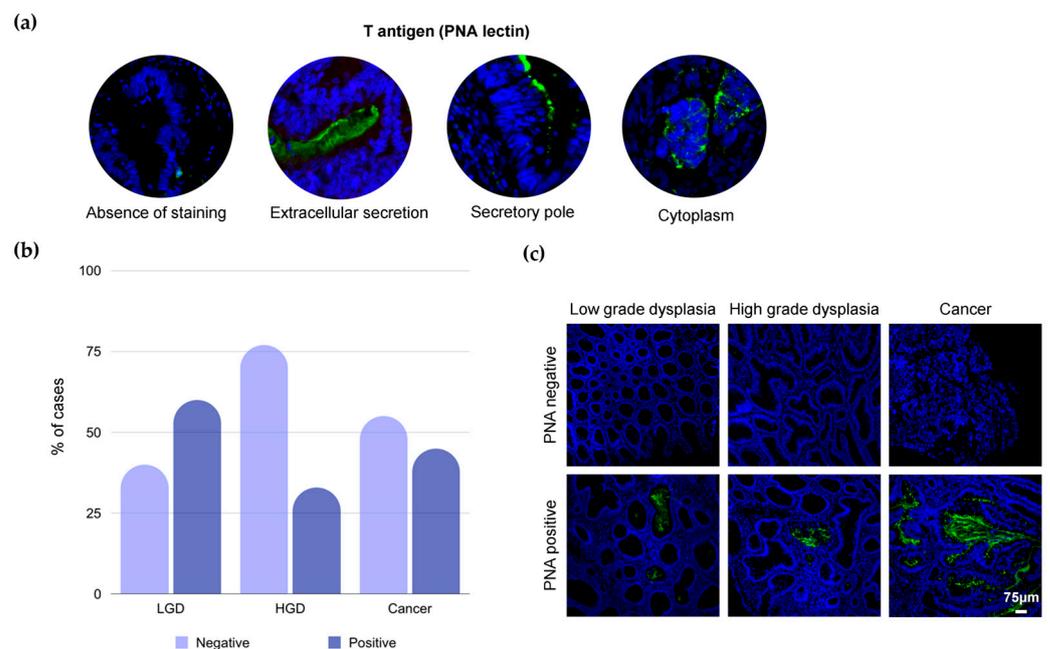


Figure 1. The T-antigen has a mild expression in cancer, as well as in pre-malignant lesions, low-grade dysplasia (LGD), and high-grade dysplasia (HGD), but not in the healthy colon mucosa. (a) The T-antigen is expressed in extracellular secretions and the secretory pole and rarely in the cytoplasm of tumors (magnification: 400×). (b) The T-antigen is expressed throughout cancer and pre-malignant lesions, although higher expression is present in cancer and LGD. (c) T-antigen expression pattern in LGD and HGD and cancer (magnification: 400×).

2.2. T-Antigen in Stool

The fact that the T-antigen was almost exclusively observed in secretions facing the lumen of the colorectal mucosa led us to hypothesize that it could be detected in stools and serve as a basis to improve non-invasive CRC screening and follow-ups. We addressed this hypothesis in preliminary settings by screening FIT-positive stools from individuals where lesions with low- and/or high-grade dysplasia were found, as well as CRC patients. We found 13% of T-positive cases among low-grade (3/24) and 67% (8/12) among high-grade dysplasia lesions (Figure 2a,b), generally considered precursors of CRC [27]. Additionally, we observed 33% of T-positive stools from CRC patients (Figure 2a,b). Furthermore, the T-antigen was not detected in the stools collected from individuals who underwent

colonoscopy and where no lesions were found. Notably, a comparative analysis between tissues and corresponding stool samples showed a 50% match (Figure 2c). Collectively, these findings reinforce the link between the T-antigen and colorectal pathogenesis. Furthermore, it highlights the potential of the T-antigen to identify high-grade dysplasia, which constitutes a significant milestone for the development of CRC, warranting frequent invasive follow-ups.

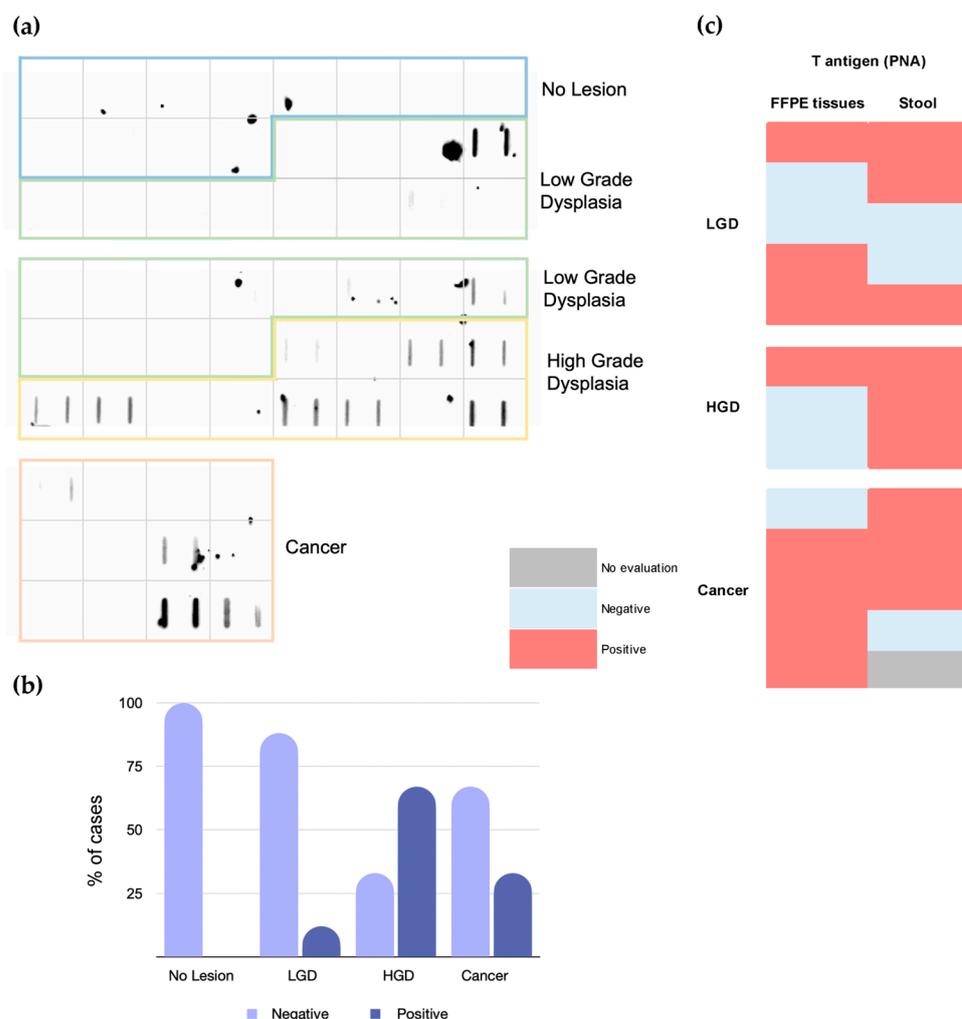


Figure 2. The T-antigen is expressed in the stools of patients with pre-malignant colon lesions (24 cases LGD, 12 cases HGD), in cancer (12 cases), and in no lesion control cases (12), being more frequent in HGD and cancer. (a) Silver squares represent each patient, and two biological replicates were used, and (b) The T-antigen is more frequent in proteins isolated from the stools of high-grade (67%) and cancer (33% of cancer cases). (c) There is a 50% match between T positivity in stools and tissues.

2.3. Proteomics and Glycoproteomics

The disparity in T-antigen expression between tissues and stool samples has prompted deeper investigations into the stool proteome and glycoproteome expressing T-antigens (Figure 3; Supplementary Figure S2). For this study, we selected T-positive stool samples from a subset of three individuals from each group (presenting low-grade dysplasia, high-grade dysplasia, or CRC), as well as individuals where no lesions were found. These samples were enriched for T-expressing glycoproteins using PNA lectin affinity chromatography. Subsequently, both the glycoprotein fractions and the flow-through, corresponding to the remaining portion of the stool proteome, were analyzed by tandem mass spectrometry. Before enrichment, the samples were reduced, alkylated, and digested with PNGase F to remove *N*-glycans. According to our experience, the removal of *N*-glycans enhances the

interaction of glycoproteins with PNA and reduces potential ambiguities in T-glycopeptides MS/MS spectra annotations.

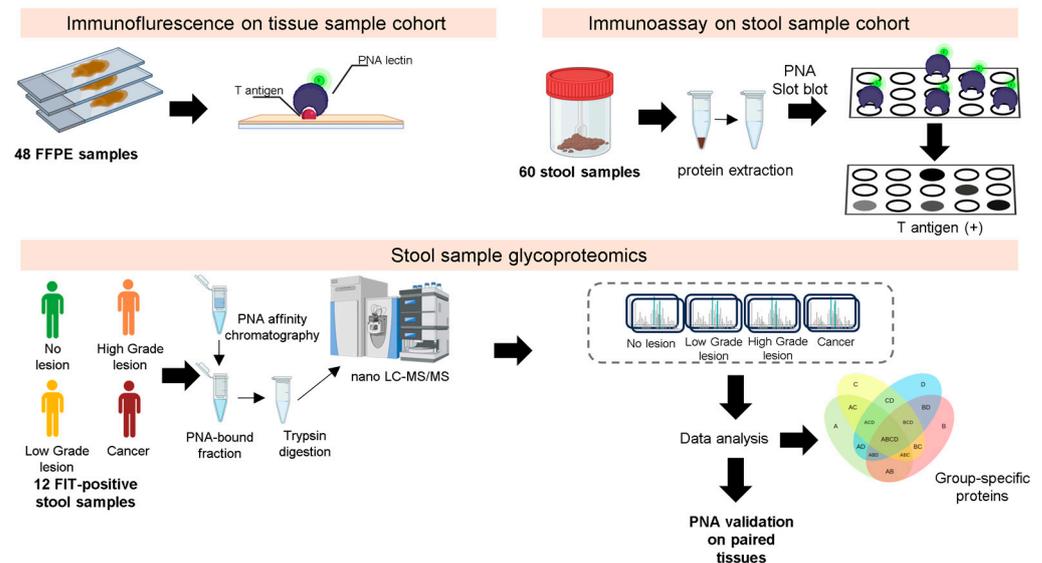


Figure 3. A representative scheme of the proteomics and glycoproteomics workflow was used in this study. Firstly, a cohort of FFPE tissues was screened for the presence of T-antigen, followed by validation in FIT-positive patient stool samples. Briefly, proteins were isolated from stools after homogenization in a lysis buffer solution. The proteins were subsequently reduced, alkylated, and digested with PNGase F to remove *N*-glycans and enriched for T-expressing glycoproteins using PNA lectin affinity chromatography. Both the glycoproteins fraction and unbound proteins were then digested with trypsin and analyzed by nanoLC-MS/MS.

Overall, we analyzed the proteome and glycoproteome of 12 stool samples, identifying a total of 262 proteins (Supplementary Table S1), 67% of which potentially carry the T-antigen (Figure 4a and Supplementary Table S2). Both the stool proteome and glycoproteome presented group-specific signatures, proteins commonly expressed protein between cancer and pre-cancerous lesions and common to all groups (Figure 4b,c and Supplementary Table S2). A higher number of group-specific glycoproteins were identified in cancer samples in relation to other pathological samples (Figure 4c). Nevertheless, particular attention was devoted to a group of 105 glycoproteins (Figure 4c), found in PNA-enriched fraction (Supplementary Table S3), consistently observed in all three samples of at least one of the groups (no lesion (NL), low-grade dysplasia (LG), high-grade dysplasia (HG), cancer). Among these, we found group-specific glycoprotein signatures for cancer (IGHG2, KRT6A, IGLC2, C9, CLCA4, ORM1, HPX, ANXA2P2, HP, A2M, LGALS4) as well as dysplasia (low-grade: COL4A2, PLB1, PRSS2; High-grade: COL4A1, NAALADL1, CAT)—Supplementary Tables S4–S6. The presence of the T-antigen in these glycoproteins was confirmed by tandem mass spectrometry, as highlighted by the MS/MS spectra for cancer-related glycoproteins alpha-1-acid glycoprotein 1 and alpha-2 macroglobulin (Supplementary Figure S1). Notably, the number of group-specific glycoproteins was higher in cancer compared to pre-cancerous lesions. Moreover, no group-specific glycoproteins could be found in the stools of FIT-positive cases without colorectal histological alterations. However, we also observed 12 glycoproteins common both to cancer and precancerous lesions groups (Figure 4c) and 43 glycoproteins present in all evaluated groups (Figure 4c; Supplementary Tables S8 and S9). Among cancer and precancerous lesions glycoproteins, we further highlight seven glycoproteins found in high-grade dysplasia as well as cancer (IGHG1, CPA2, MYH2, AZU1, PRTN3, CA1, SMPDL3B), holding potential for non-invasive detection of aggressive traits (Figure 4c; Supplementary Table S7). Overall, we found that the stool glycoproteome reflects a wide array of glycoproteins characteristic of epithelial

cells of the gastrointestinal tract, such as mucins (MUC2, MUC5AC), collagen fragments (COL4A1, COL4A2 COL4A1, COL4A2), and integrins. Multiple digestive enzymes (CPA1, CTRC, IAP, PA) as well as serum proteins (HBB, IGHG1, SERPINC1, A2M) were also identified, as expected due to the presence of bleeding in these stools. The presence of the T-antigen on these glycoproteins, supported by tandem mass spectrometry, suggests various origins besides the colorectal epithelium for the glycoproteins present in stools. This finding provides a plausible explanation for the absence of a direct correlation between the presence of the T-antigen in colorectal tissues and stools (Figure 2c).

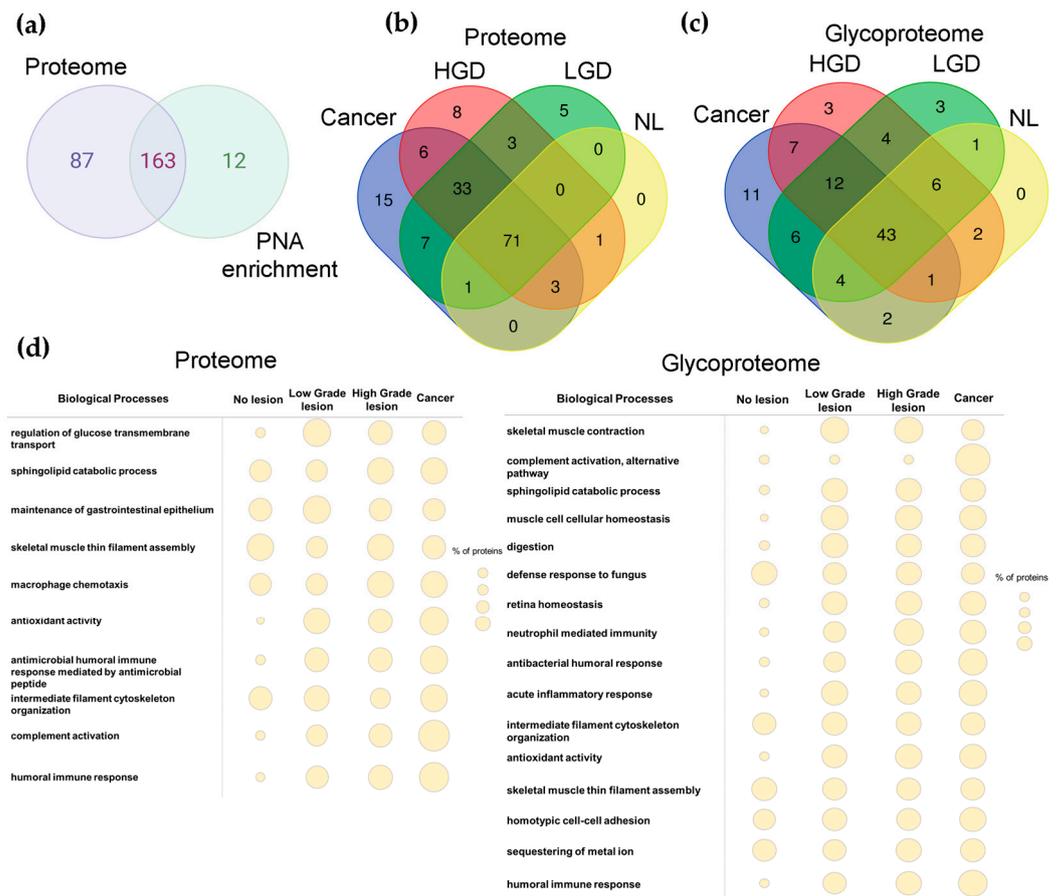


Figure 4. The stool proteomes and glycoproteomes change with disease state (low-grade and high-grade dysplasia and CRC) and present unique proteins that hold the potential for monitoring disease progression. **(a)** A total of 262 proteins were identified, 67% of which were in the glycoproteome fraction, suggesting modification with the T-antigen. **(b,c)** The stool proteome and glycoproteome show disease-state-specific glycoproteins. Regarding group specific proteome, there were identified for cancer: HBB, HPX, C7, FGA, LYPD8, TF, LGALS4, PSCA, FOLH1, FGG, C9, TKT, IGHV3-38, ORM2, SERPINB1, HGD: TNNC2, GDPD3, IGKV4-1, PIP, TFF2, CAT, SMR3B, COL4A1, LGD: HLA-DRB5, ATP2A1, CEACAM7, IGLV4-69, CEACAM1, and further highlight six proteins found in LGD and cancer: CALM3, PRTN3, IGHM, CA4, AZU1, CA1. Also, there were identified group-specific glycoprotein signatures for cancer: IGHG2, KRT6A, IGLC2, C9, CLCA4, ORM1, HPX, ANXA2P2, HP, A2M, LGALS4, LGD: COL4A2, PLB1, PRSS2, HGD: COL4A1, NAALADL1, CAT, and we further highlight seven glycoproteins found in HGD and cancer: IGHG1, CPA2, MYH2, AZU1, PRTN3, CA1, SMPDL3B. **(d)** The stool proteome and glycoproteome show disease state-specific functional signatures, by the presence of a higher percentage of proteins in patients bearing lesions and cancer, related to inflammation processes and immune response.

Finally, we comprehensively analyzed the stool proteome and glycoproteome in terms of main biological functions (Figure 4d). In general, we found that the stool proteome of pre-malignant lesions and CRC was enriched for protein species involved in the regulation of glucose transport across the cell membrane, antioxidant activity, and immune responses, namely linked to complement activation and humoral responses against bacteria. This suggests some degree of immune responses against the gut microbiome dysregulations, which become more pronounced in malignant conditions. Notably, all these biological functions were residually represented in the control group, which could be explained by a more stable state of commensal relationship between the microbiome and the host. Also, the number of glycoproteins involved in immune responses increased in cancer compared to pre-malignant lesions. By zooming in on the glycoproteome, we can see a representation of a wider number of biological functions. The presence of glycoproteins linked to immune responses was evident and particularly prominent in cancer and precancerous groups, similar to observations in the proteome. In addition to humoral responses, particularly against bacteria, there was also a relevant number of glycoproteins associated with acute inflammatory responses. Furthermore, CRC exhibited significant enrichment for biological functions linked to alternative pathway complement activation that was not observed in controls and pre-cancerous lesions, suggesting a cancer-specific signature. We also note an enrichment of proteins associated with homeostasis, muscle contraction, and digestion in cancer and precancerous groups that were not evident from proteome data. Collectively, our findings support the existence of pre-cancerous and CRC-specific proteomes and glycoproteomes that warrant further investigation towards non-invasive stool-based CRC screening.

3. Discussion

Occult blood in stool serves as a crucial initial indicator of possible alterations in the colorectal mucosa, providing a rationale for CRC screening programs and potentially life-saving interventions. However, limitations related to its sensitivity for detecting CRC and the incapacity to accurately identify advanced lesions (i.e., high-grade dysplasia) pose serious limitations for widespread adoption. Addressing these aspects, we have started by demonstrating that the cancer-associated T-antigen is present in secretions of both CRC and pre-cancerous lesions, holding the potential to be released into stools. This hypothesis was later confirmed by lectin blotting and mass spectrometry analysis, providing a new perspective for glycan-based non-invasive detection using a yet unexplored biological sample in this context.

CRC revealed a higher prevalence of the T-antigen, showing a significant association with the presence of metastases. These observations are consistent with previous reports implicating this glycan in disease spread [28,29]. Namely, the T-antigen has been described as promoting metastasis by facilitating cancer cell interaction with soluble galectins, particularly galectin-3 (Gal-3) [30,31]. These interactions have been shown to drive the adhesion of metastatic cells to the endothelium, homotypic cancer cell aggregation at primary attachment sites, the formation of intravascular metastatic deposits in lungs and bones, and the clonogenic survival and growth of metastatic cancer cell lines [30]. Our findings, together with previous reports documenting the presence of the T-antigen in liver metastases [16], which is the most common hematogenous dissemination route in CRC, reinforce the close link between this glycan and CRC dissemination. This supports ongoing efforts to target the T-antigen-Gal-3 axis with inhibitory molecules and suggests the potential for non-invasive follow-ups of disease progression.

Building on these preliminary findings, we engaged in the assessment of the T status in stool, aiming to create a rationale for glycan-based non-invasive CRC screening. Interestingly, while individuals with healthy colorectal mucosa tested negative for the T-antigen, its elevation was particularly notable in the stools of high-grade adenomas (advanced adenomas). These observations provide an opportunity to enhance FIT's capacity for aggressive lesion detection and improve CRC screening programs' effectiveness. Early

detection of high-grade dysplasia lesions based on novel biomarkers will be essential for reducing disease burden and improving patient outcomes. Screening for the T-antigen may enable a more efficient allocation of healthcare resources to individuals at higher risk of developing CRC, allowing timely diagnostic and therapeutic interventions. Additionally, this advancement could enhance follow-up care for patients with high-grade adenomas after resection, enabling easier and faster identification of relapse and facilitating closer non-invasive monitoring. Moreover, the association between the T-antigen and metastasis suggests the potential for closer and more frequent cancer progression follow-ups. Overall, our preliminary findings support further investment in assessing the T-antigen to improve stool-based identification of advanced precancerous lesions and progression to CRC. To our knowledge, this is the first report showcasing the potential of glycosylation in stool for non-invasively detecting colorectal lesions.

Finally, we employed a targeted glycoproteomics approach previously demonstrated to be effective in identifying cancer biomarkers [23,32] to investigate the glycoproteins carrying the T-antigen in stools. Our findings challenge the notion of a direct link between the detection of the T-antigen in colorectal tissues and stools. Instead, our results suggest a more complex scenario where a diverse array of glycoproteins from various sources may contribute to the presence of this glycan in stools. These sources include glycoproteins shed or secreted from cells lining the gastrointestinal tract, such as CEACAM5, integrins, cadherins, collagen fragments, and typical colon and rectum mucins like MUC2 and MUC5AC. Additionally, enzymes from secretions of the digestive glands and serum proteins from bleeding were also implicated. Notably, the T-antigen is typically associated with cancer, and its presence in serum proteins and even digestive enzymes can be deemed unusual. In fact, this glycan is widely expressed throughout the body but assumes a cryptic sialylated nature, often appearing in the form of sialyl-T or di-sialylated T-antigens. One plausible explanation for the unexpected identification of the T-antigen in these glycoproteins may involve the activity of enteric commensal and pathogenic bacteria present in the gut microbiome, which are recognized for their ability to utilize sialic acids as a nutrient source through the secretion of sialidases [33]. These are challenging concepts that suggest a connection between the presence of the T-antigen in stools and the microbiome associated with aggressive pre-malignant lesions and diseases. Nevertheless, the possibility of a systemic dysregulation of glycosylation caused by cancer cannot be overruled. Overall, more in-depth studies are needed to disclose these potential relationships fully.

Our study, while exploratory in nature, has provided valuable insights into the glycoproteomes present in both stool samples from individuals without lesions and those from individuals at various stages of disease progression. Through comprehensive analysis, we identified common functional patterns across these samples, regardless of disease severity. Particularly, we observed a notable enrichment of glycoproteins in stool samples of pre-malignant lesions and cancer associated with humoral immune responses, including responses targeting bacteria and acute inflammatory responses such as neutrophil-mediated immunity. An intriguing finding was the significant enrichment of glycoproteins linked to the complement alternative pathway in stools derived from tumors. This pathway, capable of activation in the absence of antibodies, offers a swift and efficient mechanism for enhancing immune responses against pathogens [34]. In summary, our findings suggest typical immune responses against lesions and potential responses against the gut microbiome, known to undergo alterations with disease progression. This sheds light on the intricate interplay between the immune system, microbial communities, and disease states within the gastrointestinal tract [35–38]. Other relevant biological signatures linked to the disease include changes in the levels of digestive enzymes, regulators of tissue homeostasis, and muscle contraction, suggesting organic dysregulation. In addition to these findings, we identified specific glycoproteins uniquely present in cancer samples, suggesting their potential as biomarkers for cancer detection or monitoring. Among these are HPX, C9, and LGALS4. HPX, synthesized in the liver in response to heme release, is elevated in the serum of CRC patients [39–41] and other solid tumors [41], indicating a potential pancreatic

nomic signature. Notably, HPX has also been reported in the stools of pre-cancerous and CRC patients [42,43], reinforcing its potential as a stool biomarker. However, its role in cancer remains controversial, with evidence suggesting both tumor-suppressive [44] and pro-tumorigenic effects [45]. Notably, sialyl-T-carrying glycoproteoforms of this protein have been found elevated in metastatic CRC [39] and ovarian cancer ascites fluids [46], indicating a significant association with metastasis. C9, primarily produced in the liver, has also been reported to increase in the serum of CRC patients [47,48], particularly in metastatic cases. Its role in cancer also remains poorly understood, with some studies suggesting tumor-promoting effects through inflammation and immune evasion, while others propose immune-mediated tumor destruction [49,50]. LGALS4, expressed by gastrointestinal epithelial cells, plays critical roles in cell adhesion, differentiation, and mucosal barrier formation [51]. It has been implicated in the binding and sorting of glycoproteins in the intestinal epithelium, aiding in the formation of the mucosal barrier and facilitating nutrient absorption [51,52]. Overall, LGALS4 plays a critical role in gastrointestinal physiology and may have implications for diseases such as inflammatory bowel disease and cancer [51]. LGALS4 was also found elevated in the serum and colorectal tumors and associated with poor prognosis [53,54]. In addition, we found several glycoproteins also elevated in high-grade adenomas. In this context, we highlight that AZU1 has also been previously observed as elevated in the stools of adenomas and CRC patients [42], as well as PRTN3, found in polyps and CRC tissues [55]. Additionally, elevated serum levels of PRTN3 have recently been described as an independent poor prognostic factor for progression-free survival and poor response to bevacizumab in patients with metastatic CRC [56]. Moreover, IGHG2, HPT, and A2MG identified in stool-related cancer samples, along with catalase detected in advanced pre-cancerous lesions, have been previously reported elevated in both cancerous and advanced non-cancerous lesions [43]. In summary, a significant body of evidence links many of the stool glycoproteins potentially carrying the T-antigen identified in this study (HPX, C9, LGALS4, AZU1, PRTN3, IGHG2, HPT, A2MG, CATA) to high-grade dysplasia lesions and/or CRC, providing grounds for their exploitation in the context of stool-related disease detection and improving colorectal risk assessment.

4. Materials and Methods

4.1. FFPE Tissues for Immunohistochemistry Studies

A retrospective patient sample set of 40 formalin-fixed paraffin-embedded (FFPE) CRC tissues collected from the institutional pathology biobank was selected for this study. The sample set included samples collected from non-pediatric patients who underwent surgery at the Portuguese Institute of Oncology of Porto (IPO-Porto) from 2005 to 2012 for resection of colorectal (adeno)carcinomas. Non-epithelial tumors were not considered in the selection. 18 female and 22 male patients, aged between 28 and 76 years (mean: 61 ± 11 years), were elected to participate. Moreover, FFPE tissues from individuals submitted to colonoscopy presenting precancerous lesions were also collected (Low-grade dysplasia ($n = 5$) and High-grade dysplasia ($n = 3$)). Adjacent histologically normal mucosa tissues were also analyzed whenever represented on the selected tumor sample. All procedures were approved by the institutional Ethics Board (Approval No. 64/2017 and 22/022, dated 16 March 2017 and 10 February 2022, respectively) and were conducted after signed formal patient consent. Table 2 summarizes the clinicopathological data collected from the patients' medical records.

Table 2. Clinicopathological data associated with tumor tissues were used for this study (n = 40).

	N (%)
Stage	
I	1 (2.5)
II	4 (10)
III	10 (25)
IV	25 (65)
Tumor (T)	
T1	1 (2.5)
T2	1 (2.5)
T3	31 (77.5)
T4	6 (15)
Missing Information	1 (2.5)
Lymph node metastases (N)	
N0	11 (27.5)
N1	12 (30)
N2	14 (35)
N3	3 (7.5)
Tumor Location	
Right colon	7 (17.5)
Left colon	23 (57.5)
Rectum	10 (25)

4.2. Patient Stool Sample Set

A prospective study was conducted at the Gastroenterology Department of IPO-Porto, enrolling individuals with FIT-positive results who underwent colonoscopy. Between April 2022 and July 2023, stool samples were freshly collected the day before patients' bowel preparation for colonoscopy. The collection and storage of fecal samples were adapted from previous publications [57,58]. Briefly, patients were instructed to collect feces in a sealed plastic container and transport them with an ice bag. Samples were codified and then manually homogenized with a spatula and stored at -80°C until analysis. A total of 60 samples were included in this study. The samples were selected considering the gender and age of the individuals to maintain cohort homogeneity. Samples were divided into different groups, according to the colonoscopy findings and histological characteristics of biopsy tissue collected during colonoscopy: no lesion found in colonoscopy, low-grade dysplasia, high-grade dysplasia, and patients diagnosed with adenocarcinoma. The individuals' data are summarized in Table 3. The mean age of individuals was 60.5 years (min–max: 49–83 years). The male/female ratio was approximately 50%. The study received approval from the IPO institutional Ethical Committee (Approval number CES 22/022, dated 10 February 2022) and was conducted following international guidelines, whereby all included patients signed the informed consent.

Table 3. Demographic data related to stool samples set (n = 60).

Characteristics	No Lesion (n = 12)	Low Grade Dysplasia (n = 24)	High-Grade Dysplasia (n = 12)	Cancer (n = 12)
Age (mean, min–max)	58.0 (50–72)	56.0 (49–74)	63.5 (40–85)	64.5 (51–89)
Gender				
Male	37.5%	37.5%	58.3%	54.2%
Female	62.5%	62.5%	41.7%	45.8%

4.3. Immunofluorescence for the T-Antigen

The T-antigen was evaluated in FFPE tissues from pre-malignant and CRC lesions. Briefly, FFPE tissues were prepared as previously described by us [59]. After deparaffinization, rehydration, and antigen retrieval, they were incubated with FITC-labeled PNA

lectin (1:1000; Vector Laboratories, Newark, CA, USA) for 1 h at room temperature (RT). T24 WT cells were used as negative controls, whereas T-antigen-positive T24 GCNT1 KO cells were used as positive controls [60]. 4',6'-diamidino-2-phenylindole dihydrochloride (DAPI, 2.3×10^{-3} $\mu\text{g}/\mu\text{L}$; Thermo Scientific, Waltham, MA, USA) was used as nuclear counterstain. All fluorescence images were acquired on a Leica DMI6000 FFW microscope using Las X software (version 3.0) (Leica, Wetzlar, Germany).

4.4. Protein Extraction

Proteins were extracted from 100 mg of stool samples using a lysis buffer solution (50 mM Tris at pH 8, 150 mM NaCl, 1% NP40, 0.5% deoxycholate, and 0.1% SDS) to maximize protein yields. Proteins were supplemented with a cocktail of protease inhibitors in order to maintain protein integrity. After incubation for 20 min on ice, samples were subsequently centrifuged at $2000 \times g$ for 10 min at 4 °C and supernatant was collected. Proteins were further precipitated with six volumes of acetone and incubated overnight (16 h) at -20 °C. Then, the mixture was centrifuged at $14,000 \times g$ at 4 °C for 15 min, the supernatant was discarded, and the resulting pellet was air-dried and resuspended by pipetting up and down with the lysis buffer solution supplemented with protease and phosphatase inhibitors. To improve the solubility, samples were subjected to ultrasonic bath cycles at 40 °C and ice cycles. SDS-Page was used to control protein integrity. Finally, the samples were stored at -20 °C for downstream application. Protein samples were quantified using the DC Protein Assay kit.

4.5. Lectin Blotting

Stool proteins (50 μg) were transferred in duplicates onto a nitrocellulose membrane pore size 0.45 μm (AmershamTMProtranTM, Cytiva, Marlborough, MA, USA), using a Hybri-slot apparatus (21052-014; Gibco BRL, Life Technologies, Waltham, MA, USA) under a stabilized vacuum of -0.2 bar. Proteins extracted from T-antigen expressing T24 GCNT1 KO cells were used as positive controls, whereas BSA was used as a negative control. Protein loads were determined by Ponceau S. Nonspecific bindings were blocked with a Carbo-Free Blocking Solution for 1.5 h at RT. Subsequently, the T-antigen expression was evaluated using biotinylated PNA (VectorLabs, Newark, CA, USA) at a dilution of 1:10,000 in a mixture of 50% CarboFree and 50% PBS 1x for 1 h at RT. Membranes were washed with PBS-T (PSB with 1% Tween 20) and incubated with the VECTASTAIN[®] Elite ABD-HRP Reagent (1:10; Vector Laboratories, Newark, CA, USA) for 15 min at RT. The Amersham ECL Prime Detection Reagent was used as a developing reagent. Data analysis was performed through Image Lab Software (version 6.0.1) (Bio-Rad, Hercules, CA, USA) in a ChemiDoc XRS (Bio-Rad, CA, USA).

4.6. (Glyco)Proteomics

The stool samples were selected based on the positive results in lectin blotting, except for the NL group, which was chosen randomly due to negative T-antigen expression. Stool proteins were denatured at 80 °C for 10 min, reduced with 5 mM dithiothreitol (20 μL) for 1 h at 60 °C, and alkylated with 10 mM iodoacetamide (20 μL) for 30 min at RT. The reaction was quenched with 2.5 mM DTT (20 μL). After that, proteins were N-deglycosylated with PNGase F (1 U/10 μg) overnight (16 h) at 37 °C, followed by incubation at 95 °C for 20 min. Then, protein samples were enriched for glycoproteins carrying T-antigen by lectin affinity chromatography with PNA. Eluates and flow-through for complementary proteomic analysis were digested with trypsin overnight (16 h) at 37 °C. Finally, samples acidified with 0.1% TFA were desalted using peptide desalting spin columns following the manufacturer's protocol. A nanoLC (Vanquish neoUHPLC) was coupled online to a Qexactive mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). Eluent A was aqueous formic acid (0.1%), and eluent B was 0.1% formic acid in 80% acetonitrile. Samples were injected into a trapping column (C18 PepMap 100, 5 μm particle size) and separated in the analytical column (EASY-Spray C18 PepMap, 100 Å, 150 mm \times 75 μm I.D).

and 3 μm particle size) at a flow rate of 0.25 $\mu\text{L}/\text{min}$. The column temperature was set at 35 °C. Glycopeptide separation occurred using a linear gradient of 12–46% eluent B over 50 min for T-antigen-enriched fraction and 80 min for proteomic fraction. Column wash and re-equilibration were warranted before the following injection. The mass spectrometer was operated in the positive ion mode, with an m/z range from 300 to 2000, a spray voltage of 1.9 kV, and a transfer capillary temperature of 275 °C. Q-Exactive Plus settings were a full scan resolution of 140,000, a fragment scan resolution of 17,500, a fragment scan fixed first mass at 110 m/z , and an isolation window of 4.0 m/z . Stepped normalized collision energy (nce) was set to 32, 37, and 41% for the T-antigen-enriched fraction, while nce of 30% was used for proteomic flow-through. Mass spectrometry data were processed using the SequestHT search engine and the Percolator algorithm (Proteome Discoverer 3.0, Thermo Fisher Scientific, Waltham, MA, USA) to validate protein identifications. Data were searched against the human proteome from the SwissProt database. Trypsin was selected as the digestion enzyme, considering up to 3 missed cleavage sites, a precursor ion mass tolerance of 10 ppm, and a product ion tolerance of 0.02 Da. Fixed and variable modifications included carbamidomethylcysteine (+57.021 Da) and oxidation of methionine (+15.995 Da), respectively. In addition, modification of serine and threonine with Hex(1)HexNAc(1) (+365.132 Da; core 1/T-antigen) was also allowed. The final protein list was composed of species identified with high confidence.

4.7. Bioinformatics

The whole proteome and PNA-enriched glycoproteins from the four groups addressed in this study (healthy, low-grade dysplasia, high-grade dysplasia, CRC) were characterized according to their biological functions using Cytoscape software (version 3.10.1) [61].

4.8. Statistical Analysis

Statistical Package for Social Sciences—SPSS for MacOS (version 27; IBM, Chicago, IL, USA) and GraphPad Prism software (version 9; Dotmatics, San Diego, CA, USA) were used to perform statistical analysis and data visualization. Statistical analysis comprised the Chi-square Test for categorical variables. A p -value under 0.05 was considered statistically significant.

5. Conclusions

This work presents insights into the unexploited potential of cancer-associated glycans and the stool glycoproteome for cancer screening. It demonstrates that stools from more aggressive lesions may exhibit abnormal immature glycan signatures not seen in healthy tissues, laying the foundation for future exploitation in non-invasive cancer diagnostics. Additionally, it provides a characterization of the stool glycoproteome, identifying several molecular targets shared by more aggressive non-cancerous lesions and cancer, which hold promise for future pre-clinical assessment. Nevertheless, we emphasize the exploratory nature of this study, which may limit the robustness of the results. Future efforts should involve screening larger patient cohorts for both the T-antigen and T-expressing glycoproteoforms, ideally in multicenter settings. The analysis of patients with negative FIT results is also required to fully disclose the potential of these biomarkers for non-invasive CRC screening. Finally, this work showcases the unexplored potential of the CRC glycode for non-invasive disease detection, providing a framework for addressing other potential relevant alterations in glycosylation in this context.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25073722/s1>.

Author Contributions: Conceptualization, D.F., M.D.-R., L.L. and J.A.F.; methodology, J.S., D.F., L.L. and J.A.F.; software, J.S., D.F., M.R.-S. and J.A.F.; validation, J.S., D.F., D.A.R.S., M.R.-S., B.S., L.P.A., L.L.S., L.L. and J.A.F.; formal analysis, J.S., M.E., D.F., D.A.R.S., B.S., M.G., L.L. and J.A.F.; investigation, J.S., M.E., D.F., D.A.R.S., B.S., M.G., E.F., R.V. and J.A.F.; resources, L.P.A., R.V.,

L.L.S., M.D.-R., L.L. and J.A.F.; data curation, J.S., M.E., D.F., D.A.R.S., M.R.-S., B.S., L.L. and J.A.F.; writing—original draft preparation, J.A.F.; writing—review and editing, J.S., M.E., D.F., D.A.R.S., L.L.S., M.D.-R. and L.L.; visualization, J.S., M.E., D.F., D.A.R.S. and M.R.-S.; supervision, L.L. and J.A.F.; project administration, L.L. and J.A.F.; funding acquisition, L.L.S., M.D.-R., L.L. and J.A.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work received financial support from PT national funds (FCT/MCTES, Fundação para a Ciência e Tecnologia and Ministério da Ciência, Tecnologia e Ensino Superior) through CI-IPOP project UIDP/00776/2020, FCT projects RESOLVE (<https://doi.org/10.54499/PTDC/MED-OUT/2512/2021>; accessed on 1 March 2024) and BioFaec (<http://doi.org/10.54499/2022.09051.PTDC>, accessed on 1 March 2024). This work was cofounded by the project “The Porto Comprehensive Cancer Center Raquel Seruca” with the reference NORTE-01-0145-FEDER-0726678—Consorcio PORTO. CCC—Porto.Comprehensive Cancer Center Raquel Seruca, supported by Norte Portugal Regional Operational Programme (NORTE 2020) under the PORTUGAL 2020 Partnership Agreement through the European Regional Development Fund (ERDF), Portugal. FCT is co-financed by the European Social Fund (ESF) under the Human Potential Operation Programme (POPH) from the National Strategic Reference Framework (NSRF). The authors also thank the financial support of the Portuguese Oncology Institute of Porto—Research Centre (CI-IPOP-29–2017–2024). The authors also wish to acknowledge funding from FCT for researcher positions CEECIND/03186/2017 and 2022.08311.CEECIND (JAF), and PhD grants SFRH/BD/142479/2018 and COVID/BD/153011/2023 (JS), 2020.09384.BD (DF), UI/BD/154820/2023 (DARS), SFRH/BD/146500/2019 (MRS), and BS would like to acknowledge the LPCC-NRN 2022 grant from the Portuguese League Against Cancer. LL acknowledges the junior researcher position DL 57/2016/CP1491/CT0002 from CI-IPOP.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and received approval from the Institutional Ethics Committee of IPO-Porto (Approval No. 64/2017 and 22/022, dated from 16 March 2017 and 10 February 2022, respectively).

Informed Consent Statement: Informed consent was obtained from all subjects enrolled in the study.

Data Availability Statement: The data generated in this study are included in the main article or Supplementary Material.

Acknowledgments: The authors would like to acknowledge all the individuals recruited at IPO-Porto who kindly accepted the invitation to participate in this study.

Conflicts of Interest: J.A.F. and L.L.S. are founders and associate partners of GlycoMatters Biotech. The remaining authors declare no conflict of interest.

References

1. Schreuders, E.H.; Ruco, A.; Rabeneck, L.; Schoen, R.E.; Sung, J.J.; Young, G.P.; Kuipers, E.J. Colorectal cancer screening: A global overview of existing programmes. *Gut* **2015**, *64*, 1637–1649. [[CrossRef](#)] [[PubMed](#)]
2. Simon, K. Colorectal cancer development and advances in screening. *Clin. Interv. Aging*. **2016**, *11*, 967–976. [[PubMed](#)]
3. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 Cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
4. Young, G.P.; Rabeneck, L.; Winawer, S.J. The Global Paradigm Shift in Screening for Colorectal Cancer. *Gastroenterology* **2019**, *156*, 843–851.e2. [[CrossRef](#)] [[PubMed](#)]
5. Stracci, F.; Zorzi, M.; Grazzini, G. Colorectal cancer screening: Tests, strategies, and perspectives. *Front. Public Health* **2014**, *2*, 210. [[CrossRef](#)]
6. Robertson, D.J.; Lee, J.K.; Boland, C.R.; Dominitz, J.A.; Giardiello, F.M.; Johnson, D.A.; Kaltenbach, T.; Lieberman, D.; Levin, T.R.; Rex, D.K. Recommendations on Fecal Immunochemical Testing to Screen for Colorectal Neoplasia: A Consensus Statement by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology* **2017**, *152*, 1217–1237.e3. [[CrossRef](#)] [[PubMed](#)]
7. Beniwal, S.S.; Lamo, P.; Kaushik, A.; Lorenzo-Villegas, D.L.; Liu, Y.; MohanaSundaram, A. Current Status and Emerging Trends in Colorectal Cancer Screening and Diagnostics. *Biosensors* **2023**, *13*, 926. [[CrossRef](#)] [[PubMed](#)]
8. Morikawa, T.; Kato, J.; Yamaji, Y.; Wada, R.; Mitsushima, T.; Shiratori, Y. A comparison of the immunochemical fecal occult blood test and total colonoscopy in the asymptomatic population. *Gastroenterology* **2005**, *129*, 422–428. [[CrossRef](#)] [[PubMed](#)]
9. Areia, M.; Fuccio, L.; Hassan, C.; Dekker, E.; Dias-Pereira, A.; Dinis-Ribeiro, M. Cost-utility analysis of colonoscopy or faecal immunochemical test for population-based organised colorectal cancer screening. *United Eur. Gastroenterol. J.* **2019**, *7*, 105–113. [[CrossRef](#)]

10. Saftoiu, A.; Hassan, C.; Areia, M.; Bhutani, M.S.; Bisschops, R.; Bories, E.; Cazacu, I.M.; Dekker, E.; Deprez, P.H.; Pereira, S.P.; et al. Role of gastrointestinal endoscopy in the screening of digestive tract cancers in Europe: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. *Endoscopy* **2020**, *52*, 293–304. [[CrossRef](#)]
11. Das, V.; Kalita, J.; Pal, M. Predictive and prognostic biomarkers in colorectal cancer: A systematic review of recent advances and challenges. *Biomed. Pharmacother.* **2017**, *87*, 8–19. [[CrossRef](#)] [[PubMed](#)]
12. Holst, S.; Wuhrer, M.; Rombouts, Y. Glycosylation characteristics of colorectal cancer. *Adv. Cancer Res.* **2015**, *126*, 203–256. [[PubMed](#)]
13. Madunic, K.; Mayboroda, O.A.; Zhang, T.; Weber, J.; Boons, G.J.; Morreau, H.; van Vlierberghe, R.; van Wezel, T.; Lageveen-Kammeijer, G.S.M.; Wuhrer, M. Specific (sialyl-)Lewis core 2 O-glycans differentiate colorectal cancer from healthy colon epithelium. *Theranostics* **2022**, *12*, 4498–4512. [[CrossRef](#)] [[PubMed](#)]
14. Madunic, K.; Zhang, T.; Mayboroda, O.A.; Holst, S.; Stavenhagen, K.; Jin, C.; Karlsson, N.G.; Lageveen-Kammeijer, G.S.M.; Wuhrer, M. Colorectal cancer cell lines show striking diversity of their O-glycome reflecting the cellular differentiation phenotype. *Cell Mol. Life Sci.* **2021**, *78*, 337–350. [[CrossRef](#)]
15. Cao, Y.; Stosiek, P.; Springer, G.F.; Karsten, U. Thomsen-Friedenreich-related carbohydrate antigens in normal adult human tissues: A systematic and comparative study. *Histochem. Cell Biol.* **1996**, *106*, 197–207. [[CrossRef](#)]
16. Cao, Y.; Karsten, U.R.; Liebrich, W.; Haensch, W.; Springer, G.F.; Schlag, P.M. Expression of Thomsen-Friedenreich-related antigens in primary and metastatic colorectal carcinomas. A reevaluation. *Cancer* **1995**, *76*, 1700–1708. [[CrossRef](#)]
17. Li, F.; Glinkii, O.V.; Mooney, B.P.; Rittenhouse-Olson, K.; Pienta, K.J.; Glinksky, V.V. Cell surface Thomsen-Friedenreich proteome profiling of metastatic prostate cancer cells reveals potential link with cancer stem cell-like phenotype. *Oncotarget* **2017**, *8*, 98598–98608. [[CrossRef](#)] [[PubMed](#)]
18. Mereiter, S.; Polom, K.; Williams, C.; Polonia, A.; Guergova-Kuras, M.; Karlsson, N.G.; Roviello, F.; Magalhães, A.; Reis, C.A. The Thomsen-Friedenreich Antigen: A Highly Sensitive and Specific Predictor of Microsatellite Instability in Gastric Cancer. *J. Clin. Med.* **2018**, *7*, 256. [[CrossRef](#)] [[PubMed](#)]
19. Srinivasan, N.; Bane, S.M.; Ahire, S.D.; Ingle, A.D.; Kalraiya, R.D. Poly N-acetyllactosamine substitutions on N- and not O-oligosaccharides or Thomsen-Friedenreich antigen facilitate lung specific metastasis of melanoma cells via galectin-3. *Glycoconj. J.* **2009**, *26*, 445–456. [[CrossRef](#)]
20. Takanami, I. Expression of Thomsen-Friedenreich antigen as a marker of poor prognosis in pulmonary adenocarcinoma. *Oncol. Rep.* **1999**, *6*, 341–344. [[CrossRef](#)]
21. Santos-Silva, F.; Fonseca, A.; Caffrey, T.; Carvalho, F.; Mesquita, P.; Reis, C.; Almeida, R.; David, L.; Hollingsworth, M.A. Thomsen-Friedenreich antigen expression in gastric carcinomas is associated with MUC1 mucin VNTR polymorphism. *Glycobiology* **2005**, *15*, 511–517. [[CrossRef](#)] [[PubMed](#)]
22. Trabbic, K.R.; Whalen, K.; Abarca-Heideman, K.; Xia, L.; Temme, J.S.; Edmondson, E.F.; Gildersleeve, J.C.; Barchi, J.J., Jr. A Tumor-Selective Monoclonal Antibody from Immunization with a Tumor-Associated Mucin Glycopeptide. *Sci. Rep.* **2019**, *9*, 5662. [[CrossRef](#)] [[PubMed](#)]
23. Peixoto, A.; Ferreira, D.; Azevedo, R.; Freitas, R.; Fernandes, E.; Relvas-Santos, M.; Gaitero, C.; Soares, J.; Cotton, S.; Teixeira, B.; et al. Glycoproteomics identifies HOMER3 as a potentially targetable biomarker triggered by hypoxia and glucose deprivation in bladder cancer. *J. Exp. Clin. Cancer Res.* **2021**, *40*, 191. [[CrossRef](#)] [[PubMed](#)]
24. Bojar, D.; Meche, L.; Meng, G.; Eng, W.; Smith, D.F.; Cummings, R.D.; Mahal, L.K. A Useful Guide to Lectin Binding: Machine-Learning Directed Annotation of 57 Unique Lectin Specificities. *ACS Chem. Biol.* **2022**, *17*, 2993–3012. [[CrossRef](#)] [[PubMed](#)]
25. Wang, L.; Xu, H.; Yang, H.; Zhou, J.; Zhao, L.; Zhang, F. Glucose metabolism and glycosylation link the gut microbiota to autoimmune diseases. *Front. Immunol.* **2022**, *13*, 952398. [[CrossRef](#)] [[PubMed](#)]
26. Leao, B.; Wen, X.; Duarte, H.O.; Gullo, I.; Goncalves, G.; Pontes, P.; Castelli, C.; Diniz, F.; Mereiter, S.; Gomes, J.; et al. Expression of Thomsen-Friedenreich Antigen in Colorectal Cancer and Association with Microsatellite Instability. *Int. J. Mol. Sci.* **2021**, *22*, 1340. [[CrossRef](#)] [[PubMed](#)]
27. Cayrefourcq, L.; Thomas, F.; Mazard, T.; Assenat, E.; Assou, S.; Alix-Panabieres, C. Selective treatment pressure in colon cancer drives the molecular profile of resistant circulating tumor cell clones. *Mol. Cancer* **2021**, *20*, 30. [[CrossRef](#)] [[PubMed](#)]
28. Chung, Y.S.; Yamashita, Y.; Kato, Y.; Nakata, B.; Sawada, T.; Sowa, M. Prognostic significance of T antigen expression in patients with gastric carcinoma. *Cancer* **1996**, *77*, 1768–1773. [[CrossRef](#)]
29. Springer, G.F.; Desai, P.R.; Murthy, M.S.; Scanlon, E.F. Human carcinoma-associated precursor antigens of the NM blood group system. *J. Surg. Oncol.* **1979**, *11*, 95–106. [[CrossRef](#)]
30. Yu, L.G.; Andrews, N.; Zhao, Q.; McKean, D.; Williams, J.F.; Connor, L.J.; Gerasimenko, O.V.; Hilkens, J.; Hirabayashi, J.; Kasai, K.; et al. Galectin-3 interaction with Thomsen-Friedenreich disaccharide on cancer-associated MUC1 causes increased cancer cell endothelial adhesion. *J. Biol. Chem.* **2007**, *282*, 773–781. [[CrossRef](#)]
31. Khaldoyanidi, S.K.; Glinksky, V.V.; Sikora, L.; Glinkii, A.B.; Mossine, V.V.; Quinn, T.P.; Glinksky, G.V.; Sriramarao, P. MDA-MB-435 human breast carcinoma cell homo- and heterotypic adhesion under flow conditions is mediated in part by Thomsen-Friedenreich antigen-galectin-3 interactions. *J. Biol. Chem.* **2003**, *278*, 4127–4134. [[CrossRef](#)] [[PubMed](#)]
32. Cotton, S.; Azevedo, R.; Gaitero, C.; Ferreira, D.; Lima, L.; Peixoto, A.; Fernandes, E.; Neves, M.; Neves, D.; Amaro, T.; et al. Targeted O-glycoproteomics explored increased sialylation and identified MUC16 as a poor prognosis biomarker in advanced-stage bladder tumours. *Mol. Oncol.* **2017**, *11*, 895–912. [[CrossRef](#)] [[PubMed](#)]

33. Juge, N.; Tailford, L.; Owen, C.D. Sialidases from gut bacteria: A mini-review. *Biochem. Soc. Trans.* **2016**, *44*, 166–175. [[CrossRef](#)] [[PubMed](#)]
34. Gupta, P.; Tripathy, A.S. Alternative pathway of complement activation has a beneficial role against Chandipura virus infection. *Med. Microbiol. Immunol.* **2020**, *209*, 109–124. [[CrossRef](#)] [[PubMed](#)]
35. Lee, J.W.J.; Plichta, D.R.; Asher, S.; Delsignore, M.; Jeong, T.; McGoldrick, J.; Staller, K.; Khalili, H.; Xavier, R.J.; Chung, D.C. Association of distinct microbial signatures with premalignant colorectal adenomas. *Cell Host Microbe* **2023**, *31*, 827–838.e3. [[CrossRef](#)] [[PubMed](#)]
36. Feng, Q.; Liang, S.; Jia, H.; Stadlmayr, A.; Tang, L.; Lan, Z.; Zhang, D.; Xia, H.; Xu, X.; Jie, Z.; et al. Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat. Commun.* **2015**, *6*, 6528. [[CrossRef](#)] [[PubMed](#)]
37. Radak, M.; Fallahi, H. Unraveling molecular similarities between colorectal polyps and colorectal cancer: A systems biology approach. *Intest. Res.* **2024**. [[CrossRef](#)] [[PubMed](#)]
38. Wong, C.C.; Yu, J. Gut microbiota in colorectal cancer development and therapy. *Nat. Rev. Clin. Oncol.* **2023**, *20*, 429–452. [[CrossRef](#)] [[PubMed](#)]
39. Darebna, P.; Novak, P.; Kucera, R.; Topolcan, O.; Sanda, M.; Goldman, R.; Pompach, P. Changes in the expression of N- and O-glycopeptides in patients with colorectal cancer and hepatocellular carcinoma quantified by full-MS scan FT-ICR and multiple reaction monitoring. *J. Proteom.* **2017**, *153*, 44–52. [[CrossRef](#)]
40. Ivancic, M.M.; Megna, B.W.; Sverchikov, Y.; Craven, M.; Reichelderfer, M.; Pickhardt, P.J.; Sussman, M.R.; Kennedy, G.D. Noninvasive Detection of Colorectal Carcinomas Using Serum Protein Biomarkers. *J. Surg. Res.* **2020**, *246*, 160–169. [[CrossRef](#)]
41. Fiorito, V.; Tolosano, E. Hemopexin and Cancer. *Int. J. Mol. Sci.* **2022**, *23*, 160–169. [[CrossRef](#)]
42. Bosch, L.J.W.; de Wit, M.; Pham, T.V.; Coupe, V.M.H.; Hiemstra, A.C.; Piersma, S.R.; Oudgenoeg, G.; Scheffer, G.L.; Mongera, S.; Sive Droste, J.T.; et al. Novel Stool-Based Protein Biomarkers for Improved Colorectal Cancer Screening: A Case-Control Study. *Ann. Intern. Med.* **2017**, *167*, 855–866. [[CrossRef](#)]
43. Komor, M.A.; Bosch, L.J.; Coupe, V.M.; Rausch, C.; Pham, T.V.; Piersma, S.R.; Mongera, S.; Mulder, C.J.J.; Dekker, E.; Kuipers, E.J.; et al. Proteins in stool as biomarkers for non-invasive detection of colorectal adenomas with high risk of progression. *J. Pathol.* **2020**, *250*, 288–298. [[CrossRef](#)]
44. Canesin, G.; Di Ruscio, A.; Li, M.; Ummano, S.; Hedblom, A.; Choudhury, R.; Krzyzanowska, A.; Csizmadia, E.; Palominos, M.; Stiehm, A.; et al. Scavenging of Labile Heme by Hemopexin Is a Key Checkpoint in Cancer Growth and Metastases. *Cell Rep.* **2020**, *32*, 108181. [[CrossRef](#)] [[PubMed](#)]
45. Suzuki, Y.; Takadate, T.; Mizuma, M.; Shima, H.; Suzuki, T.; Tachibana, T.; Shimura, M.; Hata, T.; Iseki, M.; Kawaguchi, K.; et al. Stromal expression of hemopexin is associated with lymph-node metastasis in pancreatic ductal adenocarcinoma. *PLoS ONE* **2020**, *15*, e0235904. [[CrossRef](#)]
46. Miyamoto, S.; Ruhaak, L.R.; Stroble, C.; Salemi, M.R.; Phinney, B.; Lebrilla, C.B.; Leiserowitz, G.S. Glycoproteomic Analysis of Malignant Ovarian Cancer Ascites Fluid Identifies Unusual Glycopeptides. *J. Proteome Res.* **2016**, *15*, 3358–3376. [[CrossRef](#)] [[PubMed](#)]
47. Chantaraamporn, J.; Champattanachai, V.; Khongmanee, A.; Verathamjamras, C.; Prasongsook, N.; Mingkwan, K.; Luevisadpibul, V.; Chutipongtanate, S.; Svasti, J. Glycoproteomic Analysis Reveals Aberrant Expression of Complement C9 and Fibronectin in the Plasma of Patients with Colorectal Cancer. *Proteomes* **2020**, *8*, 26. [[CrossRef](#)]
48. Bao, D.; Zhang, C.; Li, L.; Wang, H.; Li, Q.; Ni, L.; Lin, Y.; Huang, R.; Yang, Z.; Zhang, Y.; et al. Integrative Analysis of Complement System to Prognosis and Immune Infiltrating in Colon Cancer and Gastric Cancer. *Front. Oncol.* **2020**, *10*, 553297. [[CrossRef](#)] [[PubMed](#)]
49. Bohana-Kashtan, O.; Pinna, L.A.; Fishelson, Z. Extracellular phosphorylation of C9 by protein kinase CK2 regulates complement-mediated lysis. *Eur. J. Immunol.* **2005**, *35*, 1939–1948. [[CrossRef](#)]
50. Li, L.; Yang, H.; Li, Y.; Li, X.D.; Zeng, T.T.; Lin, S.X.; Zhu, Y.H.; Guan, X.Y. Hypoxia restrains the expression of complement component 9 in tumor-associated macrophages promoting non-small cell lung cancer progression. *Cell Death Discov.* **2018**, *4*, 63. [[CrossRef](#)]
51. Cao, Z.Q.; Guo, X.L. The role of galectin-4 in physiology and diseases. *Protein Cell.* **2016**, *7*, 314–324. [[CrossRef](#)] [[PubMed](#)]
52. Danielsen, E.M.; Hansen, G.H. Lipid raft organization and function in brush borders of epithelial cells. *Mol. Membr. Biol.* **2006**, *23*, 71–79. [[CrossRef](#)]
53. Barrow, H.; Rhodes, J.M.; Yu, L.G. Simultaneous determination of serum galectin-3 and -4 levels detects metastases in colorectal cancer patients. *Cell Oncol.* **2013**, *36*, 9–13. [[CrossRef](#)] [[PubMed](#)]
54. Barrow, H.; Guo, X.; Wandall, H.H.; Pedersen, J.W.; Fu, B.; Zhao, Q.; Chen, C.; Rhodes, J.M.; Yu, L.G. Serum galectin-2, -4, and -8 are greatly increased in colon and breast cancer patients and promote cancer cell adhesion to blood vascular endothelium. *Clin. Cancer Res.* **2011**, *17*, 7035–7046. [[CrossRef](#)]
55. Kume, H.; Muraoka, S.; Kuga, T.; Adachi, J.; Narumi, R.; Watanabe, S.; Kuwano, M.; Kodera, Y.; Matsushita, K.; Fukuoka, J.; et al. Discovery of colorectal cancer biomarker candidates by membrane proteomic analysis and subsequent verification using selected reaction monitoring (SRM) and tissue microarray (TMA) analysis. *Mol. Cell Proteom.* **2014**, *13*, 1471–1484. [[CrossRef](#)]
56. Furuya, K.; Nakajima, M.; Tsunedomi, R.; Nakagami, Y.; Xu, M.; Matsui, H.; Tokumitsu, Y.; Shindo, Y.; Watanabe, Y.; Tomochika, S.; et al. High serum proteinase-3 levels predict poor progression-free survival and lower efficacy of bevacizumab in metastatic colorectal cancer. *BMC Cancer* **2024**, *24*, 165. [[CrossRef](#)]

57. Kalimutho, M.; Blanco Gdel, V.; Gravina, P.; Cretella, M.; Mannucci, L.; Mannisi, E.; Formosa, A.; Pallone, F.; Federici, G.; Bernardini, S. Quantitative denaturing high performance liquid chromatography (Q-dHPLC) detection of APC long DNA in faeces from patients with colorectal cancer. *Clin. Chem. Lab. Med.* **2010**, *48*, 1303–1311. [[CrossRef](#)]
58. Phua, L.C.; Chue, X.P.; Koh, P.K.; Cheah, P.Y.; Chan, E.C.; Ho, H.K. Global fecal microRNA profiling in the identification of biomarkers for colorectal cancer screening among asians. *Oncol. Rep.* **2014**, *32*, 97–104. [[CrossRef](#)] [[PubMed](#)]
59. Gaitero, C.; Soares, J.; Relvas-Santos, M.; Peixoto, A.; Ferreira, D.; Paulo, P.; Brandão, A.; Fernandes, E.; Azevedo, R.; Palmeira, C.; et al. Glycoproteogenomics characterizes the CD44 splicing code associated with bladder cancer invasion. *Theranostics* **2022**, *12*, 3150–3177. [[CrossRef](#)]
60. Peixoto, A.; Ferreira, D.; Miranda, A.; Relvas-Santos, M.; Freitas, R.; Veth, T.S.; Brandão, A.; Ferreira, E.; Paulo, P.; Cardoso, M.; et al. Multilevel Plasticity and Altered Glycosylation Drive Aggressiveness in Hypoxic and Glucose-Deprived Bladder Cancer Cells. *bioRxiv* **2023**. [[CrossRef](#)]
61. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.