# Toolbox Accelerating Glycomics (TAG): Glycan Annotation from MALDI-TOF MS Spectra and Mapping Expression Variation to Biosynthetic Pathways

**Nobuaki Miura [1],\*** , **Hisatoshi Hanamatsu [2], Ikuko Yokota [2], Kazue Okada [2], Jun-Ichi Furukawa [2]** and **Yasuro Shinohara [3],\***

[1] Division of Bioinformatics, Niigata University Graduate School of Medical and Dental Sciences, 1-757 Asahimachi-dori, Chuo-ku, Niigata 951-8510, Japan

[2] Department of Advanced Clinical Glycobiology, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, Kita21, Nishi11, Kita-ku, Sapporo 001-0021, Japan; h_hanamatsu@med.hokudai.ac.jp (H.H.); iyokota@med.hokudai.ac.jp (I.Y.); okada@soyaku.co.jp (K.O.); j_furu@med.hokudai.ac.jp (J.-I.F.)

[3] Department of Pharmacy, Kinjo Gakuin University, Nagoya 463-8521, Japan

\* Correspondence: muller@med.niigata-u.ac.jp (N.M.); yshinohara@kinjo-u.ac.jp (Y.S.)

check for updates

**Abstract:** Glycans present extraordinary structural diversity commensurate with their involvement in numerous fundamental cellular processes including growth, differentiation, and morphogenesis. Unlike linear DNA and protein sequences, glycans have heterogeneous structures that differ in composition, branching, linkage, and anomericity. These differences pose a challenge to developing useful software for glycomic analysis. To overcome this problem, we developed the novel Toolbox Accelerating Glycomics (TAG) program. TAG consists of three units: 'TAG List' creates a glycan list that is used for database searching in TAG Expression; 'TAG Expression' automatically annotates and quantifies glycan signals and draws graphs; and 'TAG Pathway' maps the obtained expression information to biosynthetic pathways. Herein, we discuss the concepts, outline the TAG process, and demonstrate its potential using glycomic expression profile data from Chinese hamster ovary (CHO) cells and mutants lacking a functional *Npc1* gene (*Npc1* knockout (KO) CHO cells). TAG not only drastically reduced the amount of time and labor needed for glycomic analysis but also detected and quantified more glycans than manual analysis. Although this study was limited to the analysis of *N*-glycans and free oligosaccharides, the glycomic platform will be expanded to facilitate the analysis of *O*-glycans and glycans of glycosphingolipids.

**Keywords:** glycomics; automated annotation; MALDI-TOF MS; biosynthetic pathways; glycan diversity; informatics; expression variation

## 1. Introduction

Cell surfaces are coated with a variety of intricately arranged glycoconjugates such as glycoproteins and glycolipids, and glycosylation is thought to be essential for maintaining homeostasis in mammalian cells [1,2]. As omics approaches such as transcriptomics and proteomics have contributed enormously to our understanding of various biological processes in the postgenomic era, rapid and precise analysis of cellular glycomics is attracting a lot of attention. Many useful software packages have been developed for transcriptomics and proteomics that mine useful information from next-generation sequencer and mass spectrometry (MS) data [3–6]. However, similar programs for glycomic analysis

are scarce, which prevents the expansion of the field. Although glycan sequences and structures can be determined from tandem MS (MS/MS) data using dedicated software [7], tools for annotating glycan signals on mass spectra via large-scale quantitative analysis are not currently available. Glycome profiles differ between species, tissues, and cells, which makes it difficult to determine the number of glycans and their structures [8]. Distinguishing genuine glycan signals from contaminating signals can be difficult. There are various methods for modifying carboxylic acid of sialic acid and for derivatizing the reducing termini of glycans for glycomic analysis. Thus, glycans are empirically annotated from raw mass data, and quantitative analysis calculations are often carried out manually using spreadsheet software such as Microsoft Excel. Such manual approaches are too time-consuming and labor-intensive for large-scale glycomic analysis. Furthermore, manual approaches carry an increased risk of errors such as improper annotation, missing data, and improper sorting of results between analyses. Lack of a comprehensive glycan list and obtaining a non-standard resultant format owing to manual analyses can cause glycomics to be difficult.

Due to these difficulties, glycomic research is somewhat behind DNA- and protein-based approaches. Therefore, the development of appropriate software for automated data analysis, even if only partially automated, could accelerate glycomic research. Indeed, the automation of glycan annotation is a major goal in this field. In the present study, we developed novel software named Toolbox Accelerating Glycomics (TAG). TAG is composed of three functions that assist glycomic analysis: TAG List automatically creates a semi-comprehensive glycan list for database searching during glycan annotation of MS spectra; TAG Expression annotates glycan signals on MALDI-TOF MS spectra based on the glycan list, quantifies the annotated glycans, and provides outputs in a standard format; and TAG Pathway maps glycan expression variation to glycan biosynthetic pathways. The glycan list made with TAG List is defined as semi-comprehensive in the sense that it includes all *N*-glycans and free *N*-glycans in a certain range based on known biosynthetic pathways, but not all glycans can be listed universally. Glycan biosynthetic pathways can be visualized using the KEGG Pathway, Roche Biochemical Pathway, and Reactome databases [9–11], but these databases include only a fraction of all representative pathways. The pathways employed in TAG are not also comprehensive, but the TAG approach is modifiable and expandable by users. In addition, we employed a fused pathway in which *N*-glycans and free oligosaccharides (FOSs) are interconnected, and we viewed glycan expression variation data obtained by TAG Expression using this fused pathway. The usefulness of TAG was validated using glycomic expression profile data from Chinese hamster ovary (CHO) cells and mutants lacking the *Npc1* gene (*Npc1* knockout (KO) CHO cells).

## 2. Materials and Methods

### 2.1. Overview of Toolbox Accelerating Glycomics (TAG)

An overview of glycomic analysis using TAG is shown in Figure 1. A multiple-tabbed Excel file containing sets of peak positions and peak area obtained from MALDI-TOF MS data provides the input. In the Excel file, each tab corresponds to a separate measurement, each MS measurement is called an 'experiment', and a 'series' is defined as a group of experiments with certain conditions, such as wild type, disease, gene KO cell type, etc. For example, nine tabs are obtained for n = 3 experiments with three different series, resulting in nine comma-delimited files (CSVs). The current version of TAG has the functions listed in Table 1. TAG Expression annotates glycans using these CSV files, and glycans are listed by database searching. The TAG List constructs semi-comprehensive glycan lists that are used by TAG Expression. Annotated glycans are quantified based on the amount of added glycan internal standard. The resulting glycan structure and expression level data allow automated statistical analysis approaches such as calculating averages and standard deviations as well as performing Student's t-tests and hierarchical cluster analysis. The results are outputted as CSV files, and commercially available spreadsheet software can be used to examine the analyzed data. TAG Expression also generates an input file for cluster analysis by Cluster 3.0 software [12,13].

The results of cluster analysis can be viewed using Java Treeview [14]. The TAG Pathway visualizes the observed variation in glycan expression in the results from TAG Expression based on glycan biosynthetic pathways. The TAG Pathway can be used for *N*-glycan, FOS, and combined analyses with *N*-glycan+FOS. With respect to FOS, we focus on the free *N*-glycans (FNGs) in this study.
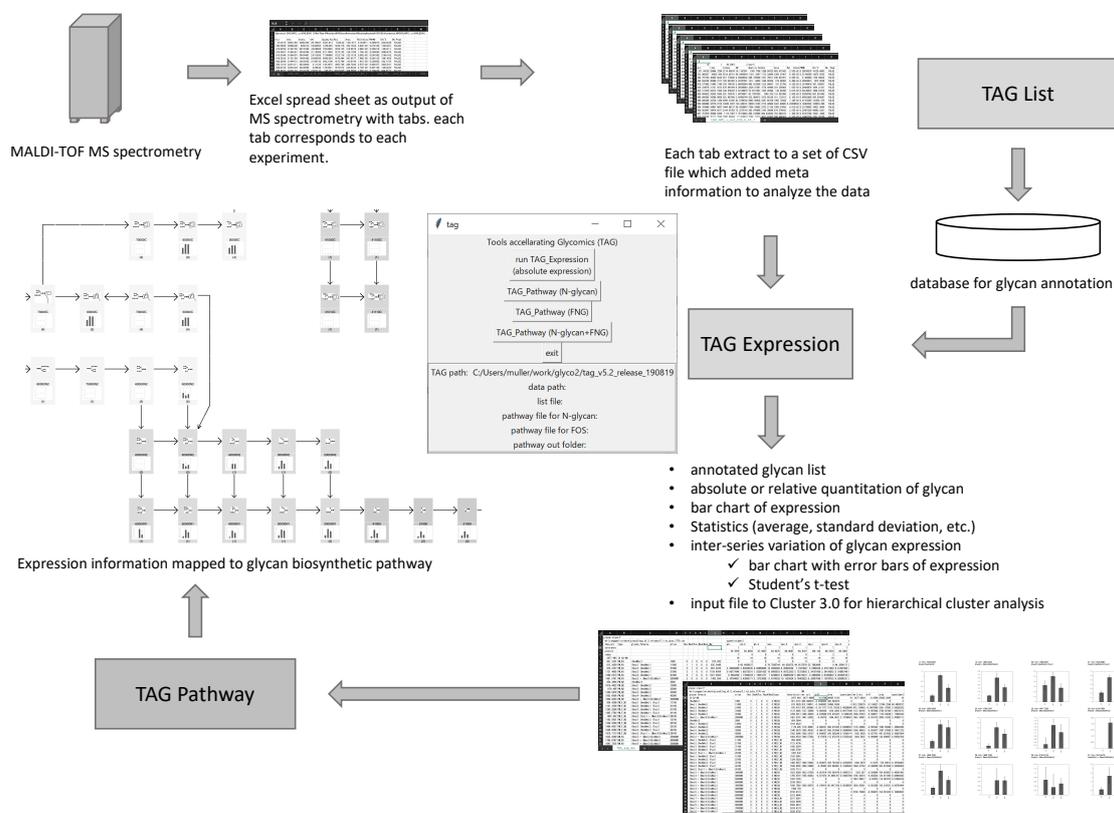


**Figure 1.** Overview of Toolbox Accelerating Glycomics (TAG).

**Table 1.** Functions of Toolbox Accelerating Glycomics (TAG).

| Button Name | Function |
| --- | --- |
| TAG List | Generating a glycan list to be employed in TAG Expression. TAG List is a separate program written in the awk script language. |
| TAG Expression | Glycan annotation, quantitation, statistical analysis, graphing. |
| TAG Pathway (*N*-glycan) | Visualizing variation in *N*-glycan expression based on *N*-glycan biosynthetic pathways. |
| TAG Pathway (FOS) | Visualizing variation in FNG expression based on FNG biosynthetic pathways. |
| TAG Pathway (*N*-glycan + FOS) | Visualizing variation in both *N*-glycans and FNGs based on *N*-glycan and FNG biosynthetic pathways. |

## 2.2. Implementation of TAG

### 2.2.1. Interface

Although TAG is essentially a set of program codes that run on a Microsoft Windows platform, some of the programs have been ported to the Mac OSX platform. The TAG interface is written in the Tcl/Tk language. The startup screen, shown in the center of Figure 1, includes various buttons for carrying out expression analysis and generating a glycan list by selecting the files needed for each

process. The current version of TAG requires Gnuplot [15], Cluster 3.0, and gawk [16] for graphing, hierarchical cluster analysis, and running the TAG List and TAG Pathway functions, respectively.

TAG Expression is written in FORTRAN 77, and the executable modules were made with GNU Fortran on the Cygwin platform. The TAG List and TAG Pathway are written as awk scripts, TAG List programs are typically run using command lines, and TAG Pathway procedures are controlled via the tcl/tk startup screen.

### 2.2.2. TAG List

Figure 2 shows the glycan list generated by the TAG List that is used by TAG Expression to automatically assign glycan signals. An 'N' or 'F' character in the first row indicated whether the list contains *N*-glycan or FNG data. The labels and *m/z* values of internal standards are in the second row and separated by commas. The glycan composition, *m/z*, and types summarized in Table S1 (*N*-glycans) and Table S2 (FNGs) are below the third row; these are also separated by commas. A typical list for *N*-glycan and FNG data is made using the TAG List program.

| | A | B | C |
|---|---|---|---|
| 1 | N | | |
| 2 | IS A2GN | 2477.963 | |
| 3 | (Hex) 1 (HexNAc) 2 | 1016.4422 | PM |
| 4 | (Hex) 2 (HexNAc) 2 | 1178.495 | PM |
| 5 | (Hex) 3 (HexNAc) 2 | 1340.5478 | PM |
| 6 | (Hex) 4 (HexNAc) 2 | 1502.6006 | PM |
| 7 | (Hex) 2 + (Man) 3 (GlcNAc) 2 | 1664.6534 | HM |
| 8 | (Hex) 1 (HexNAc) 2 (Fuc) 1 | 1162.5001 | PM_F |
| 9 | (Hex) 2 (HexNAc) 2 (Fuc) 1 | 1324.5529 | PM_F |
| 10 | (Hex) 3 (HexNAc) 2 (Fuc) 1 | 1486.6057 | PM_F |
| 11 | (Hex) 4 (HexNAc) 2 (Fuc) 1 | 1648.6585 | PM_F |
| 12 | (Hex) 2 (Fuc) 1 + (Man) 3 (GlcNAc) 2 | 1810.7113 | HM_F |
| 13 | (Hex) 3 + (Man) 3 (GlcNAc) 2 | 1826.7062 | HM |
| 14 | (Hex) 4 + (Man) 3 (GlcNAc) 2 | 1988.759 | HM |
| 15 | (Hex) 5 + (Man) 3 (GlcNAc) 2 | 2150.8118 | HM |
| 16 | (Hex) 6 + (Man) 3 (GlcNAc) 2 | 2312.8646 | HM |
| 17 | (Hex) 7 + (Man) 3 (GlcNAc) 2 | 2474.9174 | HM_G |
| 18 | (Hex) 8 + (Man) 3 (GlcNAc) 2 | 2636.9702 | HM_G |
| 19 | (Hex) 9 + (Man) 3 (GlcNAc) 2 | 2799.023 | HM_G |
| 20 | (HexNAc) 1 + (Man) 3 (GlcNAc) 2 | 1543.6271 | 1_n |

| | A | B | C |
|---|---|---|---|
| 1 | F | | |
| 2 | IS A2GN | 2477.963 | |
| 3 | (HexNAc) 1 | 651.3101 | PM_N1 |
| 4 | (Hex) 1 (HexNAc) 1 | 813.3629 | PM_N1 |
| 5 | (Hex) 2 (HexNAc) 1 | 975.4157 | PM_N1 |
| 6 | (Hex) 3 (HexNAc) 1 | 1137.4685 | PM_N1 |
| 7 | (Hex) 4 (HexNAc) 1 | 1299.5213 | PM_N1 |
| 8 | (Hex) 2 + (Man) 3 (GlcNAc) 1 | 1461.5741 | HM_N1 |
| 9 | (HexNAc) 2 | 854.3894 | PM_N2 |
| 10 | (Hex) 1 (HexNAc) 2 | 1016.4422 | PM_N2 |
| 11 | (Hex) 2 (HexNAc) 2 | 1178.495 | PM_N2 |
| 12 | (Hex) 3 (HexNAc) 2 | 1340.5478 | PM_N2 |
| 13 | (Hex) 4 (HexNAc) 2 | 1502.6006 | PM_N2 |
| 14 | (Hex) 2 + (Man) 3 (GlcNAc) 2 | 1664.6534 | HM_N2 |
| 15 | (Hex) 1 (HexNAc) 1 (Fuc) 1 | 959.4208 | PM_F_N1 |
| 16 | (Hex) 2 (HexNAc) 1 (Fuc) 1 | 1121.4736 | PM_F_N1 |
| 17 | (Hex) 3 (HexNAc) 1 (Fuc) 1 | 1283.5264 | PM_F_N1 |
| 18 | (Hex) 4 (HexNAc) 1 (Fuc) 1 | 1445.5792 | PM_F_N1 |
| 19 | (Hex) 2 (Fuc) 1 + (Man) 3 (GlcNAc) 1 | 1607.632 | HM_F_N1 |
| 20 | (Hex) 1 (HexNAc) 2 (Fuc) 1 | 1162.5001 | PM_F_N2 |

**Figure 2.** Examples of glycan lists (*N*-glycans on the left, free *N*-glycans (FNGs) on the right) generated by TAG List.

*N*-glycans that are mainly composed of mannose (Man), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), galactose (Gal), fucose (Fuc), and sialic acid (*N*-acetylneumraminic (NeuAc) acid and *N*-glycolylneuraminic acid (NeuGc) are classified into four groups: high (oligo) mannose type (HM, (Man)5-9(GlcNAc)2); pauci mannose type (PM, (Man)1-4(GlcNAc)2(Fuc)0-1); complex type containing various numbers of Gal, GlcNAc, sialic acid (NeuAc or NeuGc), Fuc, and other residues linked to a (Man)3(GlcNAc)2 *N*-glycan core structure, and hybrid type, which is a hybrid of high mannose and complex oligosaccharide. Mono- to tri-glucosylated HMs ((Glc)1-3(Man)9(GlcNAc)2), which are precursors of HM glycans, also exist. In any case, the core structure consists of three Man residues and two GlcNAc residues ((Man)3(GlcNAc)2). To evaluate the glycan expression profile, it is useful to further classify the structures of *N*-glycans according to the number of branches, neutral or acidic, presence or absence of fucose residues, and other features. For this purpose, we classified *N*-glycans into 43 types as defined in Table S1. In the case of FNGs, they were classified into 82 (Table S2). A major difference is that FNGs have two types of core structure: (Man)3 (GlcNAc)2 (N2 core) and (Man)3 (GlcNAc)1 (N1 core). In the current study, we employed high-mannose, pauci-mannose, hybrid, and complex glycans with mono- to tetra-antennae, with up to three Fuc and up to two acetyl modifications, resulting in TAG List files containing 900 *N*-glycans and 1770 FNGs.

Although the structures of complex *N*-glycans are highly diverse because the core structure is modified with various monosaccharides, there are some general rules regarding the structure of complex type *N*-glycans [2]. For example, core mannose is modified with GlcNAc(s), and the

number of GlcNAc groups determines the number of branches (typically one to four). GlcNAc is often modified with Gal (and/or GalNAc) to generate LacNAc or LacdiNAc structures, respectively. Fucosylation occurs on the non-reducing end of Gal (H-type), on GlcNAc in branches (lewis type outer arm fucosylation), and on reducing terminal GlcNAc (core fucosylation). Sialylation often occurs on terminal Gal groups with $\alpha$2-3 and $\alpha$2-6-linkages. Hybrid *N*-glycans contain mannose residues ((Man)1-3) on one side of the branch, while the other side has a complex-type sugar chain containing GlcNAc modified by $\alpha$-1,3-mannosyl-glycoprotein 2-$\beta$-*N*-acetylglucosaminyltransferase. Using the resulting information, we defined structures of complex and hybrid *N*-glycans to generate a semi-comprehensive *N*-glycan list. In addition to these structures, other modifications such as polylactosamine structure, sulfates, glucuronic acids, and $\alpha$-2-8-linked polysialic acid structures may also exist, but these will be considered in future versions of the program, and they were not considered in the current work.

To make the annotation of glycans more concise, we employed five-digit notation, in which each digit expresses the number of Hex (Man, Gal, Glc), HexNAc (GlcNAc, GalNAc), Fuc, NeuAc, and NeuGc groups, excluding core structures. In case of *N*-glycans, (Man)3(GlcNAc)2 core structures are shown as 'C', while in the case of FNGs (Man)3(GlcNAc)2, core structures are shown as 'N2', and (Man)3(GlcNAc)1 core structures are shown as 'N1'. The *m/z* value for each glycan was calculated by assuming that the glycans were derivatized with aoWR (an MS-friendly tag with a high proton affinity), and carboxylic acids of sialic acids were methylesterified. [17] Since the TAG list is a simple awk script, it is easy to modify, allowing users to analyze data obtained using other labeling reagents and modifications.

### 2.2.3. TAG Expression

CSV input files are extracted from Excel files (xls files) outputted from MS instruments. Since almost all MALDI-TOF MS instruments can export *m/z* and signal strength (e.g., height, area) values of spectra signals as xls files, we chose to use the exported xls files as source data for TAG. An example of an input file, a Tab Separated Values file produced by FlexAnalysis 3.0 (Bruker Daltonics, Bremen, Germany), is shown in Figure 3. In this mass-list file, rows 1 to 3 are used as headers, and rows 4 and below contain data obtained by MALDI-TOF MS analysis (e.g., *m/z*, time, intensity, etc.). The current version of TAG utilizes *m/z* (1st column) and peak area (7th column) values.

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Spectrum: CHO_WT_-_n=1_2¥0_A5¥1 1SRef Raw (¥¥tsclient¥C¥Users¥shinoken¥Desktop¥yokota¥150728_Kumamoto_N¥CHO_WT_-_n=1_2¥0_A5¥1) | | | | | | | | | | | |
| 2 | 5 | | 2 | 17.0462 | | 1 wt- | | | | | | |
| 3 | m/z | time | Intens. | SN | Quality Fac | Res. | Area | Rel. Intens | FWHM | Chi^2 | Bk. Peak | |
| 4 | 371.933902 | 29994.4312 | 1057.02446 | 9.27755578 | 901.318813 | 1354.78275 | 405.043875 | 5.52E-02 | 0.27453398 | 5074.75827 | FALSE | |
| 5 | 372.982979 | 30036.4032 | 2835.02979 | 24.8708036 | 6241.39797 | 1216.13387 | 1164.95528 | 0.14803197 | 0.30669566 | 5314.89986 | FALSE | |
| 6 | 394.600625 | 30888.4997 | 299.181037 | 2.57081923 | 299.967749 | 1440.81111 | 112.571559 | 1.56E-02 | 0.27387395 | 366.967776 | FALSE | |
| 7 | 416.322194 | 31721.4686 | 277.467115 | 2.36286978 | 327.523771 | 1611.67872 | 100.802826 | 1.45E-02 | 0.25831587 | 251.538758 | FALSE | |
| 8 | 460.741688 | 33359.6749 | 225.452088 | 1.91246743 | 72.3547795 | 936.852364 | 165.308706 | 1.18E-02 | 0.49179754 | 1130.0737 | FALSE | |
| 9 | 489.793767 | 34388.63 | 310.761287 | 2.64422963 | 192.092042 | 971.050996 | 209.677113 | 1.62E-02 | 0.50439552 | 683.473098 | FALSE | |
| 10 | 499.789887 | 34735.5696 | 3176.69793 | 27.0279262 | 2041.75011 | 2111.73877 | 1136.35632 | 0.16587228 | 0.23667221 | 85595.0972 | FALSE | |
| 11 | 500.773421 | 34769.517 | 10116.6128 | 86.0650008 | 19915.4782 | 2455.84702 | 3197.13522 | 0.5282421 | 0.20391067 | 105438.148 | FALSE | |
| 12 | 503.003213 | 34846.3568 | 2165.55108 | 18.4239271 | 250.215585 | 1991.42365 | 718.312988 | 0.11307493 | 0.25258473 | 68322.4512 | FALSE | |
| 13 | 503.97475 | 34879.783 | 6239.2702 | 52.9718087 | 2981.18338 | 2451.24087 | 1738.71648 | 0.32578544 | 0.20559985 | 49965.9673 | FALSE | |
| 14 | 504.771905 | 34907.1856 | 2151.39822 | 18.2967291 | 1122.92348 | 2360.26603 | 655.477557 | 0.11233593 | 0.21386229 | 16352.2451 | FALSE | |
| 15 | 505.895025 | 34945.7565 | 401.850572 | 3.41974498 | 92.5747716 | 1181.94944 | 225.88703 | 2.10E-02 | 0.42801748 | 7222.5017 | FALSE | |
| 16 | 506.797358 | 34976.714 | 3667.92051 | 31.1975458 | 5660.674 | 2203.85002 | 1160.5866 | 0.19152162 | 0.22996 | 7980.41949 | FALSE | |
| 17 | 507.727927 | 35008.6113 | 862.634885 | 7.35358755 | 1095.16542 | 956.056743 | 637.91536 | 4.50E-02 | 0.53106464 | 4419.51936 | FALSE | |
| 18 | 508.769903 | 35044.2927 | 416.548457 | 3.53618923 | 303.926967 | 2162.77263 | 155.290449 | 2.18E-02 | 0.23523966 | 1589.01111 | FALSE | |
| 19 | 509.788473 | 35079.1371 | 302.392296 | 2.57542386 | 64.5081078 | 1979.70745 | 112.533143 | 0.01578951 | 0.25750697 | 1901.78933 | FALSE | |
| 20 | 510.758895 | 35112.3021 | 730.958313 | 6.2097764 | 623.15645 | 2377.25653 | 241.7814 | 3.82E-02 | 0.21485224 | 2195.08722 | FALSE | |
| 21 | 511.757616 | 35146.4013 | 1015.05373 | 8.63497846 | 926.845481 | 2861.32574 | 279.794123 | 5.30E-02 | 0.17885332 | 2585.41625 | FALSE | |
| 22 | 512.845201 | 35183.4967 | 518.486619 | 4.41513554 | 208.530672 | 1918.12289 | 194.179237 | 2.71E-02 | 0.26736827 | 3303.65928 | FALSE | |
| 23 | 516.029872 | 35291.8938 | 7827.20523 | 66.666402 | 11814.1571 | 2708.37276 | 2054.9631 | 0.40869997 | 0.19053133 | 31690.1249 | FALSE | |
| 24 | 517.989859 | 35358.4397 | 1757.78865 | 14.9710105 | 322.745419 | 2928.50692 | 432.346579 | 9.18E-02 | 0.17687848 | 34410.5263 | FALSE | |

**Figure 3.** Input file for TAG Expression, a Tab Separated Values file produced by FlexAnalysis 3.0.

Some additional information (metadata) for analysis, such as the quantity of added internal standards (pmol), the error tolerance for precursor *m/z* values, and the amount of protein used for analysis (glycan expression is normalized per 100 µg protein), must be entered in the second row of the CSV by the user, as shown in Table 2. TAG Expression utilizes the glycan list generated by TAG List to annotate glycan signals.

**Table 2.** Metadata for Glycan Analysis Using TAG Expression.

| Cell Position | Data Type | Data |
|---|---|---|
| A2 | Real number | Quantity of internal standard. |
| B2 | Real number | Error tolerance for precursor *m/z*. |
| C2 | Real number | Total quantity (µg) of protein. |
| D2 | Integer | Number to distinguish experimental groups. For example, in the case of the current analysis (see text), wt(−) [1] is 1, NPC(−) is 2, and NPC(+) [1] is 3. |
| E2 | Strings | Short name of the experiment. For example, in the case of the current analysis (see text), wt(−), NPC(−), and NPC(+) fall under this category. |

[1] The (+) and (−) symbols indicate whether HPBCD was added or not.

We observed that the amplitude of measurement errors tends to vary depending on *m/z* and range values. TAG Expression generates a file named 'calib_ms_value_plot.html', and scatter diagrams are drawn (Figure 4) in which the x-axis is *m/z* (theoretical) and the y-axis is the deviation between theoretical and observed *m/z* values. In the diagram, the maximum cluster and outliers are considered to be signals from glycans and contaminants, respectively (Figure 4). To validate this approach, we confirm that the maximum cluster includes the signal from added internal standards after selection of the cluster.
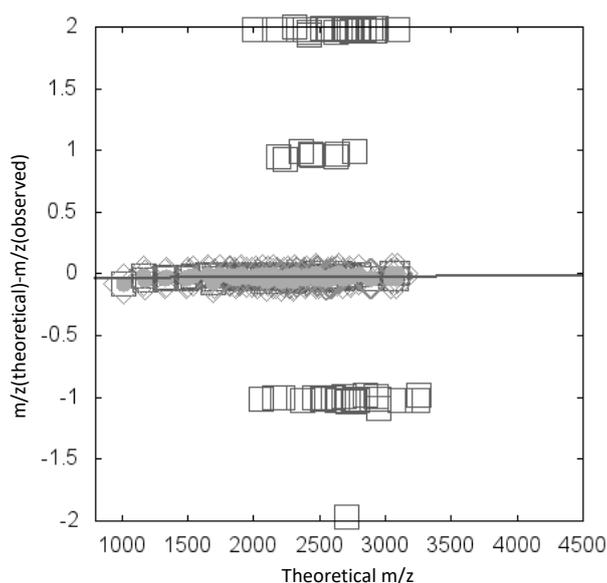


**Figure 4.** Example of a scatter diagram in which the x-axis is *m/z* (theoretical) and the y-axis is the deviation between theoretical and observed *m/z* values. Diamonds indicate the maximum cluster. Data are fitted to a linear function, and line indicates the fit. Gray circles indicate annotated glycan data points within a certain distance from the fitted line.

Following the process performed by TAG Expression, a number of files are generated in a designated folder. These are CSV sheets of the results, as well as various images for chart building. An 'out_list.csv' file stores data related to the annotation of glycans, as described in the results and discussion.

Based on previous findings that oligosaccharides within a certain molecular weight range may exhibit similar signal strengths irrespective of their structure [18,19], when examined by an appropriately calibrated MALDI-TOF MS instrument, we employed absolute quantification by comparative analysis of the areas of the MS signals derived from each glycan and a known amount of the internal standard (A2GN1). It should be noted that TAG Expression not only annotates glycan signals, but also outputs a table in which assigned glycans are comparably sorted using quantitative information in a standardized format. Specifically, the average and standard deviation for each glycan concentration are calculated automatically, along with t-test analysis of all possible combinations of experimental groups. TAG Expression also produces 'Exp_list.csv' files and 'each_glycan_quant.html' files that summarize the quantitative expression profiles and plot graphs comparing average glycan expression values with standard deviation between groups, respectively (Figure S1 (*N*-glycans) and Figure S2 (FNGs)). Additionally, script files for gnuplot chart production and Cluster 3.0 cluster analysis are produced.

### 2.2.4. TAG Pathway

When a cellular *N*-glycome is analyzed, some tens to several hundreds of glycans may be detected on MALDI-TOF spectra. If glycan expression variation can be mapped against a biosynthetic pathway, an overview of metabolic changes in the system can be obtained.

In eukaryotes, *N*-Glycan synthesis [2] begins with the synthesis of an oligosaccharide containing 14 monosaccharides on the lipid dolichol phosphate, following which the lipid-linked oligosaccharide is transferred to a specific asparagine residue of the growing polypeptide chain by the action of oligosaccharyltransferase (OST) in the endoplasmic reticulum (ER). Correctly folded glycoproteins transit to the Golgi apparatus to form high-mannose, hybrid, and complex *N*-glycans. In mammalian cells, FNGs are generated by three metabolic pathways: (i) OST-mediated release of FNGs in the ER [20,21], (ii) pyrophosphatase acting on dolichol-linked oligosaccharides [22,23], and (iii) cytoplasmic peptide:*N*-glycanase (PNGase) acting on misfolded glycoproteins via ER-associated degradation (ERAD) [24,25]. In any case, endo-β-*N*-acetylglucosaminidase (ENGase) in the cytosol metabolizes FNGs to Gn1 glycans, which have only a single GlcNAc at their reducing termini. These Gn1 glycans are susceptible to the action of a cytosolic α-mannosidase (Man2C1), giving rise to the specific Man5 GlcNAc1 structure. Man5 GlcNAc1 is transported into lysosomes where the oligosaccharides are hydrolyzed into monosaccharides for recycling. The TAG Pathway draws a fused biosynthesis pathway for *N*-glycans and FNGs; then, the variation in glycan expression is mapped onto this fused pathway.

The TAG Pathway generates an html file of glycan biosynthetic pathways. We prepared biosynthetic pathway maps for *N*-glycans and FNGs in CSV format, part of which is shown in Figure 5. The map includes cells in an Excel sheet containing information on glycans (e.g., from cells BS39 to BS42) such as glycan class (in BS39, *N*-glycans on folded proteins), type of glycan (in BS40, HM means high mannose-type glycans), five-digit notation of glycan composition (BS41), and localization of glycans (in BS42, Golgi and later) or relationships between glycans (e.g., 'ra.png' means right arrow image in cell BR38). The image files are located in the 'glycan_img' folder of TAG. The intracellular localization of glycans is expressed using cell color, where yellow, pale blue, gray, and magenta indicate ER, Golgi, cytoplasm, and lysosome, respectively. These files can be easily written or modified by users.

### 2.3. Experimental Glycomic Analysis of N-glycans and FNGs for Npc1 KO CHO Cells

To evaluate the feasibility of TAG, *N*- and FNG-glycomic data from our previously reported CHO cells and mutants lacking a functional *Npc1* gene (*Npc1* KO CHO cells), herein denoted NPC(−), were used as a test case [26]. The *Npc1* gene encodes a protein involved in lipid transport between lysosomes and the ER. The *Npc1* KO CHO cells provide a model for Niemann–Pick disease

type C (NPC) [26]. *Npc1* deficiency causes an accumulation of free cholesterol and glycolipids, and 2-Hydroxypropyl-β-cyclodextrin (HPBCD) can reduce cholesterol accumulation. Our previous study examined the effects of *Npc1* KO and HPBCD treatment on glycome expression. Glycomic analyses were carried out on wild-type CHO cells (denoted as wt(−)), *Npc1* KO CHO cells (denoted as NPC(−), and *Npc1* KO CHO cells treated with HPBCD (denoted as NPC(+)) using TAG, where the (+) and (−) symbols indicate whether HPBCD was added or not. The results were compared with those from our previous study using a macro and Microsoft Excel.

|  | BR | BS | BT | BU | BV | BW |
|---|---|---|---|---|---|---|
| 38 | ./glyco_img/ra.png | | ./glyco_img/ra.png | | ./glyco_img/ra.png | |
| 39 | | N-glycan_folded | | N-glycan_folded | | N-glycan_folded |
| 40 | | HM_GN2 | | HM_GN2 | | HM_GN2 |
| 41 | | 50000C | | 40000C | | 30000C |
| 42 | | Golgi and later | | Golgi and later | | Golgi and later |
| 43 | | | | ./glyco_img/da.png | | |
| 44 | | | | | | |

**Figure 5.** Part of the biosynthesis pathway map in comma-delimited files (CSV) format.

### 2.3.1. Extraction of Cellular Glycoproteins and Free Oligosaccharides

For the *N*-glycan and FNG analyses, glycoproteins and FNGs were extracted as previously described [17]. Approximately $1 \times 10^6$ cells were suspended in Tris-acetate buffer containing 2% sodium dodecyl sulfate and homogenized using an Ultrasonic Homogenizer (Taitec, Saitama, Japan). Reductive alkylation of the cellular proteins was performed, followed by the precipitation of proteins in the presence of a four-fold volume of ice-cold ethanol. The precipitates including proteins were dried, dissolved in ammonium bicarbonate, and digested with trypsin. Finally, *N*-glycans were prepared by PNGase F difestion. The supernatants containing FNG were completely dried and dissolved in deionized water. Then, the samples were directly subjected to the glycoblotting procedure. Detailed procedures and are provided elsewhere [26].

### 2.3.2. Glycoblotting Procedure and MALDI-TOF/TOF MS Analysis

*N*-glycans and FNGs were subjected to the glycoblotting procedure. Detailed procedures and materials are provided elsewhere [26,27]. Purified *N*-glycans, FNGs, were combined with 2,5-dihydrobenzoic acid (10 mg/mL in30% acetonitrile) and subsequently subjected to MALDI-TOF MS analysis, as previously described [17]. As all MALDI-TOF MS instruments may have a function to export the values of *m/z* and signal strength (e.g., height, area) of observed signals on the spectra as xls files, we chose to use the exported xls files as a source data of TAG.

## 3. Results and Discussion

The usefulness of TAG was explored using our previously reported *Npc1* KO CHO cell data [26].

### 3.1. Glycan Expression Analysis of Npc1 KO CHO Cells

The 'out_list.csv' file is shown in Figure 6. For signals with *m/z* values within the range of error tolerance, information about the composition, five-digit notation, glycan type as defined in Table S1, theoretical *m/z* deviation between *m/z* (theoretical) and *m/z* (observed), area, and determined quantity (pmol/100 µg protein) are shown. Values of glycans not detected in the analysis are shown as zero.

As shown in Figure 7, the 'exp_list.csv' file summarizes the quantitative expression profiles in terms of the quantity of each glycan (pmol/100 µg protein), the average, standard deviation (S.D.), and coefficient of variation (C.V.) of defined groups, and t-values and p-values (Student's t-test) between all possible combinations of series. In addition to Student's t-tests, hierarchical clustering was applied in order to group glycans based on their expression levels among samples. This was performed using the open source tool Cluster 3.0 to create clusters of data, and these clusters were represented graphically using the open source tool Java TreeView [14].

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | glycan class:N | | | | | | | | |
| 2 | wt-3.pnggmuller/work/glyco2/tag_v5.2_release/N_list_auto_877.csv | | | | | | | | |
| 3 | glycan formula | alias | Hex | HexNAc | Fuc | NeuAc | NeuGc | type | theoretical m/z |
| 4 | IS A2-GN | | | | | | | | 2477.963 |
| 5 | (Hex)1 (HexNAc)2 | 12000 | 1 | 2 | 0 | 0 | 0 | PM | 1016.4422 |
| 6 | (Hex)2 (HexNAc)2 | 22000 | 2 | 2 | 0 | 0 | 0 | PM | 1178.495 |
| 7 | (Hex)3 (HexNAc)2 | 32000 | 3 | 2 | 0 | 0 | 0 | PM | 1340.5478 |
| 8 | (Hex)4 (HexNAc)2 | 42000 | 4 | 2 | 0 | 0 | 0 | PM | 1502.6006 |
| 9 | (Hex)2 + (Man)3(GlcNAc)2 | 20000C | 2 | 0 | 0 | 0 | 0 | HM | 1664.6534 |
| 10 | (Hex)1 (HexNAc)2 (Fuc)1 | 12100 | 1 | 2 | 1 | 0 | 0 | PM_F | 1162.5001 |
| 11 | (Hex)2 (HexNAc)2 (Fuc)1 | 22100 | 2 | 2 | 1 | 0 | 0 | PM_F | 1324.5529 |
| 12 | (Hex)3 (HexNAc)2 (Fuc)1 | 32100 | 3 | 2 | 1 | 0 | 0 | PM_F | 1486.6057 |

glycan structure

five-digit notation

the number of monosaccharide residue

type of glycan in table S1/S2

theoretical m/z

| J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|
| MS | | | | | | | |
| wt- m/z | diff | area | quant(pmol) | wt-2 m/z | diff | area | quant(pmol) |
| 2477.93 | -0.0303 | 22113.89 | 5.0000 | 2477.97 | 0.0069 | 15451.19 | 5.0000 |
| 0.00 | 0.0000 | 0.00 | 0.0000 | 0.00 | 0.0000 | 0.00 | 0.0000 |
| 1178.45 | -0.0435 | 1387.11 | 1.8399 | 1178.46 | -0.0302 | 895.51 | 2.2934 |
| 1340.52 | -0.0251 | 1820.17 | 2.4143 | 1340.54 | -0.0121 | 1087.16 | 2.7843 |
| 1502.58 | -0.0179 | 1325.65 | 1.7583 | 1502.60 | -0.0033 | 798.54 | 2.0451 |
| 1664.65 | -0.0032 | 6892.52 | 9.1423 | 1664.67 | 0.0121 | 3939.24 | 10.0886 |
| 1162.42 | -0.0771 | 108.89 | 0.1444 | 1162.46 | -0.0359 | 106.23 | 0.2721 |
| 1324.53 | -0.0273 | 3971.36 | 5.2676 | 1324.54 | -0.0087 | 2877.42 | 7.3692 |
| 1486.59 | -0.0169 | 734.42 | 0.9741 | 1486.58 | -0.0209 | 405.66 | 1.0389 |

anotation information (see text)　　anotation information (see text)　…

**Figure 6.** Part of the 'out_list.csv' output file for glycan annotation by TAG Expression.

The number of unique glycans identified by TAG in the present work was compared with that reported in our previous study [26]. Since TAG annotates all glycans with the same *m/z* values but different structures, we collected only *m/z* values to represent the unique glycans annotated. *N*-glycans and FNGs detected and quantified by TAG Expression are summarized in Tables S3 and S4. A total of 61 and 36 glycans with unique *m/z* were assigned as *N*-glycans and FNGs, respectively, by TAG Expression. These numbers are greater than those previously reported (58 *N*-glycans and 30 FNGs) based on the manual picking of glycan signals [26]. It should be noted that all *N*-glycans and FNGs detected in our previous work [26] were successfully detected in the present study. Absolute quantities of glycans measured in both the previous and current study are in good agreement, with correlation coefficients ($r^2$) > 0.999. In addition, we generated glycans containing 0–2 acetyl groups for *N*-glycans and FNGs using the TAG List, resulting in two acetylated *N*-glycans and an acetylated FNG that was not detected in our previous study. One *N*-glycan and five FNG signals were also annotated in the TAG results that were similarly not detected in our previous study. This indicates the importance of the semi-comprehensive glycan list feature of TAG List. Thus, the modifiable and expandable nature of the TAG List function provides great advantages for correctly and comprehensively detecting glycan signals without the need for detailed knowledge on glycan structure.

**(a)**



**(b)**
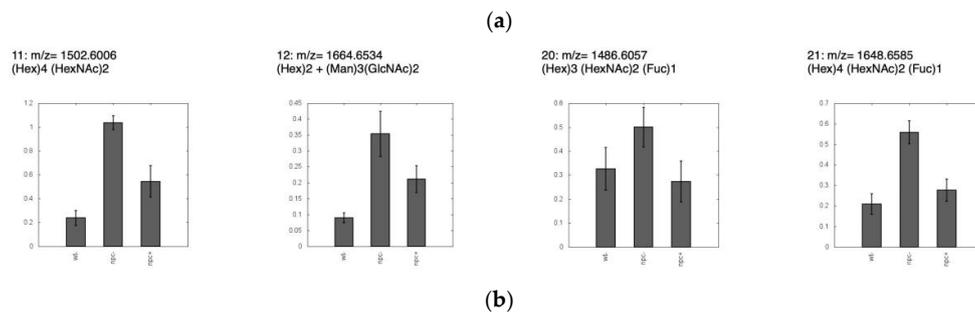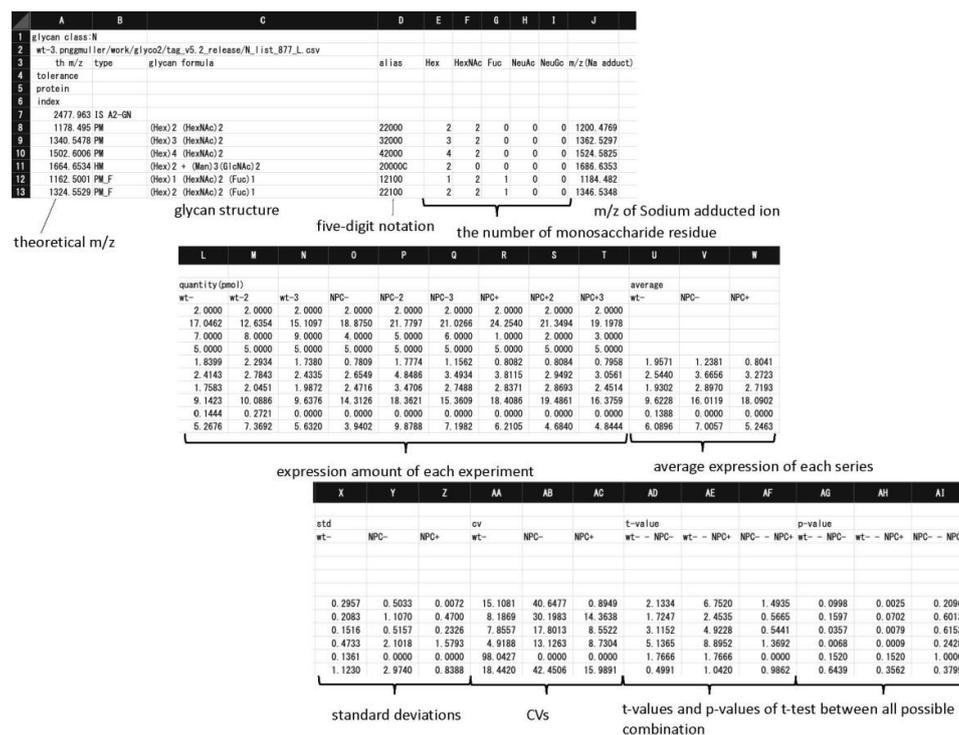
**Figure 7.** Examples showing the output files (**a**) expt_list.csv that summarize quantitative data and statistics, and (**b**) each_glycan_quant.html that graphs quantitative data for different groups. In each chart, *m/z* values and glycan composition are in the header. The vertical axis shows expression quantities (pmol/100 µg protein). Error bars represent standard deviation. Abscissa axis indicate wt(−), NPC(−), and NPC(+), respectively.

## 3.2. Glycan Expression Variation Mapped to Biosynthetic Pathways

The biosynthetic pathway used for the analysis of *N*-glycans and FNGs is shown in Figure 8. Each cell (small rectangle) corresponds to an individual glycan. *N*-glycans are shown to be attached to mature, immature, or misfolded proteins. The background color of the cell indicates the localization of glycans (e.g., cytoplasm, ER, Golgi, lysosome). The relevance of the colors is described above. The TAG Pathway creates two types of biosynthetic pathway maps; one shows variation in glycan expression as bar charts, and the other generates numerical values such as expression differences and ratios. In the latter case, significant increases and decreases ($p < 0.05$) are highlighted on the pathway in red and blue, respectively. Figure 9 shows the data in each cell, which include glycan structure, five-digit notation, and a bar chart showing glycan expression (Figure 9 left). Rather than a bar chart, we can visualize expression variability as differences, ratios, or *p*-values (Figure 9 right).
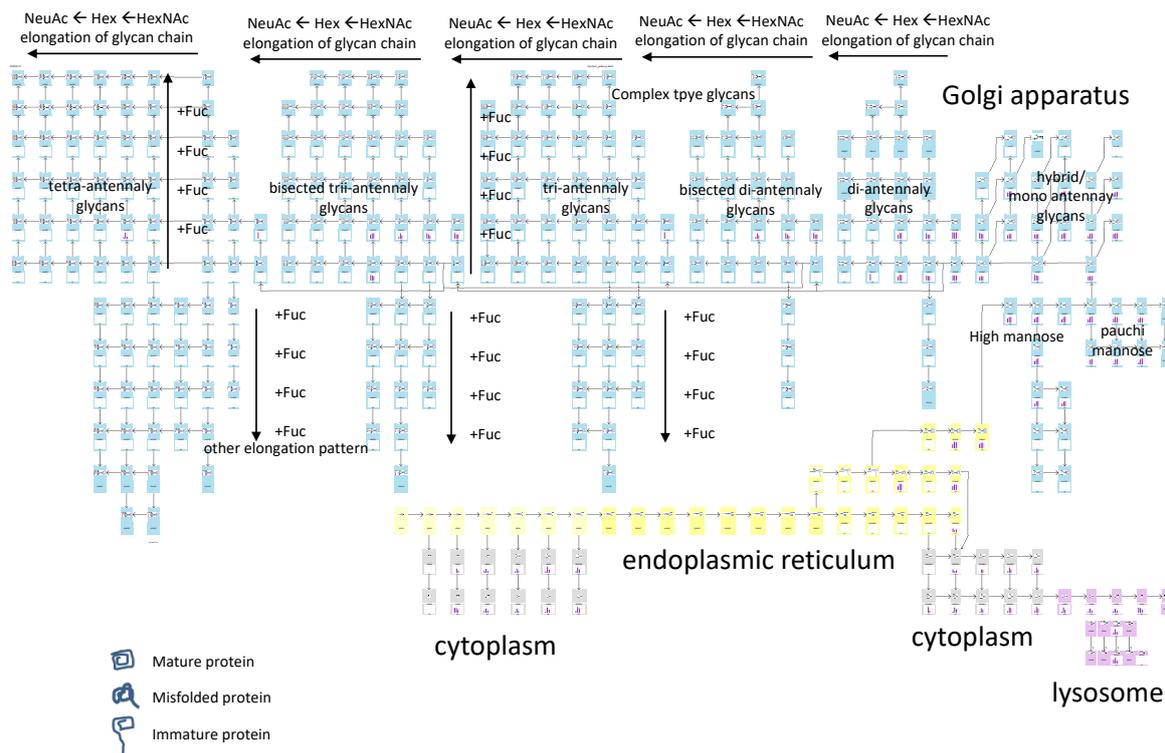
**Figure 8.** Overview of the structure for the biosynthetic pathway model for *N*-glycans and FNGs.

Several pioneering studies reported various glycomic alterations in NPC involving various glycosphingolipids (GSL) (glucosylceramides (GlcCer), lactosylceramide (LacCer), GM2, GM3, and asiao-GM2) [28,29]. In addition, the disease-specific accumulation of various sialylated glycoconjugates within endocytic compartments of *Npc1*-null and *Npc2*-deficient fibroblasts is caused by impaired recycling as opposed to altered fusion of vesicles. The treatment of either *Npc1*-null or *Npc2*-deficient cells with cyclodextrin was effective in reducing cholesterol storage, as well as the endocytic accumulation of sialylated glycoconjugates [30], although structure-intensive analysis was not performed in this study. Our previous study using *Npc1* KO CHO cells identified a number of glycomic alterations, including an increased expression of LacCer, GM1, GM2, GD1, various neolacto-series glycosphingolipids, and sialyl-T (*O*-glycan), as well as various *N*-glycans, which were typically both fucosylated and sialylated. We also observed significant increases in the total amounts of free *N*-glycans (FNGs), especially in the unique complex- and hybrid-type FNGs. The treatment of *Npc1* KO CHO cells with 2-hydroxypropyl-β-cyclodextrin (HPBCD) did not affect the glycomic alterations observed in the GSL-, *N*-, and *O*-glycans of *Npc1* KO CHO cells. However, HPBCD treatment corrected the glycomic alterations observed in FNGs to levels observed in wild-type cells [26].
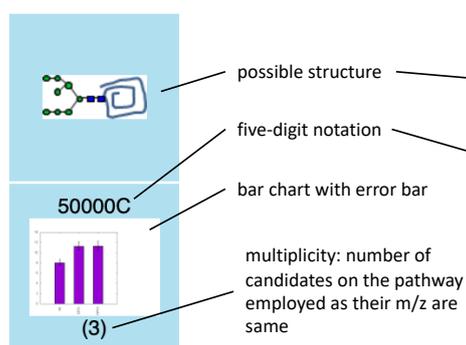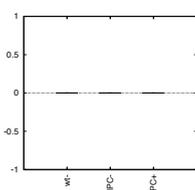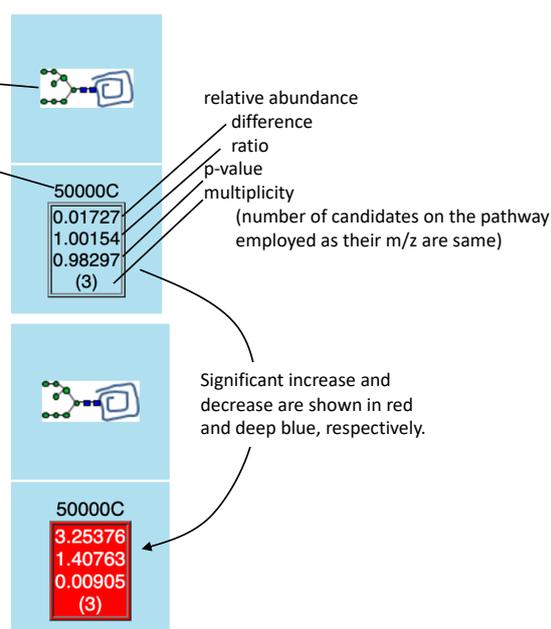
**Figure 9.** Two different methods for visualizing alterations; bar chart (**left**) and mean relative abundance and *p*-value (**right**). In the bar chart, the three bars show wt(−), NPC(−), and NPC(+) expression levels, respectively.

Glycomic alteration among wt(−), NPC(−), and NPC(+) series can be shown graphically on the map as exemplified in Figure 10. The whole graphic can be seen in Figure S3–S5. In Figure 10a, representative glycomic alterations of HM and PM glycans were shown with bar charts (upper) and mean relative abundance and *p*-value (lower). It is shown that *Npc1* knockout generally increases the glycan expression levels and that HPBCD addition has little effect on glycan expression levels. As shown in Figure 10b–d, fucosylated and sialylated complex type *N*-glycans tend to increase by *Npc1* knockout, some of which increase significantly, and the treatment of *Npc1* KO CHO cells with HPBCD did not often affect the glycomic alterations of *Npc1* KO CHO cells. In contrast, the glycomic alterations observed in FNGs (typically complex type FNGs) were corrected by the treatment of HBPCD to levels observed in wild-type cells. These observations agree well with those reported in our previous publication [26].

These functions will help researchers readily determine the specificity of alterations. Another feature of the TAG pathway is that the map template itself is written in CSV format, and it is modifiable by the user.

(**a**)



(**b**)
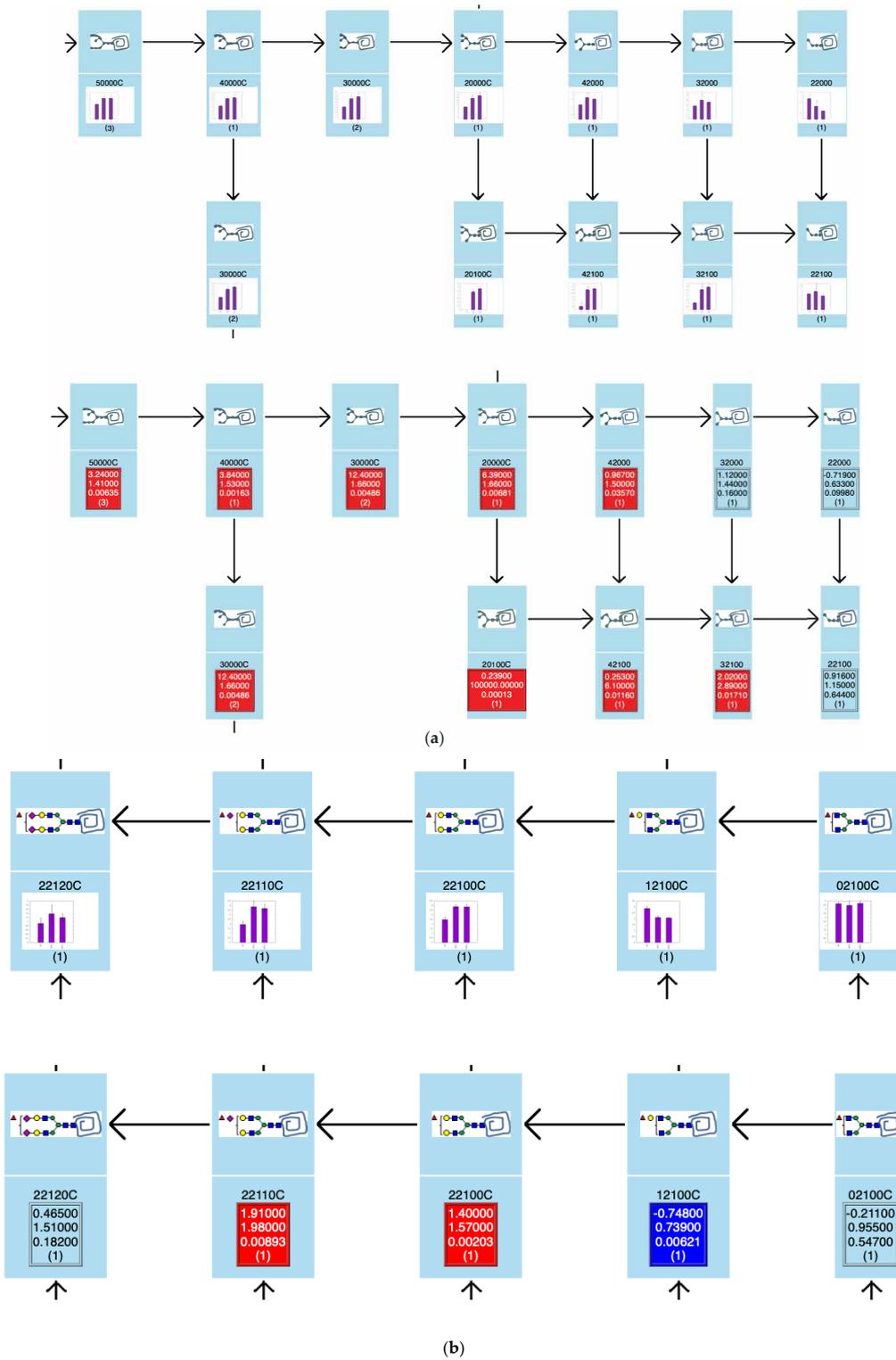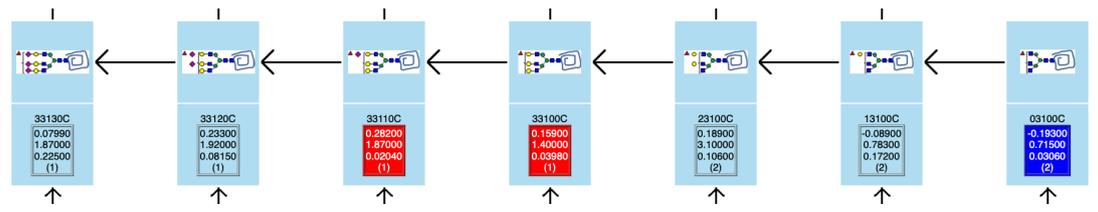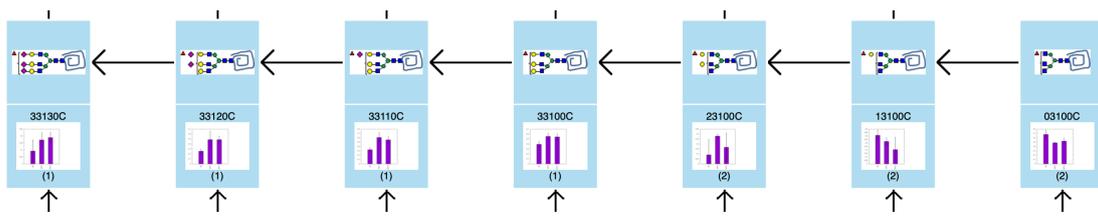
**Figure 10.** *Cont.*

(c)
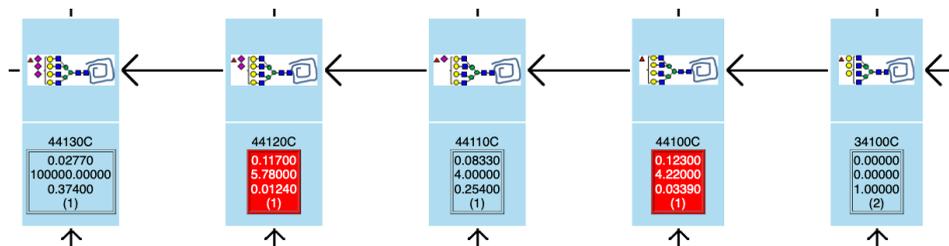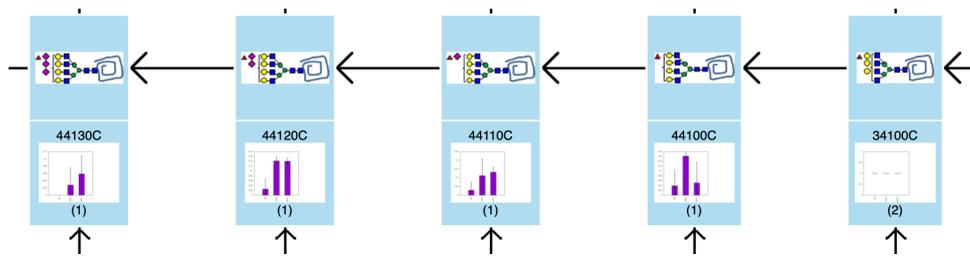


(d)

**Figure 10.** *Cont.*

(**e**)

**Figure 10.** *Cont.*

(f)



(g)
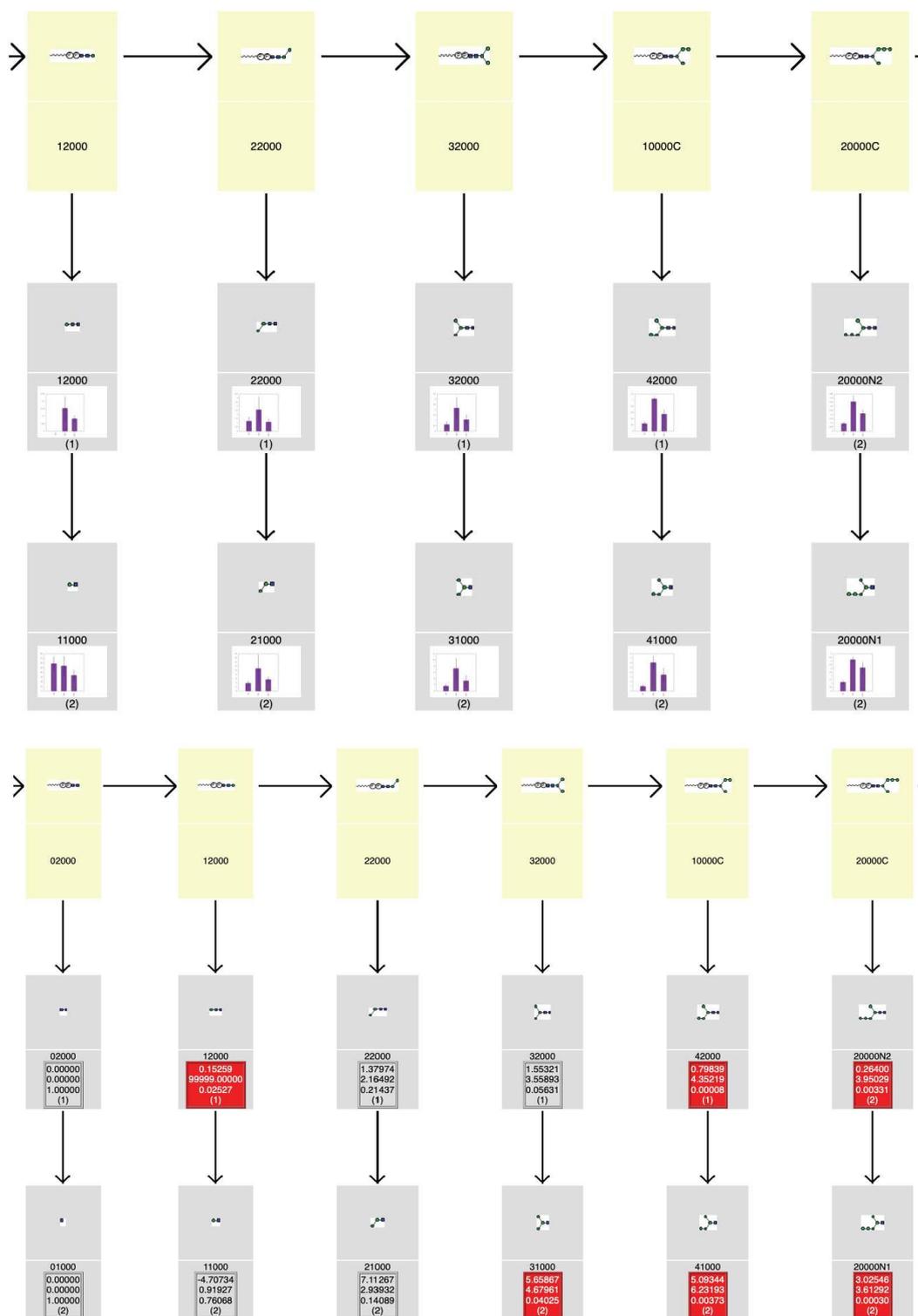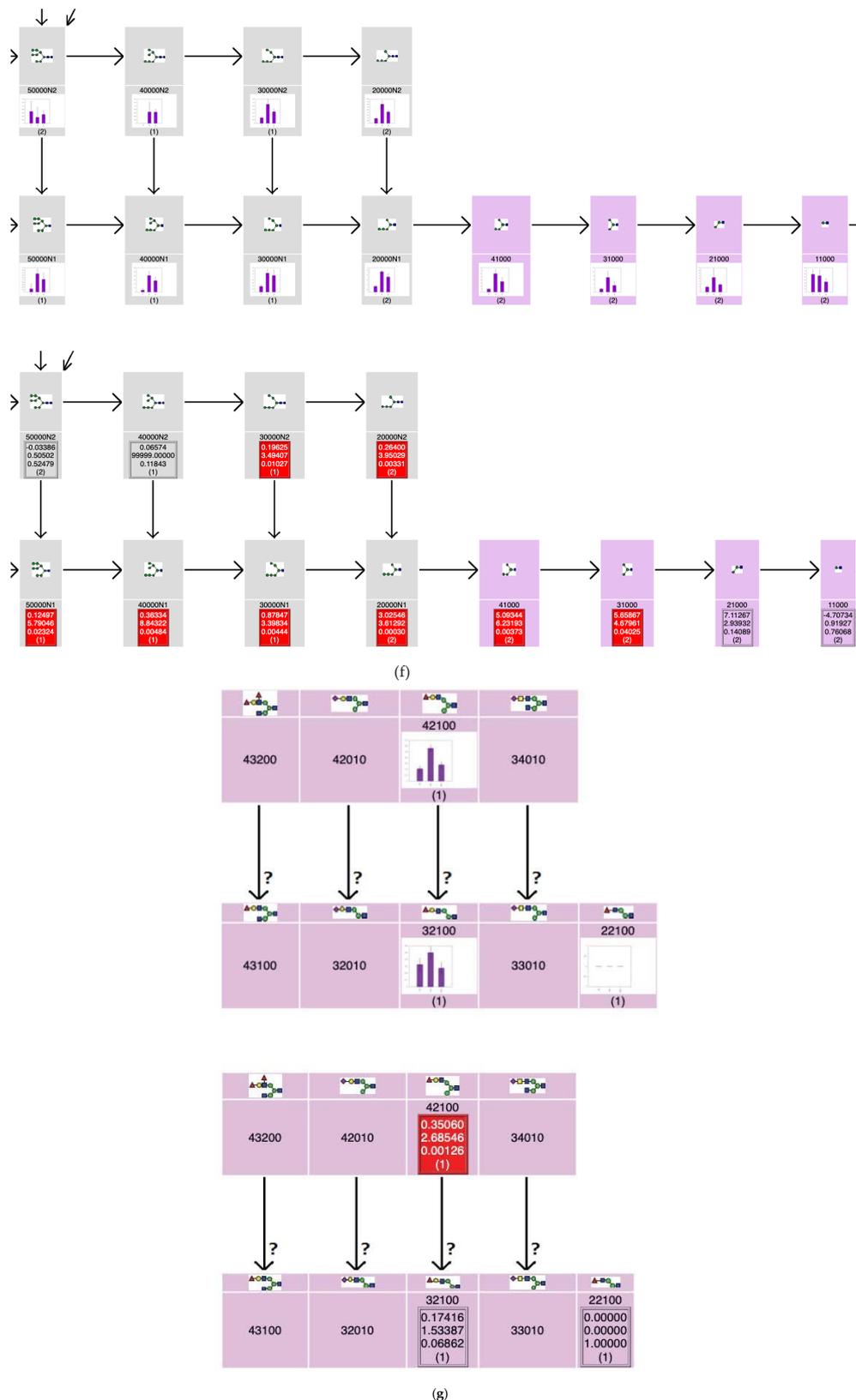
**Figure 10.** Representative alterations observed among wt(−), NPC(−), and NPC(+) series using bar charts (upper in each panel) and with mean relative abundance and *p*-value between wt(−) and NPC(−) (lower) for *N*-glycans and FNGs. (**a**) Expression levels of high (oligo) mannose type (HM) and pauci mannose type (PM) *N*-glycans. (**b**) Complex biantennary *N*-glycans. (**c**) Complex triantennary *N*-glycans. (**d**) Complex tetraantennary *N*-glycans. (**e**) FNGs possibly derived through The action of pyrophosphatase. (**f**) FNGs possibly derived through The actions of oligosaccharyltransferase (OST) and/or peptide:*N*-glycanase (PNGase). (**g**) for a part of complex FNGs.

## 4. Conclusions

In this study, we developed Toolbox Accelerating Glycomics (TAG) software that provides a modifiable and expandable glycan list, automatically annotates and quantifies the glycan signals on MALDI-TOF MS spectra, quantifies and sorts the results, and visualizes variation in glycan expression based on glycan biosynthetic pathways. Glycomic analysis without such tools requires much time and labor, whereas TAG analysis takes just a few minutes. The results can be viewed by web browsers such as Firefox, Chrome, Edge, and Safari, which reduces instrument dependency. To support the discrimination of structural isomers, TAG can provide a list of possible structure candidates with the same composition, which assists the design of further experiments such as MS/MS and exoglycosidase digestion. Considering that glycomic alterations may be mutually related to one another through biosynthetic pathways, mapping glycomic alterations against biosynthetic pathways can help determine glycan structures. There are a variety of methods to label the reducing termini of glycans and to derivatize sialic acid residue(s). Examples of the former include 2-aminobenzamide [31] and RapiFluor-MS [32]. As examples of the latter, we recently reported a sialic acid linkage-specific alkylamidation method by lactone ring-opening aminolysis (aminolysis-SALSA), for the discrimination between $\alpha$2,3- and $\alpha$2,6-linked sialylated glycan isomers by mass spectrometry [33,34]. The TAG List can be readily applied to such labels and derivatizations with minor modification. Although it is desirable to perform additional experiments to distinguish structural isomers, we think it is worth attempting to predict the correct structural isomer by accumulating glycomic data and utilizing machine learning techniques. TAG is a flexible program, and glycan lists and pathway maps can be generated by the researchers themselves. Therefore, we believe that TAG could accelerate glycomic analysis and assist the development of similar software. The TAG programs will be available to all researchers. It will be released via an open source repository service such as GitHub. In addition, we also plan to expand the TAG procedure for the analysis of other glycans, such as glycans of glycosphingolipids, *O*-glycans, and glycosaminoglycans.

**Supplementary Materials:** The following are available online at http://www.mdpi.com/2218-273X/10/10/1383/s1: Figure S1, Bar charts showing the abundance of *N*-glycans for wt(−), NPC(−), and NPC(+) series. Figure S2, Bar charts showing the abundance of FNGs for wt(−), NPC(−), and NPC(+) series. Figure S3, The entire biosynthetic pathway of *N*-glycans and FNGs mapped to bar charts. Figure S4, The entire biosynthetic pathway of *N*-glycans and FNGs mapped to differences in expression between wt(−) and NPC(−) series. Figure S5, The entire biosynthetic pathway of *N*-glycans and FNGs mapped to differences in expression between NPC(−) and NPC(+) series. Table S1, Classification of *N*-glycans. Table S2, Classification of free *N*-glycans (FNGs). Table S3, Annotated *N*-glycans with TAG expression. Annotated glycan compositions reported previously [26], determined using a macro and Microsoft Excel, are also included in this table. Table S4, Annotated FNGs with TAG expression. Annotated results reported previously [26], determined using a macro and Microsoft Excel, are also included in this table.

**Author Contributions:** Conceptualization, N.M., Y.S. and J.-I.F.; investigation, N.M., H.H., I.Y., K.O., J.-I.F. and Y.S.; software, N.M.; validation, H.H., I.Y. and K.O.; formal analysis, I.Y. and K.O.; writing the original draft, N.M., Y.S. and J.-I.F.; reviewing and editing, N.M., H.H., J.-I.F. and Y.S.; funding acquisition, J.-I.F. and Y.S. All authors have read and agreed to the published version of the manuscript.

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

## References

1. Varki, A. Biological roles of glycans. *Glycobilogy* **2017**, *27*, 3–49. [CrossRef]
2. Varki, A. (Ed.) *Essentials of Glycobiology*, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2017.

3.  Costa-Silva, J.; Domingues, D.; Lopes, F.M. RNA-seq differential expression analysis: An extended review and software tool. *PLoS ONE* **2017**, *12*, e0190152. [CrossRef] [PubMed]

4.  Tyanova, S.; Temu, T.; Cox, J. The MaxQuant Computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **2016**, *11*, 2301–2319. [CrossRef] [PubMed]

5.  Perkins, D.N.; Pappin, D.J.; Creasy, D.M.; Cottrell, J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20*, 3551–3567. [CrossRef]

6.  Craig, R.; Beavis, R.C. TANDEM: Matching proteins with tandem mass spectra. *Bioinformatics* **2004**, *20*, 1466–1467. [CrossRef]

7.  Ceroni, A.; Maass, K.; Ceyer, H.; Geyer, R.; Dell, A.; Haslam, S.M. GlycoWorkbench: A tool for the compueter-assisted annotation of mass spector of glycans. *J. Proteome Res.* **2008**, *7*, 1650–1659. [CrossRef]

8.  Svarovsky, S.A.; Joshi, L. Cancer glycan biomarkers and their detection—Past, present and future. *Anal. Methods* **2014**, *6*, 3918–3936. [CrossRef]

9.  Kanehisa, M.; Goto, S.; Furumichi, M.; Tanabe, M.; Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **2010**, *38*, D355–D360. [CrossRef]

10. Roche Biochemical Pathways. Available online: https://www.roche.com/sustainability/philanthropy/science_education/pathways.htm (accessed on 27 September 2020).

11. Croft, D.; Mundo, A.F.; Haw, R.; Milacic, M.; Weiser, J.; Wu, G.; Caudy, M.; Garapati, P.; Gillespie, M.; Kamdar, M.R.; et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* **2014**, *42*, D472–D477. [CrossRef]

12. Eisen, M.B.; Spellman, P.T.; Brown, P.O.; Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14863–14868. [CrossRef]

13. De Hoon, M.J.L.; Imoto, S.; Nolan, J.; Miyano, S. Open source clustering software. *Bioinformatics* **2004**, *20*, 1453–1454. [CrossRef] [PubMed]

14. Saldanha, A.J. Java Treeview—Extensible visualization of microarray data. *Bioinformatics* **2004**, *20*, 3246–3248. [CrossRef] [PubMed]

15. Gnuplot Homepage. Available online: http://www.gnuplot.info (accessed on 28 September 2020).

16. Gawk—GNU Project—Free Software Foundation (FSF). Available online: https://www.gnu.org/software/gawk/ (accessed on 27 September 2020).

17. Fujitani, N.; Furukawa, J.-I.; Araki, K.; Fujioka, T.; Takegawa, Y.; Piao, J.; Nishoka, T.; Tamura, T.; Nikaido, T.; Ito, M.; et al. Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. *Porc. Nat. Am. Soc.* **2013**, *110*, 2105–2110. [CrossRef]

18. Furukawa, J.-I.; Sakei, S.; Yokota, I.; Okada, K.; Hanamatsu, H.; Kobayashi, T.; Yoshida, Y.; Higashino, K.; Tamura, T.; Igarashi, Y.; et al. Quantitative GSL-glycome analysis of human whole serum based on an EGCase digestion and glycoblotting method. *J. Lipid Res.* **2015**, *56*, 2399–2407. [CrossRef]

19. Naven, T.J.; Harvey, D.J. Effect of structure on the signal strength of oligosaccharides in matrix-assisted laser desorption/ionization mass spectrometry on time-of-flight and magnetic sector instruments. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1361–1366. [CrossRef]

20. Harada, Y.; Buser, R.; Ngwa, E.M.; Hirayama, H.; Aebi, M.; Suzuki, T. Eukaryotic oligosaccharyltransferase generates free oligosaccharides during *N*-glycosylation. *J. Biol. Chem.* **2013**, *288*, 32673–32684. [CrossRef] [PubMed]

21. Harada, Y.; Masahara-Negishi, Y.; Suzuki, T. Cytosolic-free oligosaccharides are predominantly generated by the degradation of dolichol-linked oligosaccharides in mammalian cells. *Glycobiology* **2015**, *25*, 1196–1205. [CrossRef]

22. Harada, Y.; Nakajima, K.; Masahara-Negishi, Y.; Freeze, H.H.; Angata, T.; Taniguchi, N.; Suzuki, T. Metabolically programmed quality control system for dolichol-linked oligosaccharides. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19366–19371. [CrossRef]

23. Harada, Y.; Huang, C.; Yamaki, S.; Dohmae, N.; Suzuki, T. Non-lysosomal degradation of phophorylated oligosaccharides is initiated by the cytosolic endo-β-*N*-acetylglucosaminidase. *J. Biol. Chem.* **2016**, *291*, 8048–8058. [CrossRef]

24. Suzuki, T.; Huang, C.; Fujihira, H. The cytoplasmic peptide:*N*-glycanase (NGLY1); structure, expression and cellular functions. *Gene* **2016**, *577*, 1–7. [CrossRef]

25. Suzuki, T. The cytoplasmic peptide:*N*-glycanase (PNGase)—Basic science encounters a human genetic disorder. *J. Biochem.* **2015**, *157*, 23–34. [CrossRef] [PubMed]

26. Furukawa, J.-I.; Soga, M.; Okada, K.; Yokota, I.; Piao, J.; Irie, T.; Era, T.; Shinohara, Y. Impact of the Niemann-Pick c1 gene mutation on the total cellular glycomics of CHO cells. *J. Proteome Res.* **2017**, *16*, 2802–2810. [CrossRef] [PubMed]

27. Furukawa, J.-I.; Shinohara, Y.; Kuramoto, H.; Miura, Y.; Shimaoka, H.; Kurogochi, M.; Nakano, M.; Nishimura, S. Comprehensive approach to structural and functional glycomics based on chemoselective glycoblotting and sequential tag conversion. *Anal. Chem.* **2008**, *80*, 1094–1101. [CrossRef] [PubMed]

28. Fan, M.; Sidhu, R.; Fujiwara, H.; Tortelli, B.; Zhang, J.; Davidson, C.; Walkley, S.U.; Bagel, J.H.; Vite, C.; Yanjanin, N.M.; et al. Identification of Niemann-Pick C1 disease biomarkers through sphingolipid profiling. *J. Lipid Res.* **2013**, *54*, 2800–2814. [CrossRef]

29. Hashimoto, N.; Matsumoto, I.; Takahashi, H.; Ashikawa, H.; Nakamura, H.; Murayama, T. Cholesterol-dependent increases in glucosylceramide synthase activity in Niemann-Pick disease type C model cells: Abnormal trafficking of endogenously formed ceramide metabolites by inhibition of the enzyme. *Neuropharmacology* **2016**, *110*, 458–469. [CrossRef]

30. Mbua, N.E.; Flanagan-Steet, H.; Johnson, S.; Wolfert, M.A. Abnormal accumulation and recycling of glycoproteins visualized in Niemann-Pick type C cells using the chemical reporter strategy. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 10207–10212. [CrossRef]

31. Biggle, J.C.; Patel, T.P.; Bruce, J.A.; Goulding, P.N.; Charles, S.M.; Parekh, R.B. Nonselective and Efficient Fluorescent Labeling of Glycans Using 2-Amino Benzamide and Anthranilic Acid. *Anal. Bicochem.* **1995**, *230*, 229–238. [CrossRef]

32. Kaser, T.; Pavić, T.; Lauc, G.; Gornik, O. Comparison of 2-Aminobenzamide, Procainamide and *Rapi*Fluor-MS as Derivatizing Agents for High-Throughput HILIC-UPLC-FLR-MS *N*-glycan Analysis. *Front. Chem.* **2018**, *6*, 324. [CrossRef]

33. Nishikaze, T.; Tsumoto, H.; Sekiya, S.; Iwamoto, S.; Miura, Y.; Tanaka, K. Differentiation of Sialyl Linkage Isomers by One-Pot Sialic Acid Derivatization for Mass Spectrometry-Based Glycan Profiling. *Anal. Chem.* **2017**, *89*, 2353–2360. [CrossRef]

34. Hanamatsu, H.; Nishikaze, T.; Miura, N.; Piao, J.; Okada, K.; Sekiya, S.; Iwamoto, S.; Sakamoto, N.; Tanaka, K.; Furukawa, J.-I. Sialic Acid Linkage-Specific Derivatization of GlycoSphingolipid Glycans by Ring-Opening Aminolysis of Lactones. *Anal. Chem.* **2018**, *90*, 13193–13199. [CrossRef]