



Review

Insights into Bioinformatic Applications for Glycosylation: Instigating an Awakening towards Applying Glycoinformatic Resources for Cancer Diagnosis and Therapy

Manikandan Muthu ¹, Sechul Chun ¹ , Judy Gopal ¹ , Vimala Anthonydhasan ², Steve W. Haga ³, Anna Jacintha Prameela Devadoss ¹ and Jae-Wook Oh ^{4,*}

¹ Department of Environmental Health Sciences, Konkuk University, Seoul 143-701, Korea; bhagatmani@gmail.com (M.M.); scchun@konkuk.ac.kr (S.C.); jejudy777@gmail.com (J.G.); awesomeanna1@gmail.com (A.J.P.D.)

² Department of Microbiology and Immunology, Institute for Biomedicine, Gothenburg University, 413 90 Gothenburg, Sweden; vimalalisha@gmail.com

³ Department of Computer Science and Engineering, National Sun Yat Sen University, Kaohsiung 804, Taiwan; stevewhaga@yahoo.com

⁴ Department of Stem Cell and Regenerative Biotechnology, Konkuk University, Seoul 143-701, Korea

* Correspondence: ohjw@konkuk.ac.kr; Tel.: +82-2-2049-6271; Fax: +82-2-455-1044

Received: 29 October 2020; Accepted: 1 December 2020; Published: 8 December 2020



Abstract: Glycosylation plays a crucial role in various diseases and their etiology. This has led to a clear understanding on the functions of carbohydrates in cell communication, which eventually will result in novel therapeutic approaches for treatment of various disease. Glycomics has now become one among the top ten technologies that will change the future. The direct implication of glycosylation as a hallmark of cancer and for cancer therapy is well established. As in proteomics, where bioinformatics tools have led to revolutionary achievements, bioinformatics resources for glycosylation have improved its practical implication. Bioinformatics tools, algorithms and databases are a mandatory requirement to manage and successfully analyze large amount of glycobiological data generated from glycosylation studies. This review consolidates all the available tools and their applications in glycosylation research. The achievements made through the use of bioinformatics into glycosylation studies are also presented. The importance of glycosylation in cancer diagnosis and therapy is discussed and the gap in the application of widely available glyco-informatic tools for cancer research is highlighted. This review is expected to bring an awakening amongst glyco-informaticians as well as cancer biologists to bridge this gap, to exploit the available glyco-informatic tools for cancer.

Keywords: glycosylation; cancer; bioinformatics tools; databases; post translational modification; proteins

1. Introduction

Post translational modification (PTM) is the modification of proteins subsequent to protein biosynthesis. Proteins, undergo PTM to form the anticipated protein product. There are various forms of PTMs, namely hydroxylation, glycosylation, phosphorylation, methylation, ubiquitination, SUMOylation, acetylation and lipidation (Figure 1). Glycosylation is associated with a network of critical biological processes. The process involves attachment of carbohydrate molecules to lipids and protein moieties. Monosaccharides that are covalently linked by glycosidic bonds, are called glycans. N-glycosylation and O-glycosylation are the two major glycosylation types. C-glycosylation

(or C-mannosylation), glypiation (or glycosylphosphatidylinositol (GPI) anchoring) and glycation (or nonenzymatic glycosylation) also exist [1,2].

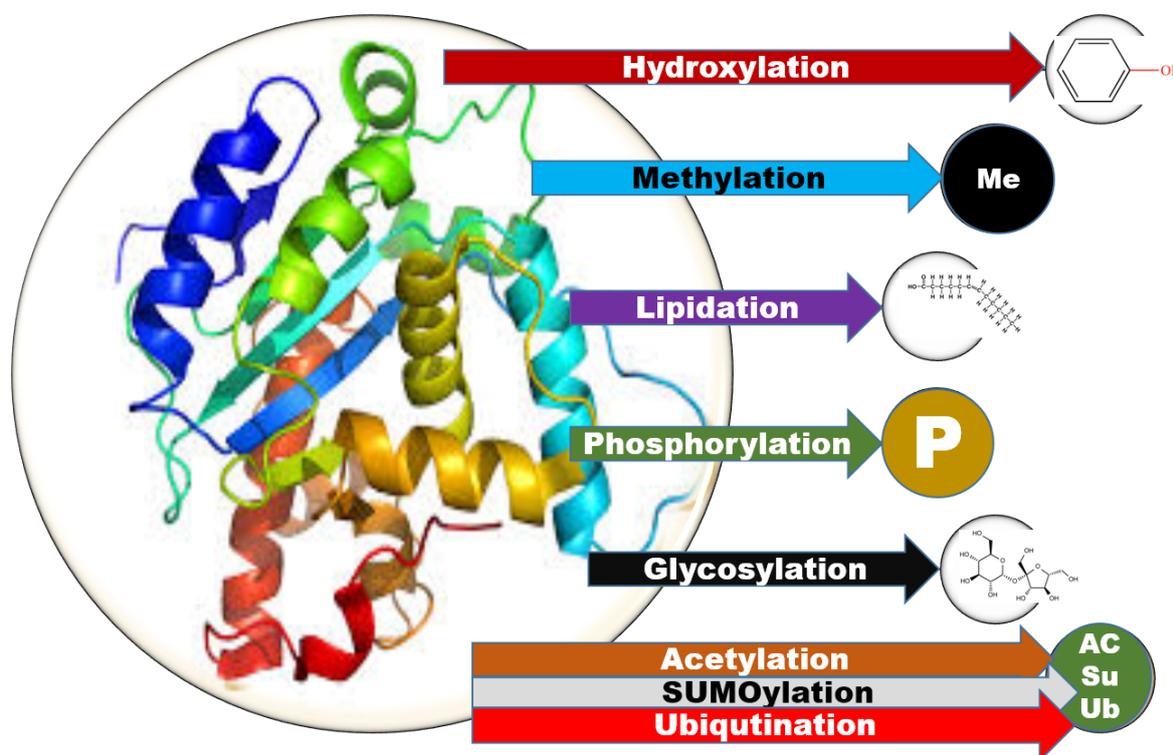


Figure 1. Schematic showing the wide range of post translational modification of proteins of biological significance.

The biological roles of glycosylation include diverse functions and influence protein folding and oligomerization [3], protein degradation [4], epitope recognition [5], protein solubility and stability [6], cell-cell interactions [7] and protein transport [8]. Subsequently, glycoproteins also have a say in cancer development and progression [9], autoimmune diseases [10] and congenital disorders [11]. Tailoring glycosylation for manufacturing biotherapeutics, enhances the safety and efficacy of monoclonal antibodies and erythropoietin. Among all PTMs, glycosylation is established to be the most complex [12]. Biologically active glycan structures are encoded indirectly into the genome and glycosidases, glycosyltransferase and carbohydrate-modifying enzymes help attaching glycan structures to proteins and lipids. [12]. Then the glycoproteins are modified with different carbohydrates resulting in several glycoforms. The recognition of glycan-occupied sites is expensive and laborious and can only be experimentally determined [13]. This is the reason why still the number of verified glycosylated residues are limited as compared to the existing knowledge of the known protein sequences [14]. This knowledge regarding glycosylation site locations is by itself a valuable asset. This knowledge is useful in improving the 3D protein structure prediction, to manipulate protein pharmacokinetic properties through glycoengineering [15]. In this direction, development of bioinformatics tools to predict glycosylation sites gain paramount significance.

Munkley et al. in their review have discussed that aberrant protein glycosylation can be linked with cancer and protein glycoforms as cancer biomarkers. Aberrant glycosylation can have a say in cancer progression, as glycans do on cell adhesion, migration, immune surveillance, interactions with cell matrix, signaling and cell metabolism [16]. Development and progression of cancer is known to result in fundamental changes to the glycome. Aberrant glycosylation has been reported in prostate cancer, where aberrations in glycan composition of prostate cancer cells are linked to disease progression. Glycan variants have been reported to be associated with breast, colon, liver, skin,

ovarian and bladder cancers and neurodegenerative diseases [17–19]. Many glycoproteins show alterations of glycosylation during cancer [20], these can act as glycan biomarkers for various cancer types. MUC-1 (CA15-3/CA27.29) [21] and plasminogen activator inhibitor (PAI-1) [22] have been reported as biomarkers for beta-human chorionic gonadotropin (Beta-hCG); breast cancer; [23] testicular cancer; alpha-fetoprotein (AFP) [24] for liver cancer and germ cell tumors. Chromogranin A (CgA) [25] is a biomarker for neuroendocrine tumors and biomarkers MUC16 (CA-125) [26] and HE4 [27] for ovarian cancer. The hallmarks of glycosylation in cancer are presented in Figure 2.

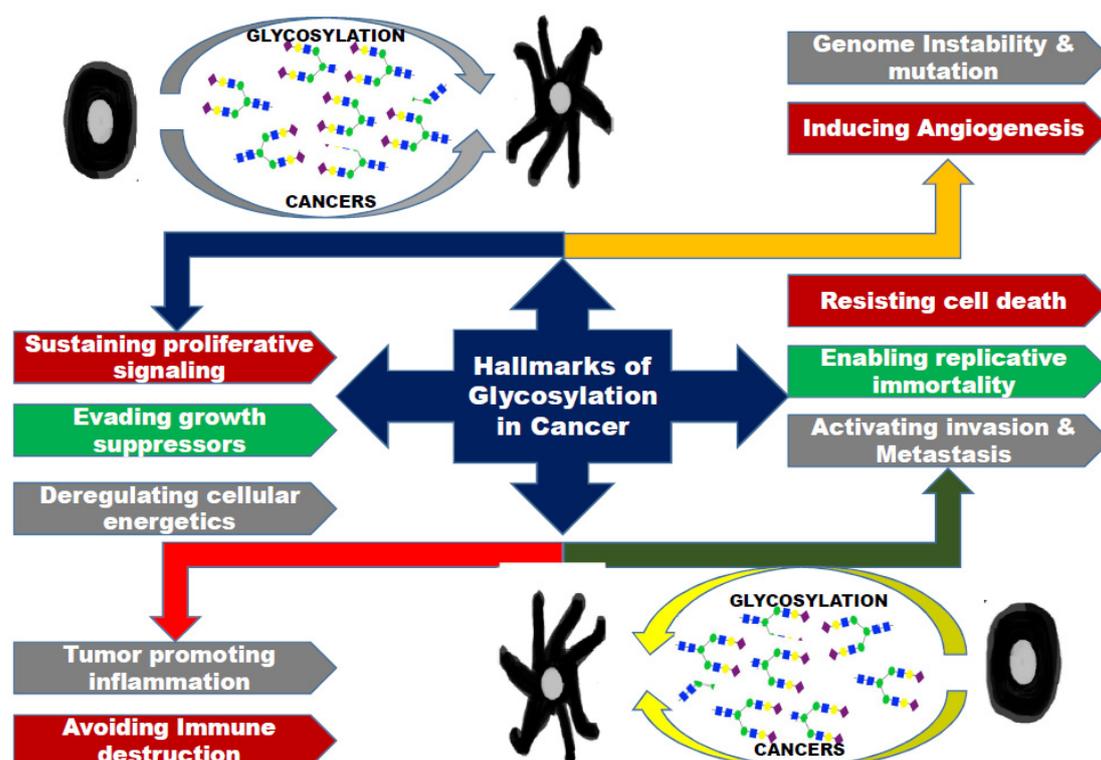


Figure 2. Snapshot of the hallmarks of glycosylation in cancer [15].

Bioinformatics resources for glycomics and glycobiology-related databases has been reviewed [28–34]. The application of bioinformatic tools in biology has led to a fundamental understanding of the biological processes involved [35,36]. Besides the storage, analysis, annotation, organization, systematization needs of genomics, proteomics, and metabolomics data have been met with the emergence and development of bioinformatics. Bioinformatics has provided glycobiologists with databases and tools to support their research. High throughput analysis of glycans requires automated analysis pipelines that bank on extensive bioinformatic support. Various bioinformatic groups are developing mathematical or statistical algorithms and computational methodologies to analyze biological data. Although bioinformatics for glycobiology or glycomics (“glyco-informatics”) [32] is yet to be established compared to genomics and proteomics [37,38], a steady increase in both the development, and use of glyco-informatics tools and databases has been evidenced [28,29,39–47].

This is the first review to look at the use of glycosylation bioinformatics data in cancer research and diagnostics. This review focusses on highlighting the importance of glycosylation in cancer. The achievements made using bioinformatics in the area of glycosylation are reviewed and presented. The scarce inputs available from the use of bioinformatics towards glycosylation with respect to cancer is reviewed for the first time. The future perspective of this integrated approach has been speculated.

2. Snapshot of Bioinformatic Resources for Glycosylation

Bioinformatic resources have become popular in glycan research for automation of analytical outputs and for high throughput processing of data. The analysis of glycans is challenging, this is why glycomics is lagging behind genomic and proteomic research. The robotic N-glycan release and labelling combined with UPLC/HPLC analytical platforms with other bioinformatics tools have enabled high-throughput glycoprofiling. High resolution NMR spectroscopy and mass spectrometry have been reported for glycoprofiling [48]. The intervention of bioinformatics resources in glycan research has impacted the successful characterization of biological samples and resulted in progress in biomarker discovery, and genome-wide association studies. Automation technology is a valuable asset when it comes to analysis, when combined with bioinformatics several orders of magnitude higher high throughput processing has been reported [43]. Figure 3 presents the workflow of glycosylation, beginning from wet lab analysis to analytical instrumentation to computation (bioinformatics).

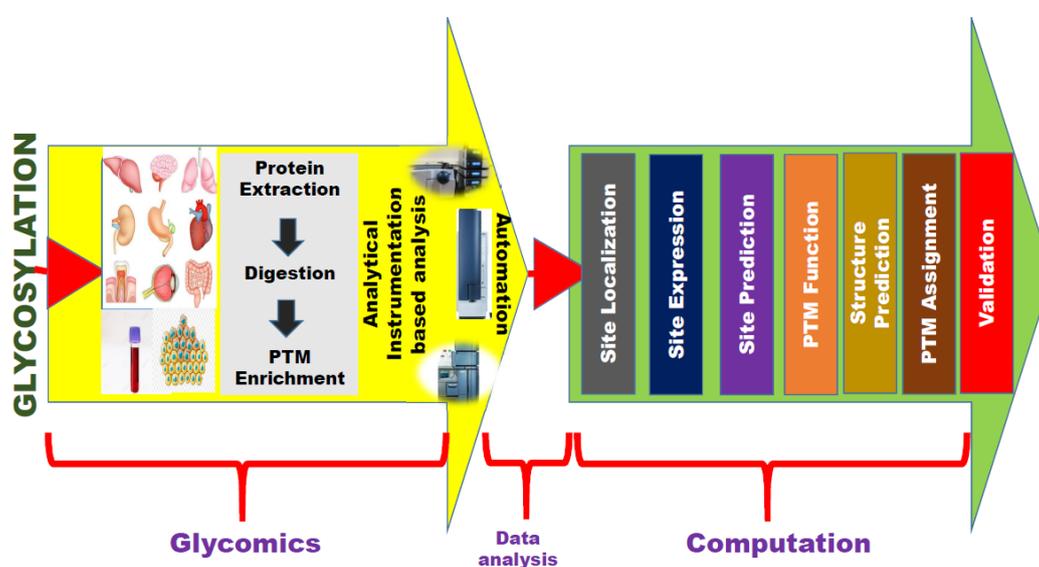


Figure 3. Work-flow of glycosylation research, beginning from wet lab analysis—instrumentational analysis culmination in bioinformatics based computation (data analysis).

Bioinformatics offers a variety of databases to the glycoscientist [30,39,42,49–51]. Glycan informatics has provided varied methods for glycosylation analysis, glycomics and glycan biomarker prediction. Tools are there for prediction of glycosylation binding sites, for statistical analysis of amino acids of a glycoprotein and mathematical modeling of glycosylation. Tools for annotation of mass spectra and for glycan structure analysis and Glyco-Gene Expression Analysis are available for users. Bioinformatics resources for glycoresearch can be grouped as: (i) databases containing information on proteins, (ii) databases storing information on enzymes and pathways building glycans, and (iii) carbohydrate structural databases [42,43]. Not many databases provide information on glycoforms of glycoproteins. The complex carbohydrate structure database (CCSD)/CarbBank was developed and maintained by the Complex Carbohydrate Research Center of the University of Georgia (USA) [52,53]. It is a reservoir of carbohydrate structures gathered through manual retrospective extraction from existing literature and all the publications in which complex carbohydrate structures were cataloged. Owing to lack of funding, the database has been no longer updated. However, holding 50,000 records from 14,000 publications and 23,500 different carbohydrate sequences, it remains as the largest carbohydrate-related repository. Offshoots of CCSD data have resulted in open access databases, such as GLYCOSCIENCES.de [54], KEGG Glycan database [55], CFG Glycan Database [56], Bacterial Carbohydrate Structure Database (BCSDB) [57] and GlycoBase (currently inactive replaced by GlyStore) [58]. The Japan Consortium for Glycobiology and Glycotechnology DataBase (JCGGDB),

which assembles CabosDB, Galaxy, LipidBank, GlycoEpitope, LfDB, and Structural Glycomics Calculation (SGCAL) (1490 carbohydrate structures) and GlycoSuiteDB (3300 unique carbohydrate structures) [59] are freely accessible as well. The EUROCarbDB project aims at developing a new infrastructure, where scientists themselves can upload carbohydrate structure-related data. A prototype that can store carbohydrate structures, has been developed. Primary experimental data (MS, HPLC, and NMR) that serves as reference data for the carbohydrate structure in question can also be uploaded.

Mutual exchange between carbohydrate databases is reported to be a challenging task [60], since various databases use different sequence formats to encode carbohydrate structures [61]. This is why glyco-informatics exists as multiple disconnected and incompatible islands of experimental data, data resources and specific applications, managed by various independent consortia, institutions and research groups [43,62] and no comprehensive database for carbohydrate structures exists. Such isolated islands of information make it difficult for the user to access and derive required information. The difficulty encountered by the users is that they have to visit different database portals to retrieve information on a specific carbohydrate structure and they need to acclimatize themselves with the different local query options, which demand prior knowledge on encoding of the residues in that respective database. In 2005, a comprehensive index of all available structures was initiated by translating the freely available databases to the GlycoCT sequence format [61] and storing in a new database known as GlycomeDB [63]. A web interface was developed as a single query point for all carbohydrate structure databases [64].

A novel bioinformatics method called GlycoMinestruct has been developed. This improves the prediction of human N- and O-linked glycosylation sites by combining sequence and structural features in an integrated computational framework with a two-step feature-selection strategy [65]. *GlycoMinestruct* is established to outperform NGlycPred, which is the only predictor that incorporated both sequence and structure features, achieving high AUC values (in the field of pharmacokinetics, the area under the curve (AUC) defines the variation of a drug concentration in blood plasma as a function of time) for N- and O-linked glycosylation. *GlycoMinestruct* has been used to screen the human proteome and obtain high-confidence predictions for N- and O-linked glycosylation sites and expedited discovery of glycosylation events and substrates facilitating hypothesis-driven experimental studies. There are other glycosylation online predictors available too, which predict N-Glycosylation, O-glycosylation and C-mannosylation sites in the sequence of interest. For example NetNGlyc [66] used for N-Glycosylation site prediction, NetOGlyc [67], YinOYang [68] for O-Glycosylation sites prediction, GPP [69] for N-Glycosylation and O-Glycosylation sites prediction likewise NetCGlyc [70] for C-mannosylation sites prediction) and GlycoMine [71] used for N-, C- and O-linked glycosylation sites prediction. These servers are designed based on different machine learning algorithms like SVM (Support Vector Machine), ANN (Artificial Neural Networks) and RFA (Random Forest Algorithm). OGlycBase [72] is the database of O- and C-glycosylated proteins, contains the experimentally verified O- and C-glycosylated proteins. Whereas the tools like GPI-SOM [73], PredGPI [74] and BIG-P [75] are used for Glycosylphosphatidylinositol modification site (GPI-anchors) prediction. These tools are designed based on SVM and Hidden Markov Model (HMM) algorithms. GlycoFragment and Distance Mapping are also used for glycosylation 3D structure prediction [76].

Bioinformatic tools have also been able to enhance the prediction of glycosylation sites. Glycosylation is able to affect the function of the modified protein and so methods have been developed to predict the glycosylation sites based on the amino acid sequences. The statistical analysis of amino acids surrounding the glycosylation binding sites is dealt with by the being investigated by German Cancer Research Center. One of their tools GlySeq [77], statistically analyzes the amino acids surrounding the glycosylation sites based on protein sequences acquired from Swiss-Prot and Protein Data Bank (PDB), available on the GlySeqDB database. In addition to analyzing the surround sequence, a tool called GlyVicinity performs a statistical analysis of a PDB entry. At Johns Hopkins University, a model to mathematically formulate N-glycosylation was developed [78] whereby, given a set of expressed genes, the list of possible glycans synthesized by the input can be predicted. An enhanced

version of this model is currently in use. Recent developments in MS and ionization techniques have enabled the effective structural analysis of glycans [78,79]. The development of glyco-bioinformatics databases and tools such as UniCarb KB [80], GlycoMod [81], GlycoWorkBench [82], GlycReSoft [83] and Multiglycan [84] have accelerated the speed of glycan characterization [28].

Glycosciences DB [85] is a comprehensive data collection linking glycomics and proteomics data, which consists of the databases (<http://www.glycosciences.de/db-overview.php>) such as GlycoCD-glycoproteins/glycosphingolipids/carbohydrate recognition sites carbohydrate binding lectins/GlycoMapsDB-conformational maps of glycans and MonosaccharideDB-comprehensive resource of these monosaccharides and the tools like Sweet-II, GlyProt, PDB CARbohydrate REsidue check (pdb-care), pdb2linucs, Carbohydrate Ramachandran Plot (CARP), LINUCS, LiGraph, SUGar MOTif search (sumo), extraction of carbohydrate information from pdb-files, glycosidic linkage torsions, glycosylation sites statistics, oligosaccharides 2D structure drawings of which used for display the glycan structures and N- and O-glycan cores motif identification. In 2008, the work group "Frontiers in glycomics" emphasized the need for a curated, sustainably funded glyco-structure database [62,86]. This proposed structural database contains associated information about experimental and biosynthetic data accounting for various analytical methods used to assign glycan structures. Pioneering attempts for a comprehensive glycomics database resulted in Carbank [52] in the 1980s. Although this attempt came to a halt, its assembled data lived on in the next-generation databases including SWEET-DB [87] and GlycosuiteDB [88] (later incorporated into UniCarbKB [89] and GlyConnect [90]), having the same agenda as their ancestor. Then, integrative initiatives arose with the goal of centralizing scattered data [e.g., GlycomeDB [63]], and combining it with in silico analytical tools such as GLYCOSCIENCES.de [54], KEGG GLYCAN [55,91] and repositories provided by the Consortium for Functional Glycomics (CFG). The progress in the field was chaotic at the beginning of the century, but progressed considerably through the deployment of a new generation of centralized and integrative resources. These are GlyGen in the US, Glycomics@ExPASy [92] in Europe and GlyCosmos in Japan. GlyCosmos includes GlyTouCan [93], a registry providing glycan structures and develops complementary repositories. Each unique glycan can be correlated with MS data, HPLC retention times and NMR spectra using GlyTouCan.

UniCarb-DB was launched in 2011, to capture information from glycomics MS/MS data, [94,95]. Several versions of UniCarb-DB have been released with the primary aim focusing on improving the quality of glycomics data. UniCarb-DB is now integrated with Glycomics@ExPASy and provides the framework for accessing experimental MS data of biological origin. MS fragmentation glycan spectra are included in the NIST Glycan Mass Spectral Reference Library [96]. Compared to the progress made in glycan structure databases, slow progress has only been made towards the development of softwares for data analysis of glycomics data. GlycosidIQ automated the comparison of observed fragments with theoretical glycofragments from structural databases [97], this automation has been used in a commercial software [98]. GlycoReSoft was used for glycan detection from LC-MS data to compare different samples [99]. Other approaches were able to convert mass spectra into structures [95,100]. More advanced tools for glycomics analysis use partial de novo sequencing [101] including GlycoDeNovo [102] and the recently published Glycoforest [103]. High throughput glycomics MS annotation GRITS Toolbox [104] and quantitation tools [105] are now available.

Entire glycomes of cell systems have been characterized using advanced analytical tools and these data have been stored in glycosylation-specific databases like the GlycomeDB [106] and the Consortium for Functional Glycomics (CFG) website [56]. Bioinformatics and statistical approaches also offer an avenue to study glycosylation [107]. Systems-based mathematical modelling represents an alternate methodology [108,109]. A new MATLAB-based toolbox called Glycosylation Network Analysis Toolbox (GNAT) is able to construct and analyze glycosylation reaction networks. This enables construction, visualization and simulation of glycosylation reaction networks. It enables storing glycan structure information in SBML format files [110] and provides basic manipulation tools to tailor glycans and associated networks. It provides an interactive GUI (Graphical User Interface) that connects individual glycans represented in glycosylation networks and corresponding data in glycomics repositories.

Finally, the toolbox facilitates simulation of glycosylation reaction networks [111]. Table 1 presents the overview of the glyco-informatic tools and resources available to date.

Table 1. Bioinformatic tool applications in various Glycosylation associated studies.

Bioinformatics Resources	URL	Application	Reference
Glycosylation Structure Related Resources			
UniCarb KB	http://unicarbk.org/	Repository for glycan structures of glycoproteins	[64]
GlycoMod	http://web.expasy.org/glycomod/	Software tool for N-linked and O-linked glycan structures prediction	[84]
<i>GlycosuiteDB</i>	http://www.glycosuite.com	Relational database of glycoprotein glycan structures and their biological sources	[88]
<i>GlycoDeNovo</i>	https://www.cs.brandeis.edu/~hong/Research/GlycoDeNovo/GlycoDeNovo.htm	Algorithm for Accurate de novo Glycan Topology Reconstruction from Tandem Mass Spectras	[102]
<i>Glycoforest</i>	https://glycoforest.expasy.org/	Partial de-novo algorithm for sequencing glycan structures based on MS/MS spectra	[103]
Sweet-II	http://www.glycosciences.de/modeling/sweet2/doc/index.php	Tool to construct 3D models of saccharides from their sequences using standard nomenclature	[112]
GlyProt	http://www.glycosciences.de/modeling/glyprot/php/main.php	Tool to connect N-glycans in silico to a given 3D protein structure.	[113]
PDB CArbohydrate REsidue check (pdb-care)	http://www.glycosciences.de/tools/pdb-care/	Identify and assign carbohydrate structures using atom types and their 3D atom coordinates from PDB-files	[114]
Glyco3D	http://glyco3d.cermav.cnrs.fr/home.php	3D structures of monosaccharides, disaccharides, oligosaccharides, polysaccharides, glycosyltransferases, lectins, monoclonal antibodies against carbohydrates, and glycosaminoglycan-binding proteins	[115]
LiGraph	http://www.glycosciences.de/tools/LiGraph/	Convert a sugar graph to ASCII IUPAC sugar nomenclature or as a graph	[85]
KEGG Glycan database	https://www.genome.jp/kegg/glycan/	Database for glycan structures	[55]
<i>GlyTouCan</i>	https://glytoucan.org/	Repository for glycan structures.	[93]
Resources for Analytical Glycosylation			
<i>UniCarb-DB</i>	https://unicarb-db.expasy.org/	Repository for glycomic MS data	[89]
GlycanBuilder, GlycoWorkBench	http://www.eurocarbdb.org/applications/structure-tools	Tool to annotate Mass Spectra of Glycans	[116]

Table 1. Cont.

Bioinformatics Resources	URL	Application	Reference
Resources for Analytical Glycosylation			
Multiglycan	https://bio.tools/multiglycan	Information provider for glycan profile information from LC-MS Spectra.	[117]
GlycoFragment and Distance Mapping	www.dkfz.de/spec/projekte/fragments/	Web tool to interpret mass spectra of complex carbohydrates	[76]
Prediction of Glycosylation			
NGlycPred	https://bioinformatics.niaid.nih.gov/nglycpred/	Server to predict N-linked Glycosylation sites	[118]
UNIPEP	http://www.unipep.org/	Human and mouse N-glycosylated proteins and their N-glycosylation sites for biomarker discovery	[119]
GlycoEP	https://bio.tools/GlycoEP	prediction of N-linked, O-linked and C-linked glycosites in eukaryotic glycoproteins	[120]
GlySeq	http://www.glycosciences.de/tools/glyseq/	Analyze the sequences around glycosylation sites.	[77]
SPRINT-Gly	https://sparks-lab.org/server/sprint-gly/	predicting N- and O-linked glycosylation sites	[121]
DictyOGlyc	http://www.cbs.dtu.dk/services/DictyOGlyc/	Prediction of O-glycosylation sites in Dictyostelium discoideum proteins	[73]
GlycoMine ^{struct}	https://glycomine.erc.monash.edu/Lab/GlycoMine_Struct/index.jsp	Highly accurate mapping of the human N-linked and O-linked glycoproteomes	[65]
GlycoForm	http://www.boxer.tcd.ie/gf/	Mathematical model software for N-linked glycosylation	[78]
Glycopep	http://hexose.chem.ku.edu/predictiontable.php	N-linked glycosylation based on target protein and CID spectra analysis	[122]
Consortium for Functional Glycomics (CFC)	http://www.functionalglycomics.org/static/consortium/consortium.shtml	Glycomics resources of glycans and glycan-binding protein	[63]
NetNGlyc	http://www.cbs.dtu.dk/services/NetNGlyc/	N-Glycosylation site predictor	[123]
NetOGlyc	http://www.cbs.dtu.dk/services/NetOGlyc/	O-Glycosylation sites predictor	[75]
GPP	https://comp.chem.nottingham.ac.uk/glyco/	N- and O-Glycosylation site predictor	[73]
Big-PI	https://mendel.imp.ac.at/gpi/gpi_server.html	GPI anchors predictor	[124]
GPI-SOM	http://gpi.unibe.ch/	GPI anchors predictor	[125]
PredGPI	http://gpcr.biocomp.unibo.it/predgpi/	GPI anchors predictor	[85]
NetCGlyc	http://www.cbs.dtu.dk/services/NetCGlyc/	C-mannosylation sites prediction	[123]

Table 1. Cont.

Bioinformatics Resources	URL	Application	Reference
Carbohydrate Knowledgebases			
GlycoCT	www.eurocarbdb.org	Tool to convert unifying sequence format for carbohydrates	[61]
SWEET-DB	http://www.pdg.cnb.uam.es/cursos/Leon2002/pages/software/DatabasesListNAR2002/summary/300.html	Repository for annotated carbohydrates	[87]
SUGar MOTif search (sumo)	http://www.glycosciences.de/tools/sumo/	Tool to search sugar motif regions from carbohydrate structures	[126]
Bacterial Carbohydrate Structure Database (BCSDB, 6789)	http://csdb.glycoscience.ru/bacterial/	Repository for prokaryotic carbohydrate structures, taxonomy, bibliography, NMR data, etc.	[57]
pdb2linucs	http://www.glycosciences.de/tools/pdb2linucs/	Extract carbohydrate information from pdb-files and display it using the LINUCS-Code	[77]
CCSD	https://cordis.europa.eu/project/id/BIOT0184	Information system of carbohydrate science	[52]
Carbohydrate Ramachandran Plot (CARP)	http://www.glycosciences.de/tools/carp/	Analyze carbohydrate data from PDB files using the pdb2linucs algorithm	[127]
LINUCS	http://www.glycosciences.de/tools/linucs/	Linear Notation for Unique description of Carbohydrate Sequences	[128]
Miscellaneous Resources for Glycosylation			
RESID	https://proteininformationresource.org/resid/togm.shtml	Table of Glycosylation Modifications	[129]
GlyGen	https://www.glygen.org/	Informatics Resources for Glycoscience	[130]
LfDB	https://acgg.asia/lfdb2/	Lectin Frontier DataBase	[131]
CFG	http://www.functionalglycomics.org/	Comprehensive resource for functional glycomics research	[56]
GlycoStore	https://www.glycostore.org	A curated database of information on glycan retention properties with chromatographic, electrophoretic and mass-spectrometry composition data.	[132]
Galaxy, LipidBank	https://jcgdb.jp/database_en.html	Consortium for Glycobiology and Glycobiotechnology database	[133]
Glycosylation Network Analysis Toolbox (GNAT)	http://gnatmatlab.sourceforge.net/	MATLAB-based environment for systems glycobiology	[111]
Glycomics@ExPASy	https://www.expasy.org/search/glycomics	Expasy resources for glycomics data.	[134]
MonosaccharideDB	http://www.monosaccharidedb.org/	comprehensive resource of monosaccharides	[135]

Table 1. Cont.

Bioinformatics Resources	URL	Application	Reference
Miscellaneous Resources for Glycosylation			
GlycReSoft	http://www.bumc.bu.edu/msr/glycresoft/	Software for Glycomics and Glycoproteomics	[99]
CarbBank	https://www.genome.jp/dbget-bin/www_bfind?carbbank	Database management program and the project system of CCSD	[53]
GlyCosmos	https://glycosmos.org/	Comprehensive web resource for the glycosciences	[136]
SysPTM	http://lifecenter.sgst.cn/SysPTM/	Resource for post-translational modification	[137]

3. Scarce List of Bioinformatic Resources for Glycosylation in Cancers

It is now known that a plethora of alterations in protein glycosylation can profoundly impact disease and cell phenotypes. The biosynthesis of cancer-associated glycans and the subsequent glycoproteome act synergistically towards disease evolution. It is now proven that there exists a cross talk between glycans and tumor cells [138]. This cross talk leads to the activation of specific oncogenic pathways and brings about invasion and disease dissemination, becoming an important source of relevant glyco(neo)epitopes, holding tremendous potential towards clinical intervention. Glycosylation modifications in cancer include, alterations in glycan length, often toward shorter O-glycans and more branched N-glycans, often marked by distinct changes in glycan sialylation and fucosylation influencing terminal epitopes. Aberrant glycosylation has become the fundamental basis for glycan based cancer biomarker discovery, which has been extensively reviewed and reported by many authors [16,139–165]. Apart from these, many changes in glycan chains are reported for glycosaminoglycans (GAG). Aberrant glycosylation thus plays a crucial role in tumor progression by regulating tumor proliferation, invasion, metastasis, and angiogenesis [166,167], being frequently cited as a hallmark of cancer [117]. In ovarian cancer, increased branching of the N-glycans attached to glycoproteins [168] and increased sialylation of N-glycopeptides [169] have been confirmed. Because 90% of ovarian cancers are of epithelial origin [170], the glycosylation landscape on cancer cell surface membrane glycoproteins has become diagnostically, prognostically, and therapeutically meaningful [171].

With the increasing evidence of the expression of aberrant glycans being intimately linked with cancer [172], other studies have also shown that altered glycosylation patterns on the surface of cancer cells disrupt normal cellular functions. Lauc et al., 2013 [173] have provided experimental evidence that glycosylation of immunoglobulin G (IgG) influences IgG effector function by modulating binding to Fc receptors. Genetic loci associated with IgG glycosylation were identified using ultra-performance liquid chromatography (UPLC) and MALDI-TOF mass spectrometry (MS). Meta-analysis of genome-wide association study (GWAS) identified 9 genome-wide significant loci strongly associated with various pathological conditions including haematological cancers such as acute lymphoblastic leukaemia, Hodgkin lymphoma, and multiple myeloma. Their study confirmed that it was possible to identify new loci that control glycosylation of a single plasma protein using GWAS in autoimmune diseases and haematological cancers.

Glycosylation is reported to be a global target for androgen control in prostate cancer patients. RNA sequencing analysis of prostate cancer cells and patients by Munkley, 2017 [174] identified 700 androgen-regulated genes. Gene ontology (GO) defined glycosylation as an androgen-regulated process in prostate cancer cells and identified 25 genes with roles in the glycosylation process.

Walsh et al., [48] have described a pipeline and Saldova et al. [175] report the successful identification of biomarkers in breast cancer and ovarian cancer [175] and still others in lung cancer [176]

and arthritis [177] using bioinformatics resources. They report an automated LC platform that includes bioinformatics tools such as GlycoBase, GlycoDigest. GlycoMarker, helps identify biomarkers in LC profiles. It is able to statistically test markers, visualize and model algorithms. GlycoMarker was effectively tested on a dataset of 62 breast cancer and 107 control LC profiles [175] and biomarkers for breast cancer were identified. Glycomarker is no longer active.

Breast cancer is the most common cancer in women worldwide, and resistance to the current therapeutics, is an increasing medical challenge. Recently, there is strong evidence pointing out towards the role of glycosylation in tumor formation and metastasis. Kaur et al. [178] have studied the role of sialoglycoproteins in breast cancer by using bioinformatics tools, based on UNIPROT database. By collection of FASTA sequences of breast cancer gene sequence, a database in the form of a word document was constructed. Glycosylation was studied using YinOYang tool on EXPASY, followed by inclusion of differentially expressed genes using KEGG, DAVID and Ingenuity databases. Significant changes in the expression profiling of glycosylation patterns of various proteins associated with breast cancer was also identified. Bioinformatics (involving yinOyang tool on expasy, followed by involvement of differentially expressed genes in important molecular and signaling casades using KEGG, DAVID and Ingenuity databases) worked hand in hand to prove that differential aberrant glycosylated proteins in breast cancer cells are a crucial factor for the overall progression and development of cancer. These were the limited reports that were available involving the use of informatics tools for glycosylation in cancers as summarized in Table 2. As observed from the handful of information in Table 2, this review stress on the fact that more work needs to be done in this area.

Table 2. Bioinformatic tools for glycosylation in cancer applications.

Bioinformatics Resources	Cancer Type	URL	Application	Reference
RNA sequencing analysis of genes	prostate cancer cell lines and patients	A composite sequencing server	RNA sequencing analysis of identified a set of 700 androgen-regulated genes.	[174]
Gene ontology (GO)	prostate cancer cell lines and patients	Gene ontology (GO) is general functional annotation server	identified 72 terms with significant gene enrichment ($p < 0.05$) and defined glycosylation as an androgen-regulated process in prostate cancer cells.	[174]
GlycoBase (inactive now)	Breast cancer cells	(https://glycobase.nibr.tie):	database of experimentally determined glycan structures originally developed from the EurocarbDB project	[58]
GlycoDigest	Breast cancer cells	(http://www.glycodigest.org):	a tool that simulates exoglycosidase digestion based on controlled rules acquired from expert knowledge and experimental evidence available in GlycoBase	[179]
GlycoMarker	Breast cancer cells	https://glycobase.nibr.tie/glycomarker	Web application/server for Biomarker discovery, identifies markers in LC profiles	[48]

Table 2. Cont.

Bioinformatics Resources	Cancer Type	URL	Application	Reference
using KEGG, DAVID and Ingenuity databases, uniprot database	Breast cancer cells	KEGG, DAVID and Ingenuity databases, uniprot database these are very general functional analysis and sequence databases	Significant change in the expression profiling of glycosylation patterns of various proteins associated with Triple negative breast cancer was identified. Differential aberrant glycosylated proteins in breast cancer cells with respect to non-neoplastic cells	[178]
genome-wide association study (GWAS)	haematological cancers such as acute lymphoblastic leukaemia, Hodgkin lymphoma, and multiple myeloma	GWAS is a general study for genome wide citation	identify new loci that control glycosylation of a single plasma protein using GWAS	[178]
YinOYang	Breast cancer cells	http://www.cbs.dtu.dk/services/YinOYang/	O-Glycosylation sites predictor	[125]
ANCOVA/MANCOVA statistics	Lung Cancer cells	https://www.statisticssolutions.com/multivariate-analysis-of-covariance-mancova/	Statistical analysis	[176]
Student's <i>t</i> test, orthogonal partial least squares discriminant analysis and receiver operating characteristic curve	colorectal cancer tissues (CRC) in Chinese patients	Statistical tools	Statistical analysis of MS data from Linear ion trap quadrupole-electrospray ionization mass spectrometry, on CRC tissues	[180]
Perseus™ (Max Planck Institute of Biochemistry, Berlin, Germany)	Breast cancer	https://pubmed.ncbi.nlm.nih.gov/29344888/	Perseus™ 1.5.2.6 was used for hierarchical clustering, principal component analysis (PCA), and plotting for visualization and statistics	[181]

4. Future Perspective

The transversal nature of glycans and the existence of crosstalk between glycans and stromal components of tumor microenvironments is well established. No the focus is on including glycans and glycoconjugates in comprehensive panomic models, for bringing about progress in precision medicine. Precision medicine is the new age medical approach whereby personalized individual care and medical counsel is given. Glycotope GmbH, founded in 2001, developed GlycoExpress technology [182], which can produce and optimize humanized glycoproteins [183]. They have developed novel technologies for production of biopharmaceuticals based on fully humanized glycoproteins. [182,183]. Thus, the importance and impact of glycosylation on cancer and its diagnosis and therapy are clearly expanding. Glycosylation thus holds promises towards cancer diagnosis, early detection as well as cancer therapy. Bioinformatics is a core component of any pipeline and greatly speed up the storage, annotation and analysis of the data. Bioinformatics programs enable even inexperienced researchers to handle large data sets and performing analysis with ease and efficacy. Bioinformatic applications have

been pivotal in enlarging the visions of proteomics in clinical disease detection and management and they would and should in Glycosylation applications too, From what Aoki-Kinoshita [28] reported as three major databases for complex carbohydrates [Glycosciences.de, KEGG GLYCAN, database by the Consortium for Functional Glycomics (CFG)], we have come a long way. This is evidenced by the exhaustive list of tools presented in Table 1. Through reviewing the current scenario in this area, we discovered that many tools and glyco-informatic resources have become inactive, redundant and void. Some tools have merely been developed and demonstrated by a research group and nothing much has been done on it and it has gone inactive. We speculate the reason for this is the gap between tool development and knowledge of existing tools and practical implementation aspects.

As listed in Section 2, numerous bioinformatics resources are available for glycosylation studies, however, this review found a huge vacuum with respect to reports on the use of these glyco-informatic tools for glycosylation studies with respect to cancer. Aberrant glycosylation actively contributes to tumor progression by regulating tumor proliferation, invasion, metastasis, and angiogenesis and could easily benefit from the use of glyco-informatic tools. With a wide range of bioinformatics tools specifically for glycosylation developed and proven for their efficacy, the fact that not even 1% of these available options have been used for glycosylation in cancers is an alarming inference this review aims to focus. This review emphasizes the need to apply the available glyco-informatics options towards cancer for a better cutting-edge technology transfer. The sad fact in science is when we hold the sophistication and yet refrain from applying it towards practical implementation, especially when the practical implementation pertains to human health and welfare.

When cancer has become a rising concern and constant remedies and milestones are being continuously evaluated, given the fact that glycosylation could shed light on early detection and treatment regimens, bioinformatics tools will no doubt be an added asset. Bioinformatics has eased various tedious protocols and will prove handy in this aspect too. Only when the existing tools are utilized and exploited to the full extent, can any improvisations or improvements to suit the requirements for cancer applications be made. The available options can by themselves bring about benchmark progress in glycosylation in cancer detection and therapy. This review puts forth the need for testing the bioinformatics glyco-informatic resources available for cancer research, which may eventually be the key to a crucial fundamental understanding of this complicated disease [184].

5. Conclusions

The bioinformatics options available for glycosylation which is a variant of the post-translation modification of proteins have been reviewed and a consolidated list of the tools and resources available have been presented. The lack of application of the existing glyco-informatic tools for cancer research has been speculated and the need to do so has been emphasized. The major emphasis this review is highlighting is that given all the available bioinformatics resources and developments in glyco-informatics, there is very less that has been applied into the field of cancer biology. This review exposes the huge gap between available bioinformatics tools and technologies and real time applications. This review is expected to bring about an awakening amidst the glyco-informaticians as well as cancer biologists leading to significant bridging of this gap.

Author Contributions: Conceptualization, writing—original draft preparation, M.M., J.G., J.-W.O. and V.A.; writing—review and editing, S.C., A.J.P.D. and S.W.H.; funding acquisition, J.-W.O. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through the Agri-Bio Industry Technology Development Program and through the Agri-food R&D Performance Follow-up Support Program, funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA, 116084-03-3-HD020 and 120026011HD020).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ramamurthy, B.; Hook, P.; Larsson, L. An overview of carbohydrate-protein interactions with specific reference to myosin and ageing. *Acta Physiol. Scand.* **1999**, *167*, 327–329. [[CrossRef](#)]
2. Mazola, Y.; Chinea, G.; Musacchio, A. Glycosylation and Bioinformatics: Current status for glycosylation prediction tools. *Biotechnol. Appl.* **2011**, *28*, 6–12.
3. Mitra, N.; Sinha, S.; Ramya, T.N.; Surolia, A. N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem. Sci.* **2006**, *31*, 156–163. [[CrossRef](#)] [[PubMed](#)]
4. Mbonye, U.R.; Yuan, C.; Harris, C.E.; Sidhu, R.S.; Song, I.; Arakawa, T.; Smith, W.L. Two distinct pathways for cyclooxygenase-2 protein degradation. *J. Biol. Chem.* **2008**, *283*, 8611–8623. [[CrossRef](#)]
5. Specks, U.; Fass, D.N.; Finkelman, J.D.; Hummel, A.M.; Viss, M.A.; Litwiller, R.D.; McDonald, C.J. Functional significance of Asn-linked glycosylation of proteinase 3 for enzymatic activity, processing, targeting, and recognition by anti-neutrophil cytoplasmic antibodies. *J. Biochem.* **2007**, *141*, 101–112. [[CrossRef](#)] [[PubMed](#)]
6. Sola, R.J.; Griebenow, K. Effects of glycosylation on the stability of protein pharmaceuticals. *J. Pharm. Sci.* **2009**, *98*, 1223–1245. [[CrossRef](#)] [[PubMed](#)]
7. Troy, F.A., 2nd. Polysialylation: From bacteria to brains. *Glycobiology* **1992**, *2*, 5–23. [[CrossRef](#)]
8. Vagin, O.; Kraut, J.A.; Sachs, G. Role of N-glycosylation in trafficking of apical membrane proteins in epithelia. *Am. J. Physiol. Renal Physiol.* **2009**, *296*, F459–F469. [[CrossRef](#)]
9. Couldrey, C.; Green, J.E. Metastases: The glycan connection. *Breast Cancer Res.* **2000**, *2*, 321–323. [[CrossRef](#)]
10. Corthay, A.; Backlund, J.; Holmdahl, R. Role of glycopeptide-specific T cells in collagen-induced arthritis: An example how post-translational modification of proteins may be involved in autoimmune disease. *Ann. Med.* **2001**, *33*, 456–465. [[CrossRef](#)]
11. Freeze, H.H. Update and perspectives on congenital disorders of glycosylation. *Glycobiology* **2001**, *11*, 129R–143R. [[CrossRef](#)] [[PubMed](#)]
12. Varki, A.; Cummings, R.D.; Esko, J.D.; Freeze, H.H.; Hart, G.W.; Etzler, M.E. *Essentials of Glycobiology*, 2nd ed.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2008.
13. Zaia, J. Mass spectrometry and the emerging field of glycomics. *Chem. Biol.* **2008**, *15*, 881–892. [[CrossRef](#)] [[PubMed](#)]
14. Chen, Y.Z.; Tang, Y.R.; Sheng, Z.Y.; Zhang, Z. Prediction of mucin-type O-glycosylation sites in mammalian proteins using the composition of k-spaced amino acid pairs. *BMC Bioinform.* **2008**, *9*, 101. [[CrossRef](#)] [[PubMed](#)]
15. Kianmehr, A.; Mohammadi, H.S.; Shokrgozar, M.A.; Omidinia, E. In silico design and analysis of a new hyperglycosylated analog of erythropoietin to improve drug efficacy. *Adv. Biomed. Res.* **2015**, *4*, 142. [[CrossRef](#)]
16. Munkley, J.; Elliott, D.J. Hallmarks of glycosylation in cancer. *Oncotarget* **2016**, *7*, 35478–35489. [[CrossRef](#)]
17. Azevedo, R.; Peixoto, A.; Gaiteiro, C.; Fernandes, E.; Neves, M.; Lima, L.; Santos, L.L.; Ferreira, J.A. Over forty years of bladder cancer glycobiology: Where do glycans stand facing precision oncology? *Oncotarget* **2017**, *8*, 91734–91764. [[CrossRef](#)]
18. 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](#)]
19. Palmigiano, A.; Barone, R.; Sturiale, L.; Sanfilippo, C.; Bua, R.O.; Romeo, D.A.; Messina, A.; Capuana, M.L.; Maci, T.; Le Pira, F.; et al. CSF N-glycoproteomics for early diagnosis in Alzheimer's disease. *J. Proteom.* **2016**, *131*, 29–37. [[CrossRef](#)]
20. Lauc, G.; Pezer, M.; Rudan, I.; Campbell, H. Mechanisms of disease: The human N-glycome. *Biochim. Biophys. Acta* **2016**, *1860*, 1574–1582. [[CrossRef](#)]
21. Brockhausen, I.; Yang, J.M.; Burchell, J.; Whitehouse, C.; Taylor-Papadimitriou, J. Mechanisms underlying aberrant glycosylation of MUC1 mucin in breast cancer cells. *Eur. J. Biochem.* **1995**, *233*, 607–617. [[CrossRef](#)]
22. Gils, A.; Pedersen, K.E.; Skottrup, P.; Christensen, A.; Naessens, D.; Deinum, J.; Enghild, J.J.; Declerck, P.J.; Andreasen, P.A. Biochemical importance of glycosylation of plasminogen activator inhibitor-1. *Thromb. Haemost.* **2003**, *90*, 206–217. [[CrossRef](#)] [[PubMed](#)]

23. Lempiainen, A.; Hotakainen, K.; Blomqvist, C.; Alfthan, H.; Stenman, U.H. Hyperglycosylated human chorionic gonadotropin in serum of testicular cancer patients. *Clin. Chem.* **2012**, *58*, 1123–1129. [[CrossRef](#)] [[PubMed](#)]
24. Sato, Y.; Nakata, K.; Kato, Y.; Shima, M.; Ishii, N.; Koji, T.; Taketa, K.; Endo, Y.; Nagataki, S. Early recognition of hepatocellular carcinoma based on altered profiles of alpha-fetoprotein. *N. Engl. J. Med.* **1993**, *328*, 1802–1806. [[CrossRef](#)] [[PubMed](#)]
25. Gadroy, P.; Stridsberg, M.; Capon, C.; Michalski, J.C.; Strub, J.M.; Van Dorsselaer, A.; Aunis, D.; Metz-Boutigue, M.H. Phosphorylation and O-glycosylation sites of human chromogranin A (CGA79-439) from urine of patients with carcinoid tumors. *J. Biol. Chem.* **1998**, *273*, 34087–34097. [[CrossRef](#)]
26. Jankovic, M.M.; Milutinovic, B.S. Glycoforms of CA125 antigen as a possible cancer marker. *Cancer Biomark.* **2008**, *4*, 35–42. [[CrossRef](#)]
27. Hua, L.; Liu, Y.; Zhen, S.; Wan, D.; Cao, J.; Gao, X. Expression and biochemical characterization of recombinant human epididymis protein 4. *Protein Expr. Purif.* **2014**, *102*, 52–62. [[CrossRef](#)]
28. Aoki-Kinoshita, K.F. An introduction to bioinformatics for glycomics research. *PLoS Comput. Biol.* **2008**, *4*, e1000075. [[CrossRef](#)]
29. von der Lieth, C.W.; Bohne-Lang, A.; Lohmann, K.K.; Frank, M. Bioinformatics for glycomics: Status, methods, requirements and perspectives. *Brief. Bioinform.* **2004**, *5*, 164–178. [[CrossRef](#)]
30. Frank, M.; Schloissnig, S. Bioinformatics and molecular modeling in glycobiology. *Cell Mol. Life Sci.* **2010**, *67*, 2749–2772. [[CrossRef](#)]
31. Yu, U.; Lee, S.H.; Kim, Y.J.; Kim, S. Bioinformatics in the post-genome era. *J. Biochem. Mol. Biol.* **2004**, *37*, 75–82. [[CrossRef](#)]
32. Von der Lieth, C.W.; Lutteke, T.; Frank, M. *Bioinformatics for Glycobiology and Glycomics: An Introduction*, 1st ed.; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2009.
33. Abrahams, J.L.; Taherzadeh, G.; Jarvas, G.; Guttman, A.; Zhou, Y.; Campbell, M.P. Recent advances in glycoinformatic platforms for glycomics and glycoproteomics. *Curr. Opin. Struct. Biol.* **2020**, *62*, 56–69. [[CrossRef](#)] [[PubMed](#)]
34. Li, X.; Xu, Z.; Hong, X.; Zhang, Y.; Zou, X. Databases and Bioinformatic Tools for Glycobiology and Glycoproteomics. *Int. J. Mol. Sci.* **2020**, *21*, 6727. [[CrossRef](#)] [[PubMed](#)]
35. Abdurakhmonov, I.Y. *Bioinformatics: Basics, Development, and Future, Bioinformatics—Updated Features and Applications*; Abdurakhmonov, I.Y., Ed.; IntechOpen: Rijeka, Croatia, 2016.
36. Zhang, S.-Y.; Liu, S.-L. *Bioinformatics*, 2nd ed.; Academic Press: New York, NY, USA, 2013; pp. 338–340. [[CrossRef](#)]
37. Kersey, P.; Apweiler, R. Linking publication, gene and protein data. *Nat. Cell Biol.* **2006**, *8*, 1183–1189. [[CrossRef](#)] [[PubMed](#)]
38. Mulder, N.J.; Kersey, P.; Pruess, M.; Apweiler, R. In silico characterization of proteins: UniProt, InterPro and Integr8. *Mol. Biotechnol.* **2008**, *38*, 165–177. [[CrossRef](#)]
39. Lutteke, T. Web resources for the glycoscientist. *Chembiochem* **2008**, *9*, 2155–2160. [[CrossRef](#)]
40. Mahal, L.K. Glycomics: Towards bioinformatic approaches to understanding glycosylation. *Anti Cancer Agents Med. Chem.* **2008**, *8*, 37–51. [[CrossRef](#)]
41. Mamitsuka, H. Informatic innovations in glycobiology: Relevance to drug discovery. *Drug Discov. Today* **2008**, *13*, 118–123. [[CrossRef](#)]
42. Ranzinger, R.; Herget, S.; Lutteke, T.; Frank, M. *Handbook of Glycomics*; Cummings, R.D., Pierce, J.M., Eds.; Elsevier: Amsterdam, The Netherlands, 2009.
43. von der Lieth, C.W. *Comprehensive Glycoscience—From Chemistry to Systems Biology*; Elsevier: Oxford, UK, 2007; Volume 2.
44. von der Lieth, C.W.; Lutteke, T.; Frank, M. The role of informatics in glycobiology research with special emphasis on automatic interpretation of MS spectra. *Biochim. Biophys. Acta* **2006**, *1760*, 568–577. [[CrossRef](#)]
45. Aoki-Kinoshita, K.F.; Kanehisa, M. Bioinformatics approaches in glycomics and drug discovery. *Curr. Opin. Mol. Ther.* **2006**, *8*, 514–520.
46. Perez, S.; Mulloy, B. Prospects for glycoinformatics. *Curr. Opin. Struct. Biol.* **2005**, *15*, 517–524. [[CrossRef](#)]
47. Marchal, I.; Golfier, G.; Dugas, O.; Majed, M. Bioinformatics in glycobiology. *Biochimie* **2003**, *85*, 75–81. [[CrossRef](#)]

48. Walsh, I.; O’Flaherty, R.; Rudd, P.M. Bioinformatics applications to aid high-throughput glycan profiling. *Perspect. Sci.* **2017**, *11*, 31–39. [[CrossRef](#)]
49. Brooksbank, C.; Camon, E.; Harris, M.A.; Magrane, M.; Martin, M.J.; Mulder, N.; O’Donovan, C.; Parkinson, H.; Tuli, M.A.; Apweiler, R.; et al. The European Bioinformatics Institute’s data resources. *Nucleic Acids Res.* **2003**, *31*, 43–50. [[CrossRef](#)] [[PubMed](#)]
50. Wheeler, D.L.; Barrett, T.; Benson, D.A.; Bryant, S.H.; Canese, K.; Chetvernin, V.; Church, D.M.; Dicuccio, M.; Edgar, R.; Federhen, S.; et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **2008**, *36*, D13–D21. [[CrossRef](#)] [[PubMed](#)]
51. Whitfield, E.J.; Pruess, M.; Apweiler, R. Bioinformatics database infrastructure for biotechnology research. *J. Biotechnol.* **2006**, *124*, 629–639. [[CrossRef](#)] [[PubMed](#)]
52. Doubet, S.; Bock, K.; Smith, D.; Darvill, A.; Albersheim, P. The Complex Carbohydrate Structure Database. *Trends Biochem. Sci.* **1989**, *14*, 475–477. [[CrossRef](#)]
53. Doubet, S.; Albersheim, P. CarbBank. *Glycobiology* **1992**, *2*, 505. [[CrossRef](#)]
54. Lutteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; von der Lieth, C.W. GLYCOSCIENCES.de: An Internet portal to support glycomics and glycobiology research. *Glycobiology* **2006**, *16*, 71R–81R. [[CrossRef](#)]
55. Hashimoto, K.; Goto, S.; Kawano, S.; Aoki-Kinoshita, K.F.; Ueda, N.; Hamajima, M.; Kawasaki, T.; Kanehisa, M. KEGG as a glycome informatics resource. *Glycobiology* **2006**, *16*, 63R–70R. [[CrossRef](#)]
56. Raman, R.; Venkataraman, M.; Ramakrishnan, S.; Lang, W.; Raguram, S.; Sasisekharan, R. Advancing glycomics: Implementation strategies at the consortium for functional glycomics. *Glycobiology* **2006**, *16*, 82R–90R. [[CrossRef](#)]
57. Toukach, F.V.; Knirel, Y.A. New database of bacterial carbohydrate structures. *Glycoconj. J.* **2005**, *22*, 216–217.
58. Campbell, M.P.; Royle, L.; Radcliffe, C.M.; Dwek, R.A.; Rudd, P.M. GlycoBase and autoGU: Tools for HPLC-based glycan analysis. *Bioinformatics* **2008**, *24*, 1214–1216. [[CrossRef](#)] [[PubMed](#)]
59. Cooper, C.A.; Joshi, H.J.; Harrison, M.J.; Wilkins, M.R.; Packer, N.H. GlycoSuiteDB: A curated relational database of glycoprotein glycan structures and their biological sources. 2003 update. *Nucleic Acids Res.* **2003**, *31*, 511–513. [[CrossRef](#)] [[PubMed](#)]
60. Toukach, P.; Joshi, H.J.; Ranzinger, R.; Knirel, Y.; von der Lieth, C.W. Sharing of worldwide distributed carbohydrate-related digital resources: Online connection of the Bacterial Carbohydrate Structure DataBase and GLYCOSCIENCES.de. *Nucleic Acids Res.* **2007**, *35*, D280–D286. [[CrossRef](#)]
61. Herget, S.; Ranzinger, R.; Maass, K.; Lieth, C.W. GlycoCT—a unifying sequence format for carbohydrates. *Carbohydr. Res.* **2008**, *343*, 2162–2171. [[CrossRef](#)]
62. Packer, N.H.; von der Lieth, C.W.; Aoki-Kinoshita, K.F.; Lebrilla, C.B.; Paulson, J.C.; Raman, R.; Rudd, P.; Sasisekharan, R.; Taniguchi, N.; York, W.S. Frontiers in glycomics: Bioinformatics and biomarkers in disease. An NIH white paper prepared from discussions by the focus groups at a workshop on the NIH campus, Bethesda MD (September 11–13, 2006). *Proteomics* **2008**, *8*, 8–20. [[CrossRef](#)]
63. Ranzinger, R.; Herget, S.; Wetter, T.; von der Lieth, C.W. GlycomeDB—Integration of open-access carbohydrate structure databases. *BMC Bioinform.* **2008**, *9*, 384. [[CrossRef](#)]
64. Ranzinger, R.; Frank, M.; von der Lieth, C.W.; Herget, S. Glycome-DB.org: A portal for querying across the digital world of carbohydrate sequences. *Glycobiology* **2009**, *19*, 1563–1567. [[CrossRef](#)]
65. Li, F.; Li, C.; Revote, J.; Zhang, Y.; Webb, G.I.; Li, J.; Song, J.; Lithgow, T. GlycoMine(struct): A new bioinformatics tool for highly accurate mapping of the human N-linked and O-linked glycoproteomes by incorporating structural features. *Sci. Rep.* **2016**, *6*, 34595. [[CrossRef](#)]
66. Gupta, R.; Jung, E.; Brunak, S. Prediction of N-glycosylation sites in human proteins. 2004. in preparation.
67. Steentoft, C.; Vakhrushev, S.Y.; Joshi, H.J.; Kong, Y.; Vester-Christensen, M.B.; Schjoldager, K.T.; Lavrsen, K.; Dabelsteen, S.; Pedersen, N.B.; Marcos-Silva, L.; et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* **2013**, *32*, 1478–1488. [[CrossRef](#)]
68. Gupta, R.; Brunak, S. Prediction of Glycosylation across the Human Proteome and the Correlation to Protein Function. In Proceedings of the Pacific Symposium on Biocomputing, Lihue, Hawaii, 3–7 January 2002; pp. 310–322.
69. Gupta, R.; Birch, H.; Rapacki, K.; Brunak, S.; Hansen, J.E. O-GLYCBASE version 4.0: A revised database of O-glycosylated proteins. *Nucleic Acids Res.* **1999**, *27*, 370–372. [[CrossRef](#)]

70. Julenius, K. NetCGlyc 1.0: Prediction of mammalian C-mannosylation sites. *Glycobiology* **2007**, *17*, 868–876. [[CrossRef](#)]
71. Li, F.; Li, C.; Wang, M.; Webb, G.I.; Zhang, Y.; Whisstock, J.C.; Song, J. GlycoMine: A machine learning-based approach for predicting N-, C- and O-linked glycosylation in the human proteome. *Bioinformatics* **2015**, *31*, 1411–1419. [[CrossRef](#)] [[PubMed](#)]
72. Hamby, S.E.; Hirst, J.D. Prediction of glycosylation sites using random forests. *BMC Bioinform.* **2008**, *9*, 500. [[CrossRef](#)] [[PubMed](#)]
73. Fankhauser, N.; Maser, P. Identification of GPI anchor attachment signals by a Kohonen self-organizing map. *Bioinformatics* **2005**, *21*, 1846–1852. [[CrossRef](#)] [[PubMed](#)]
74. Pierleoni, A.; Martelli, P.L.; Casadio, R. PredGPI: A GPI-anchor predictor. *BMC Bioinform.* **2008**, *9*, 392. [[CrossRef](#)] [[PubMed](#)]
75. Eisenhaber, B.; Bork, P.; Yuan, Y.; Löffler, G.; Eisenhaber, F. Automated annotation of GPI anchor sites: Case study *C. elegans*. *Trends Biochem. Sci.* **2000**, *25*, 340–341. [[CrossRef](#)]
76. Lohmann, K.K.; von der Lieth, C.W. GlycoFragment and GlycoSearchMS: Web tools to support the interpretation of mass spectra of complex carbohydrates. *Nucleic Acids Res.* **2004**, *32*, W261–W266. [[CrossRef](#)]
77. Lutteke, T.; Frank, M.; von der Lieth, C.W. Carbohydrate Structure Suite (CSS): Analysis of carbohydrate 3D structures derived from the PDB. *Nucleic Acids Res.* **2005**, *33*, D242–D246. [[CrossRef](#)]
78. Krambeck, F.J.; Betenbaugh, M.J. A mathematical model of N-linked glycosylation. *Biotechnol. Bioeng.* **2005**, *92*, 711–728. [[CrossRef](#)]
79. Umana, P.; Bailey, J.E. A mathematical model of N-linked glycoform biosynthesis. *Biotechnol. Bioeng.* **1997**, *55*, 890–908. [[CrossRef](#)]
80. Lohmann, K.K.; von der Lieth, C.W. GLYCO-FRAGMENT: A web tool to support the interpretation of mass spectra of complex carbohydrates. *Proteomics* **2003**, *3*, 2028–2035. [[CrossRef](#)] [[PubMed](#)]
81. Goldberg, D.; Sutton-Smith, M.; Paulson, J.; Dell, A. Automatic annotation of matrix-assisted laser desorption/ionization N-glycan spectra. *Proteomics* **2005**, *5*, 865–875. [[CrossRef](#)] [[PubMed](#)]
82. Tang, H.; Mechref, Y.; Novotny, M.V. Automated interpretation of MS/MS spectra of oligosaccharides. *Bioinformatics* **2005**, *21* (Suppl. S1), i431–i439. [[CrossRef](#)]
83. Irungu, J.; Go, E.P.; Dalpathado, D.S.; Desaire, H. Simplification of mass spectral analysis of acidic glycopeptides using GlycoPep ID. *Anal. Chem.* **2007**, *79*, 3065–3074. [[CrossRef](#)]
84. Cooper, C.A.; Gasteiger, E.; Packer, N.H. GlycoMod—A software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics* **2001**, *1*, 340–349. [[CrossRef](#)]
85. Bohm, M.; Bohne-Lang, A.; Frank, M.; Loss, A.; Rojas-Macias, M.A.; Lutteke, T. Glycosciences.DB: An annotated data collection linking glycomics and proteomics data (2018 update). *Nucleic Acids Res.* **2019**, *47*, D1195–D1201. [[CrossRef](#)]
86. Rojas-Macias, M.A.; Mariethoz, J.; Andersson, P.; Jin, C.; Venkatakrishnan, V.; Aoki, N.P.; Shinmachi, D.; Ashwood, C.; Madunic, K.; Zhang, T.; et al. Towards a standardized bioinformatics infrastructure for N- and O-glycomics. *Nat. Commun.* **2019**, *10*, 3275. [[CrossRef](#)]
87. Loss, A.; Bunsmann, P.; Bohne, A.; Loss, A.; Schwarzer, E.; Lang, E.; von der Lieth, C.W. SWEET-DB: An attempt to create annotated data collections for carbohydrates. *Nucleic Acids Res.* **2002**, *30*, 405–408. [[CrossRef](#)]
88. Cooper, C.A.; Harrison, M.J.; Wilkins, M.R.; Packer, N.H. GlycoSuiteDB: A new curated relational database of glycoprotein glycan structures and their biological sources. *Nucleic Acids Res.* **2001**, *29*, 332–335. [[CrossRef](#)]
89. Campbell, M.P.; Peterson, R.; Mariethoz, J.; Gasteiger, E.; Akune, Y.; Aoki-Kinoshita, K.F.; Lisacek, F.; Packer, N.H. UniCarbKB: Building a knowledge platform for glycoproteomics. *Nucleic Acids Res.* **2014**, *42*, D215–D221. [[CrossRef](#)] [[PubMed](#)]
90. Alocchi, D.; Mariethoz, J.; Gastaldello, A.; Gasteiger, E.; Karlsson, N.G.; Kolarich, D.; Packer, N.H.; Lisacek, F. GlyConnect: Glycoproteomics Goes Visual, Interactive, and Analytical. *J. Proteome Res.* **2019**, *18*, 664–677. [[CrossRef](#)] [[PubMed](#)]
91. Kanehisa, M. *KEGG GLYCAN*; Aoki-Kinoshita, K.F., Ed.; Springer: Berlin/Heidelberg, Germany, 2017.
92. Mariethoz, J.; Alocchi, D.; Gastaldello, A.; Horlacher, O.; Gasteiger, E.; Rojas-Macias, M.; Karlsson, N.G.; Packer, N.H.; Lisacek, F. Glycomics@ExPASy: Bridging the Gap. *Mol. Cell. Proteom.* **2018**, *17*, 2164–2176. [[CrossRef](#)] [[PubMed](#)]

93. Tiemeyer, M.; Aoki, K.; Paulson, J.; Cummings, R.D.; York, W.S.; Karlsson, N.G.; Lisacek, F.; Packer, N.H.; Campbell, M.P.; Aoki, N.P.; et al. GlyYouCan: An accessible glycan structure repository. *Glycobiology* **2017**, *27*, 915–919. [[CrossRef](#)]
94. Hayes, C.A.; Karlsson, N.G.; Struwe, W.B.; Lisacek, F.; Rudd, P.M.; Packer, N.H.; Campbell, M.P. UniCarb-DB: A database resource for glycomics discovery. *Bioinformatics* **2011**, *27*, 1343–1344. [[CrossRef](#)] [[PubMed](#)]
95. Campbell, M.P.; Nguyen-Khuong, T.; Hayes, C.A.; Flowers, S.A.; Alagesan, K.; Kolarich, D.; Packer, N.H.; Karlsson, N.G. Validation of the curation pipeline of UniCarb-DB: Building a global glycan reference MS/MS repository. *Biochim. Biophys. Acta* **2014**, *1844*, 108–116. [[CrossRef](#)]
96. Remoroza, C.A.; Mak, T.D.; De Leoz, M.L.A.; Mirokhin, Y.A.; Stein, S.E. Creating a Mass Spectral Reference Library for Oligosaccharides in Human Milk. *Anal. Chem.* **2018**, *90*, 8977–8988. [[CrossRef](#)]
97. Joshi, H.J.; Harrison, M.J.; Schulz, B.L.; Cooper, C.A.; Packer, N.H.; Karlsson, N.G. Development of a mass fingerprinting tool for automated interpretation of oligosaccharide fragmentation data. *Proteomics* **2004**, *4*, 1650–1664. [[CrossRef](#)]
98. Apte, A.; Meitei, N.S. Bioinformatics in glycomics: Glycan characterization with mass spectrometric data using SimGlycan. *Methods Mol. Biol.* **2010**, *600*, 269–281. [[CrossRef](#)]
99. Maxwell, E.; Tan, Y.; Tan, Y.; Hu, H.; Benson, G.; Aizikov, K.; Conley, S.; Staples, G.O.; Slysz, G.W.; Smith, R.D.; et al. GlycReSoft: A software package for automated recognition of glycans from LC/MS data. *PLoS ONE* **2012**, *7*, e45474. [[CrossRef](#)]
100. Ashline, D.J.; Hanneman, A.J.; Zhang, H.; Reinhold, V.N. Structural documentation of glycan epitopes: Sequential mass spectrometry and spectral matching. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 444–453. [[CrossRef](#)] [[PubMed](#)]
101. Sun, W.; Lajoie, G.A.; Ma, B.; Zhang, K. Bioinformatics Research and Applications. In Proceedings of the 7th International Symposium (ISBRA 2011), Changsha, China, 27–29 May 2011.
102. Hong, P.; Sun, H.; Sha, L.; Pu, Y.; Khatri, K.; Yu, X.; Tang, Y.; Lin, C. GlycoDeNovo—An Efficient Algorithm for Accurate de novo Glycan Topology Reconstruction from Tandem Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2288–2301. [[CrossRef](#)] [[PubMed](#)]
103. Horlacher, O.; Jin, C.; Alocci, D.; Mariethoz, J.; Muller, M.; Karlsson, N.G.; Lisacek, F. Glycoforest 1.0. *Anal. Chem.* **2017**, *89*, 10932–10940. [[CrossRef](#)] [[PubMed](#)]
104. Weatherly, D.B.; Arpinar, F.S.; Porterfield, M.; Tiemeyer, M.; York, W.S.; Ranzinger, R. GRITS Toolbox—a freely available software for processing, annotating and archiving glycomics mass spectrometry data. *Glycobiology* **2019**, *29*, 452–460. [[CrossRef](#)] [[PubMed](#)]
105. Jansen, B.C.; Reiding, K.R.; Bondt, A.; Hipgrave Ederveen, A.L.; Palmblad, M.; Falck, D.; Wührer, M. MassyTools: A High-Throughput Targeted Data Processing Tool for Relative Quantitation and Quality Control Developed for Glycomics and Glycoproteomic MALDI-MS. *J. Proteome Res.* **2015**, *14*, 5088–5098. [[CrossRef](#)] [[PubMed](#)]
106. Ranzinger, R.; Herget, S.; von der Lieth, C.W.; Frank, M. GlycomeDB—a unified database for carbohydrate structures. *Nucleic Acids Res.* **2011**, *39*, D373–D376. [[CrossRef](#)]
107. Konishi, Y.; Aoki-Kinoshita, K.F. The GlycomeAtlas tool for visualizing and querying glycome data. *Bioinformatics* **2012**, *28*, 2849–2850. [[CrossRef](#)]
108. Krambeck, F.J.; Bennun, S.V.; Narang, S.; Choi, S.; Yarema, K.J.; Betenbaugh, M.J. A mathematical model to derive N-glycan structures and cellular enzyme activities from mass spectrometric data. *Glycobiology* **2009**, *19*, 1163–1175. [[CrossRef](#)]
109. Liu, G.; Marathe, D.D.; Matta, K.L.; Neelamegham, S. Systems-level modeling of cellular glycosylation reaction networks: O-linked glycan formation on natural selectin ligands. *Bioinformatics* **2008**, *24*, 2740–2747. [[CrossRef](#)]
110. Hucka, M.; Finney, A.; Sauro, H.M.; Bolouri, H.; Doyle, J.C.; Kitano, H.; Arkin, A.P.; Bornstein, B.J.; Bray, D.; Cornish-Bowden, A.; et al. The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics* **2003**, *19*, 524–531. [[CrossRef](#)]
111. Liu, G.; Puri, A.; Neelamegham, S. Glycosylation Network Analysis Toolbox: A MATLAB-based environment for systems glycomics. *Bioinformatics* **2013**, *29*, 404–406. [[CrossRef](#)] [[PubMed](#)]
112. Bohne, A.; Lang, E.; von der Lieth, C.W. SWEET—WWW-based rapid 3D construction of oligo- and polysaccharides. *Bioinformatics* **1999**, *15*, 767–768. [[CrossRef](#)] [[PubMed](#)]

113. Bohne-Lang, A.; von der Lieth, C.W. GlyProt: In silico glycosylation of proteins. *Nucleic Acids Res.* **2005**, *33*, W214–W219. [[CrossRef](#)] [[PubMed](#)]
114. Lutteke, T.; von der Lieth, C.W. pdb-care (PDB carbohydrate residue check): A program to support annotation of complex carbohydrate structures in PDB files. *BMC Bioinform.* **2004**, *5*, 69. [[CrossRef](#)] [[PubMed](#)]
115. Perez, S.; Sarkar, A.; Rivet, A.; Breton, C.; Imberty, A. Glyco3D: A portal for structural glycosciences. *Methods Mol. Biol.* **2015**, *1273*, 241–258. [[CrossRef](#)] [[PubMed](#)]
116. Ceroni, A.; Maass, K.; Geyer, H.; Geyer, R.; Dell, A.; Haslam, S.M. GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.* **2008**, *7*, 1650–1659. [[CrossRef](#)]
117. Everest-Dass, A.V.; Kolarich, D.; Campbell, M.P.; Packer, N.H. Tandem mass spectra of glycan substructures enable the multistage mass spectrometric identification of determinants on oligosaccharides. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 931–939. [[CrossRef](#)]
118. Chuang, G.Y.; Boyington, J.C.; Joyce, M.G.; Zhu, J.; Nabel, G.J.; Kwong, P.D.; Georgiev, I. Computational prediction of N-linked glycosylation incorporating structural properties and patterns. *Bioinformatics* **2012**, *28*, 2249–2255. [[CrossRef](#)]
119. Zhang, H.; Loriaux, P.; Eng, J.; Campbell, D.; Keller, A.; Moss, P.; Bonneau, R.; Zhang, N.; Zhou, Y.; Wollscheid, B.; et al. UniPep—A database for human N-linked glycosites: A resource for biomarker discovery. *Genome Biol.* **2006**, *7*, R73. [[CrossRef](#)]
120. Chauhan, J.S.; Rao, A.; Raghava, G.P. In silico platform for prediction of N-, O- and C-glycosites in eukaryotic protein sequences. *PLoS ONE* **2013**, *8*, e67008. [[CrossRef](#)]
121. Taherzadeh, G.; Dehzangi, A.; Golchin, M.; Zhou, Y.; Campbell, M.P. SPRINT-Gly: Predicting N- and O-linked glycosylation sites of human and mouse proteins by using sequence and predicted structural properties. *Bioinformatics* **2019**, *35*, 4140–4146. [[CrossRef](#)]
122. Go, E.P.; Rebecchi, K.R.; Dalpathado, D.S.; Bandu, M.L.; Zhang, Y.; Desaire, H. GlycoPep DB: A tool for glycopeptide analysis using a “Smart Search”. *Anal. Chem.* **2007**, *79*, 1708–1713. [[CrossRef](#)]
123. Carrot-Zhang, J.; Chambwe, N.; Damrauer, J.S.; Knijnenburg, T.A.; Robertson, A.G.; Yau, C.; Zhou, W.; Berger, A.C.; Huang, K.L.; Newberg, J.Y.; et al. Comprehensive Analysis of Genetic Ancestry and Its Molecular Correlates in Cancer. *Cancer Cell* **2020**, *37*, 639–654. [[CrossRef](#)]
124. Johansen, M.B.; Kiemer, L.; Brunak, S. Analysis and prediction of mammalian protein glycation. *Glycobiology* **2006**, *16*, 844–853. [[CrossRef](#)] [[PubMed](#)]
125. Gupta, R.; Jung, E.; Gooley, A.A.; Williams, K.L.; Brunak, S.; Hansen, J. Scanning the available Dictyostelium discoideum proteome for O-linked GlcNAc glycosylation sites using neural networks. *Glycobiology* **1999**, *9*, 1009–1022. [[CrossRef](#)] [[PubMed](#)]
126. Seu, C.S.; Chen, Y. Identification of SUMO-binding motifs by NMR. *Methods Mol. Biol.* **2009**, *497*, 121–138. [[CrossRef](#)] [[PubMed](#)]
127. Lutteke, T.; Frank, M.; von der Lieth, C.W. Data mining the protein data bank: Automatic detection and assignment of carbohydrate structures. *Carbohydr. Res.* **2004**, *339*, 1015–1020. [[CrossRef](#)]
128. Bohne-Lang, A.; Lang, E.; Forster, T.; von der Lieth, C.W. LINUCS: Linear notation for unique description of carbohydrate sequences. *Carbohydr Res.* **2001**, *336*, 1–11. [[CrossRef](#)]
129. Farriol-Mathis, N.; Garavelli, J.S.; Boeckmann, B.; Duvaud, S.; Gasteiger, E.; Gateau, A.; Veuthey, A.L.; Bairoch, A. Annotation of post-translational modifications in the Swiss-Prot knowledge base. *Proteomics* **2004**, *4*, 1537–1550. [[CrossRef](#)]
130. York, W.S.; Mazumder, R.; Ranzinger, R.; Edwards, N.; Kahsay, R.; Aoki-Kinoshita, K.F.; Campbell, M.P.; Cummings, R.D.; Feizi, T.; Martin, M.; et al. GlyGen: Computational and Informatics Resources for Glycoscience. *Glycobiology* **2020**, *30*, 72–73. [[CrossRef](#)]
131. Hirabayashi, J.; Tateno, H.; Shikanai, T.; Aoki-Kinoshita, K.F.; Narimatsu, H. The Lectin Frontier Database (LfDB), and data generation based on frontal affinity chromatography. *Molecules* **2015**, *20*, 951–973. [[CrossRef](#)] [[PubMed](#)]
132. Maes, E.; Bonachera, F.; Strecker, G.; Guerardel, Y. SOACS index: An easy NMR-based query for glycan retrieval. *Carbohydr Res.* **2009**, *344*, 322–330. [[CrossRef](#)] [[PubMed](#)]
133. Lisacek, F.; Mariethoz, J.; Alocci, D.; Rudd, P.M.; Abrahams, J.L.; Campbell, M.P.; Packer, N.H.; Stahle, J.; Widmalm, G.; Mullen, E.; et al. Databases and Associated Tools for Glycomics and Glycoproteomics. *Methods Mol. Biol.* **2017**, *1503*, 235–264. [[CrossRef](#)] [[PubMed](#)]

134. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R.D.; Bairoch, A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **2003**, *31*, 3784–3788. [[CrossRef](#)]
135. Lutteke, T.; von der Lieth, C.W. MonoSaccharideDB: A reference resource to unify the notation of carbohydrate residues. *Glycobiology* **2005**, *15*, 1209–1210.
136. Yamada, I.; Shiota, M.; Shinmachi, D.; Ono, T.; Tsuchiya, S.; Hosoda, M.; Fujita, A.; Aoki, N.P.; Watanabe, Y.; Fujita, N.; et al. The GlyCosmos Portal: A unified and comprehensive web resource for the glycosciences. *Nat. Methods* **2020**, *17*, 649–650. [[CrossRef](#)]
137. Li, H.; Xing, X.; Ding, G.; Li, Q.; Wang, C.; Xie, L.; Zeng, R.; Li, Y. SysPTM: A systematic resource for proteomic research on post-translational modifications. *Mol. Cell. Proteom.* **2009**, *8*, 1839–1849. [[CrossRef](#)]
138. Anugraham, M.; Jacob, F.; Nixdorf, S.; Everest-Dass, A.V.; Heinzelmann-Schwarz, V.; Packer, N.H. Specific glycosylation of membrane proteins in epithelial ovarian cancer cell lines: Glycan structures reflect gene expression and DNA methylation status. *Mol. Cell. Proteom.* **2014**, *13*, 2213–2232. [[CrossRef](#)]
139. Munkley, J. The Role of Sialyl-Tn in Cancer. *Int. J. Mol. Sci.* **2016**, *17*, 275. [[CrossRef](#)]
140. Tuccillo, F.M.; de Laurentiis, A.; Palmieri, C.; Fiume, G.; Bonelli, P.; Borrelli, A.; Tassone, P.; Scala, I.; Buonaguro, F.M.; Quinto, I.; et al. Aberrant glycosylation as biomarker for cancer: Focus on CD43. *BioMed Res. Int.* **2014**, *2014*, 742831. [[CrossRef](#)]
141. Yin, B.W.; Dnistrian, A.; Lloyd, K.O. Ovarian cancer antigen CA125 is encoded by the MUC16 mucin gene. *Int. J. Cancer* **2002**, *98*, 737–740. [[CrossRef](#)] [[PubMed](#)]
142. O'Brien, T.J.; Beard, J.B.; Underwood, L.J.; Dennis, R.A.; Santin, A.D.; York, L. The CA 125 gene: An extracellular superstructure dominated by repeat sequences. *Tumour Biol.* **2001**, *22*, 348–366. [[CrossRef](#)]
143. Nustad, K.; Bast, R.C., Jr.; Brien, T.J.; Nilsson, O.; Seguin, P.; Suresh, M.R.; Saga, T.; Nozawa, S.; Borner, O.P.; de Bruijn, H.W.; et al. Specificity and affinity of 26 monoclonal antibodies against the CA 125 antigen: First report from the ISOBM TD-1 workshop. International Society for Oncodevelopmental Biology and Medicine. *Tumour Biol.* **1996**, *17*, 196–219. [[CrossRef](#)] [[PubMed](#)]
144. Yin, B.W.; Lloyd, K.O. Molecular cloning of the CA125 ovarian cancer antigen: Identification as a new mucin, MUC16. *J. Biol. Chem.* **2001**, *276*, 27371–27375. [[CrossRef](#)] [[PubMed](#)]
145. Zurawski, V.R., Jr.; Orjaseter, H.; Andersen, A.; Jellum, E. Elevated serum CA 125 levels prior to diagnosis of ovarian neoplasia: Relevance for early detection of ovarian cancer. *Int. J. Cancer* **1988**, *42*, 677–680. [[CrossRef](#)] [[PubMed](#)]
146. Bast, R.C., Jr.; Klug, T.L.; St John, E.; Jenison, E.; Niloff, J.M.; Lazarus, H.; Berkowitz, R.S.; Leavitt, T.; Griffiths, C.T.; Parker, L.; et al. A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer. *N. Engl. J. Med.* **1983**, *309*, 883–887. [[CrossRef](#)] [[PubMed](#)]
147. Bast, R.C., Jr.; Badgwell, D.; Lu, Z.; Marquez, R.; Rosen, D.; Liu, J.; Baggerly, K.A.; Atkinson, E.N.; Skates, S.; Zhang, Z.; et al. New tumor markers: CA125 and beyond. *Int. J. Gynecol. Cancer* **2005**, *15* (Suppl. S3), 274–281. [[CrossRef](#)]
148. Gostout, B.S.; Brewer, M.A. Guidelines for referral of the patient with an adnexal mass. *Clin. Obstet. Gynecol.* **2006**, *49*, 448–458. [[CrossRef](#)]
149. Pauler, D.K.; Menon, U.; McIntosh, M.; Symecko, H.L.; Skates, S.J.; Jacobs, I.J. Factors influencing serum CA125II levels in healthy postmenopausal women. *Cancer Epidemiol. Biomark. Prev.* **2001**, *10*, 489–493.
150. Lauro, S.; Trasatti, L.; Bordin, F.; Lanzetta, G.; Bria, E.; Gelibter, A.; Reale, M.G.; Vecchione, A. Comparison of CEA, MCA, CA 15-3 and CA 27-29 in follow-up and monitoring therapeutic response in breast cancer patients. *Anticancer Res.* **1999**, *19*, 3511–3515.
151. Uehara, M.; Kinoshita, T.; Hojo, T.; Akashi-Tanaka, S.; Iwamoto, E.; Fukutomi, T. Long-term prognostic study of carcinoembryonic antigen (CEA) and carbohydrate antigen 15-3 (CA 15-3) in breast cancer. *Int. J. Clin. Oncol.* **2008**, *13*, 447–451. [[CrossRef](#)]
152. Kumpulainen, E.J.; Kesikuru, R.J.; Johansson, R.T. Serum tumor marker CA 15.3 and stage are the two most powerful predictors of survival in primary breast cancer. *Breast Cancer Res. Treat.* **2002**, *76*, 95–102. [[CrossRef](#)] [[PubMed](#)]
153. Ebeling, F.G.; Stieber, P.; Untch, M.; Nagel, D.; Konecny, G.E.; Schmitt, U.M.; Fateh-Moghadam, A.; Seidel, D. Serum CEA and CA 15-3 as prognostic factors in primary breast cancer. *Br. J. Cancer* **2002**, *86*, 1217–1222. [[CrossRef](#)] [[PubMed](#)]

154. Harris, L.; Fritsche, H.; Mennel, R.; Norton, L.; Ravdin, P.; Taube, S.; Somerfield, M.R.; Hayes, D.F.; Bast, R.C., Jr.; American Society of Clinical, O. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J. Clin. Oncol.* **2007**, *25*, 5287–5312. [[CrossRef](#)] [[PubMed](#)]
155. Locker, G.Y.; Hamilton, S.; Harris, J.; Jessup, J.M.; Kemeny, N.; Macdonald, J.S.; Somerfield, M.R.; Hayes, D.F.; Bast, R.C., Jr.; Asco. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J. Clin. Oncol.* **2006**, *24*, 5313–5327. [[CrossRef](#)]
156. Safi, F.; Schlosser, W.; Kolb, G.; Beger, H.G. Diagnostic value of CA 19-9 in patients with pancreatic cancer and nonspecific gastrointestinal symptoms. *J. Gastrointest. Surg.* **1997**, *1*, 106–112. [[CrossRef](#)]
157. Ychou, M.; Duffour, J.; Kramar, A.; Gourgou, S.; Grenier, J. Clinical significance and prognostic value of CA72-4 compared with CEA and CA19-9 in patients with gastric cancer. *Dis. Markers* **2000**, *16*, 105–110. [[CrossRef](#)]
158. Marrelli, D.; Pinto, E.; De Stefano, A.; de Manzoni, G.; Farnetani, M.; Garosi, L.; Roviello, F. Preoperative positivity of serum tumor markers is a strong predictor of hematogenous recurrence of gastric cancer. *J. Surg. Oncol.* **2001**, *78*, 253–258. [[CrossRef](#)]
159. Duraker, N.; Celik, A.N. The prognostic significance of preoperative serum CA 19-9 in patients with resectable gastric carcinoma: Comparison with CEA. *J. Surg. Oncol.* **2001**, *76*, 266–271. [[CrossRef](#)]
160. Reiter, W.; Stieber, P.; Reuter, C.; Nagel, D.; Cramer, C.; Pahl, H.; Fateh-Moghadam, A. Prognostic value of preoperative serum levels of CEA, CA 19-9 and CA 72-4 in gastric carcinoma. *Anticancer Res.* **1997**, *17*, 2903–2906.
161. Kirwan, A.; Utratna, M.; O'Dwyer, M.E.; Joshi, L.; Kilcoyne, M. Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. *BioMed Res. Int.* **2015**, *2015*, 490531. [[CrossRef](#)] [[PubMed](#)]
162. Kailemia, M.J.; Park, D.; Lebrilla, C.B. Glycans and glycoproteins as specific biomarkers for cancer. *Anal. Bioanal. Chem.* **2017**, *409*, 395–410. [[CrossRef](#)] [[PubMed](#)]
163. Reis, C.A.; Osorio, H.; Silva, L.; Gomes, C.; David, L. Alterations in glycosylation as biomarkers for cancer detection. *J. Clin. Pathol.* **2010**, *63*, 322–329. [[CrossRef](#)] [[PubMed](#)]
164. Peixoto, A.; Relvas-Santos, M.; Azevedo, R.; Santos, L.L.; Ferreira, J.A. Protein Glycosylation and Tumor Microenvironment Alterations Driving Cancer Hallmarks. *Front. Oncol.* **2019**, *9*, 380. [[CrossRef](#)]
165. Wang, M.; Zhu, J.; Lubman, D.M.; Gao, C. Aberrant glycosylation and cancer biomarker discovery: A promising and thorny journey. *Clin. Chem. Lab. Med.* **2019**, *57*, 407–416. [[CrossRef](#)]
166. Paulson, J.C.; Colley, K.J. Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *J. Biol. Chem.* **1989**, *264*, 17615–17618.
167. Nakano, M.; Saldanha, R.; Gobel, A.; Kavallaris, M.; Packer, N.H. Identification of glycan structure alterations on cell membrane proteins in desoxyepothilone B resistant leukemia cells. *Mol. Cell. Proteom.* **2011**, *10*. [[CrossRef](#)]
168. Alley, W.R.; Vasseur, J.A.; Goetz, J.A.; Syoboda, M.; Mann, B.F.; Matei, D.E.; Menning, N.; Hussein, A.; Mechref, Y.; Novotny, M.V. N-linked Glycan Structures and Their Expressions Change in the Blood Sera of Ovarian Cancer Patients. *J. Proteome Res.* **2012**, *11*, 2282–2300. [[CrossRef](#)]
169. Shetty, V.; Hafner, J.; Shah, P.; Nickens, Z.; Philip, R. Investigation of ovarian cancer associated sialylation changes in N-linked glycopeptides by quantitative proteomics. *Clin. Proteom.* **2012**, *9*, 10. [[CrossRef](#)]
170. Colombo, N.; Peiretti, M.; Parma, G.; Lapresa, M.; Mancari, R.; Carinelli, S.; Sessa, C.; Castiglione, M.; Group, E.G.W. Newly diagnosed and relapsed epithelial ovarian carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **2010**, *21* (Suppl. S5), v23–v30. [[CrossRef](#)]
171. Schiess, R.; Wollscheid, B.; Aebersold, R. Targeted proteomic strategy for clinical biomarker discovery. *Mol. Oncol.* **2009**, *3*, 33–44. [[CrossRef](#)] [[PubMed](#)]
172. Dube, D.H.; Bertozzi, C.R. Glycans in cancer and inflammation—potential for therapeutics and diagnostics. *Nat. Rev. Drug Discov.* **2005**, *4*, 477–488. [[CrossRef](#)] [[PubMed](#)]
173. Lauc, G.; Huffman, J.E.; Pucic, M.; Zgaga, L.; Adamczyk, B.; Muzinic, A.; Novokmet, M.; Polasek, O.; Gornik, O.; Kristic, J.; et al. Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. *PLoS Genet.* **2013**, *9*, e1003225. [[CrossRef](#)] [[PubMed](#)]
174. Munkley, J. Glycosylation is a global target for androgen control in prostate cancer cells. *Endocr. Relat. Cancer* **2017**, *24*, R49–R64. [[CrossRef](#)] [[PubMed](#)]

175. Saldova, R.; Shehni, A.A.; Haakensen, V.D.; Steinfeld, I.; Hilliard, M.; Kifer, I.; Helland, A.; Yakhini, Z.; Borresen-Dale, A.L.; Rudd, P.M. Association of N-Glycosylation with Breast Carcinoma and Systemic Features Using High-Resolution Quantitative UPLC. *J. Proteome Res.* **2014**, *13*, 2314–2327. [[CrossRef](#)]
176. Arnold, J.N.; Saldova, R.; Galligan, M.C.; Murphy, T.B.; Mimura-Kimura, Y.; Telford, J.E.; Godwin, A.K.; Rudd, P.M. Novel Glycan Biomarkers for the Detection of Lung Cancer. *J. Proteome Res.* **2011**, *10*, 1755–1764. [[CrossRef](#)]
177. Albrecht, S.; Unwin, L.; Muniyappa, M.; Rudd, P.M. Glycosylation as a marker for inflammatory arthritis. *Cancer Biomark.* **2014**, *14*, 17–28. [[CrossRef](#)]
178. Kaur, N.; Mathur, A.; Aggarwal, A.; Gupta, S.; Rashmi, T. Bioinformatic Analysis of Aberrant Glycosylation in Triple Negative Breast Cancer. In Proceedings of the 6th World Congress on Breast Cancer & Therapy, San Francisco, CA, USA, 16–18 October 2017.
179. Gotz, L.; Abrahams, J.L.; Mariethoz, J.; Rudd, P.M.; Karlsson, N.G.; Packer, N.H.; Campbell, M.P.; Lisacek, F. GlycoDigest: A tool for the targeted use of exoglycosidase digestions in glycan structure determination. *Bioinformatics* **2014**, *30*, 3131–3133. [[CrossRef](#)]
180. Zhang, D.; Xie, Q.; Wang, Q.; Wang, Y.; Miao, J.; Li, L.; Zhang, T.; Cao, X.; Li, Y. Mass spectrometry analysis reveals aberrant N-glycans in colorectal cancer tissues. *Glycobiology* **2019**, *29*, 372–384. [[CrossRef](#)]
181. Lee, S.B.; Bose, S.; Ahn, S.H.; Son, B.H.; Ko, B.S.; Kim, H.J.; Chung, I.Y.; Kim, J.; Lee, W.; Ko, M.S.; et al. Breast cancer diagnosis by analysis of serum N-glycans using MALDI-TOF mass spectroscopy. *PLoS ONE* **2020**, *15*, e0231004. [[CrossRef](#)]
182. Oriol, R.; Mollicone, R.; Cailleau, A.; Balanzino, L.; Breton, C. Divergent evolution of fucosyltransferase genes from vertebrates, invertebrates, and bacteria. *Glycobiology* **1999**, *9*, 323–334. [[CrossRef](#)] [[PubMed](#)]
183. Goletz, S.; Danielczyk, A.; Stahn, R.; Karsten, U.; Stoeckl, L.; Loeffler, A.; Hillemann, A.; Baumeister, H. GlycoOptimization for Fully Human and Largely Improved Biopharmaceutical Antibodies and Proteins. In Proceedings of the 1st Beilstein Glyco-Bioinformatics Symposium—Bits ‘n’ Bytes of Sugars, Potsdam, Germany, 4–8 October 2009; pp. 87–100.
184. Nadeem, T.; Khan, M.A.; Ijaz, B.; Ahmed, N.; Rahman, Z.U.; Latif, M.S.; Ali, Q.; Rana, M.A. Glycosylation of Recombinant Anticancer Therapeutics in Different Expression Systems with Emerging Technologies. *Cancer Res.* **2018**, *78*, 2787–2798. [[CrossRef](#)] [[PubMed](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).