

Chapter 15

RINGS: A Web Resource of Tools for Analyzing Glycomics Data

Kiyoko F. Aoki-Kinoshita

Abstract This chapter introduces the web resource called RINGS which can be accessed at <http://www.rings.t.soka.ac.jp/> and provides freely available tools for analyzing and data mining glycomics data. These include DrawRINGS, an applet to draw glycan structures and obtain the KCF (KEGG Chemical Function)-formatted representation of the structure; ProfilePSTMM, a tool based on a probabilistic model which can extract patterns within groups of glycan structures; Glycan Miner, a tool for finding the common glycan substructures within a group of glycan structures; MCAW, a multiple alignment tool for glycans; Kernel Tool, a tool to find glycan substructures that are particular to one glycan structure data set compared to another; GPP, glycan pathway predictor for generating *N*-glycan biosynthesis pathways; and utilities for converting between various glycan structure representations. Each tool will be described in its usage and application, along with tips for using the tools most effectively. Moreover, RINGS provides a data management system as well as a feedback system to allow users to store their data on the RINGS server as well as to interact with developers to improve RINGS functionality.

Keywords Glycans • Glycobiology • Glycoinformatics • Data mining • Web resource

15.1 Introduction

Due to the development of semi-high-throughput technologies for generating glycomics data in the glycosciences, such as mass spectrometry (MS), nuclear magnetic resonance (NMR), and glycan arrays, numerous glycomics databases and algorithms have been developed to analyze this data in the past decade. The latter (algorithms), however, have been mainly published in the informatics literature, and few have been implemented as easy-to-use tools for the wet-lab glycomicrobiologist. Therefore, RINGS was developed since 2006 in order to provide a

K.F. Aoki-Kinoshita (✉)

Faculty of Science and Engineering, Soka University, 1-236 Tangi-machi, 192-8577, Hachioji, Tokyo, Japan

e-mail: kkiyoko@soka.ac.jp

user-friendly interface to use these tools. In the current version of RINGS, the main text representation of glycan structures used is KCF (KEGG Chemical Function Format) (Akune et al. 2010). As new formats are being developed, utilities to translate between KCF and other formats have also been incorporated into RINGS. In the future, however, RINGS tools will be designed to allow any format to be used as input so that the user need not be concerned with the glycan representation to use.

15.2 DrawRINGS

DrawRINGS is an easy-to-use, web-based drawing tool for searching glycans using KCaM (Aoki et al. 2004). Users can also obtain KCF text for glycans by drawing structures on the canvas.

15.2.1 DrawRINGS Input

The DrawRINGS interface is a Java applet that runs in web browsers. Because of security risks, many browsers these days do not permit applets to run by default. However, it is still possible to run them with the appropriate browser settings, which are described in the *Tips* Sect. 15.2.3. The main drawing interface for DrawRINGS is illustrated in Fig. 15.1. There are nine buttons across the top, which

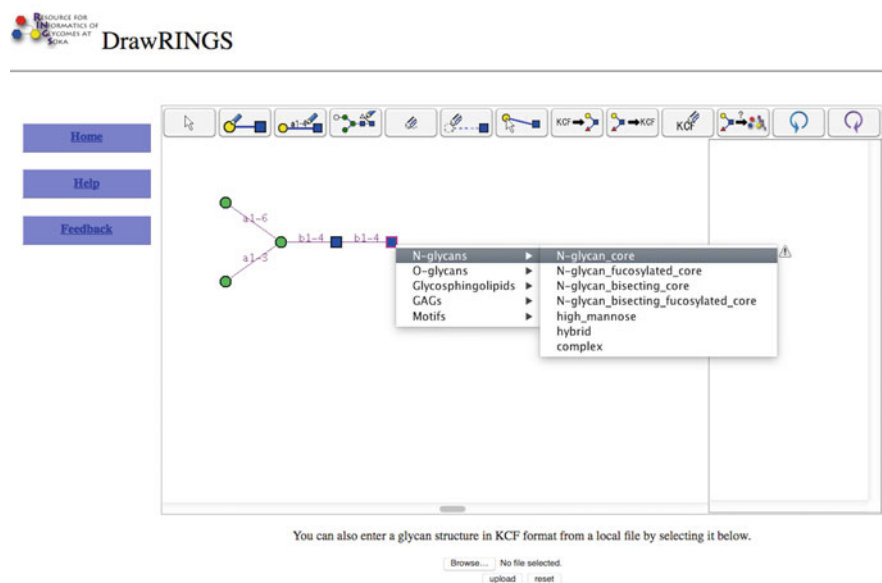









Fig. 15.1 The DrawRINGS input screen, illustrating the menu displayed when using the *core structure* button







can be used for drawing monosaccharides (called nodes), glycosidic bonds (called edges), erasing nodes, making queries, etc. Table 15.1 describes each button and its functionality.

Table 15.1 Each button in DrawRINGS allows users to perform the following tasks

Button	Function	Description
	Select	Select monosaccharides on the canvas. Multiple monosaccharides can be selected at once by dragging the mouse over the monosaccharides. Selected items can be moved around on the canvas
	Draw monosaccharide	A selection menu of monosaccharide names appears when the drawing area is clicked, and the selected monosaccharide will be drawn. Moreover, a pop-up window to input a monosaccharide name appears when "Other" is selected such that names of components not appearing in the list can be entered
	Draw glycosidic bond	A glycosidic bond can be drawn between two monosaccharides by clicking on the corresponding monosaccharides. A selection menu of glycosidic conformations appears when a bond is clicked. Moreover, a pop-up menu to input user-defined conformations appears by selecting "Other"
	Core structure	Core structures include the common <i>N</i> -glycan core structures (including bisecting GlcNAc and core fucosylation), <i>O</i> -glycan structures, glycosaminoglycans (GAGs), and common motifs such as Lewis structures, glycolipids, lactosamine repeats, etc. Any of these basic structures can be drawn at once by selecting this button and clicking on the canvas, where a hierarchical menu of structures will be displayed
	Clear canvas	The entire canvas will be cleared of monosaccharides and bonds
	Delete monosaccharide	Delete a particular monosaccharide and its attached bond(s)
	Move monosaccharide	A selected monosaccharide can be dragged across the canvas

(continued)

Table 15.1 (continued)

Button	Function	Description
	Draw KCF	The structure corresponding to the KCF in the text area will be drawn on the canvas
	Output KCF	The structure being drawn will be output into the text area in KCF format
	Clear KCF	The KCF in the text area will be deleted
	Run query	A pop-up menu of options to use when querying the RINGS databases for the drawn structure will appear. Options to select a score matrix, the database to search (RINGS, GlycomeDB, or both), and similarity score options are available (similarity to sort results by percentage matched, or matched to sort by number of matching components). If a score matrix is selected, the algorithm will use the values from the score matrix to weight the linkages appropriately. Please refer to the text for details
	Undo	The last action performed on the canvas will be undone
	Redo	The last action undone will be re-performed

15.2.1.1 Drawing/Inputting Structures onto the Canvas

To draw on the canvas, first use the buttons for either drawing monosaccharides or drawing structures. Then additional monosaccharides and glycosidic bonds can be added as necessary. Adjustments to the topology of the glycan can be made by using the select button and dragging the necessary components around on the canvas. Selected components can also be copied using Ctrl-C (or Command-C on Mac OS) and pasted using Ctrl-V (or Command-V on Mac OS).

A file containing a glycan in KCF format can also be loaded into DrawRINGS by selecting it using the *Browse...* button located below the canvas. Once selected, the structure can be drawn on the canvas by clicking the *Upload* button.

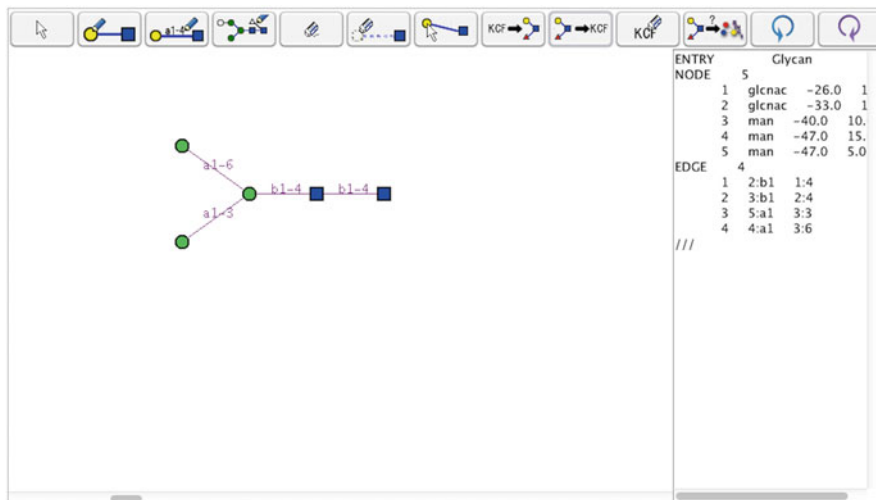


Fig. 15.2 The DrawRINGS screen after KCF has been outputted. The text area containing the KCF data can be edited manually as needed. The updated glycan image can then be obtained by clicking the *Draw KCF* button, which will update the canvas with the glycan represented by the KCF text

15.2.1.2 Obtaining KCF

Once a structure is drawn on the canvas, the KCF-formatted text representing it can be obtained by clicking the *Output KCF* button. Figure 15.2 illustrates how DrawRINGS looks when the KCF text is shown in the text area on the right-hand side of the canvas. This text area can be edited manually to modify the glycan displayed as well. The updated glycan figure can then be obtained by clicking the *Draw KCF* button, which will update the canvas with the glycan represented by the KCF text.

15.2.1.3 Running Queries

The RINGS database stores all glycan structures derived from KEGG GLYCAN (Aoki-Kinoshita and Kanehisa 2015) and GlycomeDB (Ranzinger et al. 2011), in addition to several structures that were manually curated from the literature, up to 2009. The structures in these data can be queried using the *Query* button in DrawRINGS.

The algorithm behind this query tool is KCaM (Aoki et al. 2004), which implements an efficient tree structure matching algorithm. Three options are provided by this query tool: (1) score matrix, (2) database, and (3) score type. A glycan score matrix can be analogized to BLOSUM or PAM for protein sequence alignment, where particular glycosidic linkages and monosaccharides that should be considered

similar to one another, in particular within a certain glycan class, can be scored more highly to obtain higher-scoring glycans within the same class. Predefined score matrices are available, including *N*-glycans, *O*-glycans, and sphingolipids. A fourth matrix, called *Link_similarity*, is a hand-generated score matrix based on expert knowledge of glycosidic linkages and monosaccharides that may be substituted more frequently with other glycosidic linkages and monosaccharides, respectively. For example, β 1-3 may be considered more similar to β 1-4 than α 1-2.

For the target database, users may choose between querying either the data RINGS, which includes KEGG GLYCAN and manually curated structures, or GlycomeDB, or both of these databases. Because the latter contains almost 40,000 structures, it takes a few minutes longer to query it.

Finally, the score type allows users to specify the scoring system to use, and the results are always displayed in descending order of score. The *Similarity* option gives a maximum score of 100, whereas the *Matched* option sums the score of matching monosaccharides and linkages, where each matching glycosidic linkage (and its monosaccharide) receives a score of 100. In fact, the *Similarity* option actually divides the *Matched* score by the larger of the two structures. Because each glycosidic linkage can receive a maximum score of 100, the maximum possible score would be 100 times the maximum number of linkages. For example, if a glycan having nine glycosidic linkages is compared against one of four, the maximum possible score is 100 times the larger of the structures (in this case, $100 \times 9 = 900$). Thus, if the number of matching linkages (and their monosaccharides) is, say, five, then the *Similarity* score between the two structures would be 56% (= $500 \div 900$).

15.2.2 DrawRINGS Query Output

As a result of running a query, a new browser window listing the matching results in order of score will be displayed, as in Fig. 15.3. This figure is a screenshot of a query using the structure drawn in Fig. 15.2 and using the *Link_Similarity* score matrix on the GlycomeDB data. The results list the structures in order of similarity, where it is apparent that the most similar structures are displayed at the top. Each result is linked to the Glycan Entry page, shown in Fig. 15.4, which displays the various text representations of the structure. This page also provides a link to the KEGG GLYCAN database and reaction information about the glycan if available.

15.2.3 DrawRINGS Tips

The following are some tips regarding using DrawRINGS:

- Java needs to be enabled in the browser with appropriate security settings (usually at low levels).



DrawRINGS Result

NEXT 1 2 3 4 5 1 / 106 Go Text file

No.	Score	Glycan id	Reaction id	Structure ?
1	100.0	G00311		
2	100.0	G05952		
3	93.8	G06968		
4	87.6	G06965		

Fig. 15.3 The DrawRINGS result screen using the structure drawn in Fig. 15.2 and using the Link_Similarity score matrix on the GlycomeDB data. The results list the structures in order of similarity, where it is apparent that the most similar structures are displayed at the top. Each result is linked to the Glycan Entry page of RINGS (Fig. 15.4), which displays detailed information about the structure



Glycan Entry

Entry	G06928
Interaction ID	
Class	N-Glycan
KCF	<pre> ENTRY G06928 Glycan COMPOSITION (GlcNAc)2 (Man)3 MASS 910.8 CLASS Glycoprotein; N-Glycan DBLINKS CCSD: 32909 GlycomeDB: 21902 JCGGDB: JCGG-STR012821 NODE 5 1 GlcNAc 0 0 2 GlcNAc -10 0 3 Man -20 0 4 Man -30 5 5 Man -30 -5 EDGE 4 1 2:b1 1:4 2 3:a1 2:4 3 4:b1 3:6 4 5:b1 3:3 /// </pre>
LINUCS Code	<chem>[][D-GlcNAc]{((4+1))[b-D-GlcNAc]{((4+1))[a-D-Man]{((3+1))[b-D-Man]{((6+1))[b-D-Man]{}}}}}</chem>
GlycoCT	<pre> RES 1b:x-dglc-HEX-1:5 2s:n-acetyl 3b:b-dglc-HEX-1:5 4s:n-acetyl 5b:a-dman-HEX-1:5 6b:b-dman-HEX-1:5 7b:b-dman-HEX-1:5 LIN 1:1d(2+1)2n 2:1o(4+1)3d 3:3d(2+1)4n 4:3o(4+1)5d 5:5o(3+1)6d 6:5o(6+1)7d </pre>
Structure	

Fig. 15.4 A Glycan Entry page for the fourth matching structure in the DrawRINGS query results shown in Fig. 15.3. The various text representations of this structure are listed, along with a link to the original KEGG GLYCAN database

- Google Chrome may not work, but DrawRINGS should work on Safari, Firefox, and Internet Explorer.
- Set the URLs <http://www.rings.t.soka.ac.jp> and <http://rings.t.soka.ac.jp> in the list of allowed sites in Java settings.
- At the time of this writing, selecting the text area and using Ctrl-C (or Command-C on Macs) to copy the KCF are disabled on many browsers. Thus, in order to download the KCF text, it is recommended that DrawRINGS be used as a registered user having logged in and then a query be made on the RINGS database, in order to store the KCF in the user's data library. The User Management system is described later in Sect. 15.9 and instructions on running queries are described in the next subsection.

15.3 ProfilePSTMM

The ProfilePSTMM Tool was built based on the probabilistic model proposed called PSTMM (probabilistic sibling-dependent tree Markov model) (Ueda et al. 2004), which was then modified such that profiles could be directly output from the model (Aoki-Kinoshita et al. 2006). In simple terms, this model *learns* patterns from the input data in a process called *training*. Once a model is *trained* on a particular data set, it can essentially be used to *test* other glycans to determine how well it matches the pattern that has been learned by the model. In other words, during the *training* process, ProfilePSTMM attempts to find patterns inherent within a given glycan data set containing many disparate glycan structures. These patterns are then visualized in the results as distributions of monosaccharides at particular positions of a branched topology. Once a pattern has been learned, it can potentially be used to assess, or test, whether a given glycan structure that was not in the original data set should belong to the set or not. That is, it assesses the glycan to see how well it matches the pattern learned.

As for the ProfilePSTMM Tool, it performs the *training* process and simply outputs the *trained* pattern that was *learned* from the data. The model is currently not stored for later *testing* on other data. However, such advanced analyses are being planned at the time of this writing.

The ProfilePSTMM model was originally developed to aid in the understanding of glycan patterns being recognized by glycan-binding agents such as proteins, viruses, and antibodies, whose data can be obtained from glycan array databases. It uses a probabilistic model to overcome the noise that can be found in many of the glycan-binding data available, as well as weak-binding patterns which may skew analytical results. This is in comparison to tree alignment algorithms, which do not allow for such flexibility.

RESOURCE FOR
INFORMATICS OF
GLYCOMES AT
SRIKA

ProfilePSTMM Training

[Home](#)

[Help](#)

[Feedback](#)

Profile name

Glycan data
Direct input or load glycan data
(KCF or KEGG Glycan ID)

G00015
G00016
G00017
G00018
G00019
G00020
G00021
G00022
G00171
G00178
G00198
G00199
G00200
G00202

選択されていません

KCF Glycan ID
Check the inputted glycan data type

Glycan data type

Number of times to shuffle

Fig. 15.5 A snapshot of the default ProfilePSTMM Tool input page. A list of KEGG GLYCAN IDs or KCF data can be specified

15.3.1 ProfilePSTMM Input

The ProfilePSTMM Tool takes as input a data set of glycans. The intention of this tool was to allow for the analysis of glycans that exhibit binding to a particular agent. Thus, data such as that from lectin and glycan arrays can be analyzed with this tool. A snapshot of the input screen is given in Fig. 15.5. A list of KEGG GLYCAN IDs or a concatenated list of KCF data can be specified. The *shuffle number* refers to the number of times the computation should be repeated to avoid local optima. Because this tool is a probabilistic model, it is suggested that it is run at least five times in order to find the optimal solution. A file can also be specified containing the input data, either as KEGG GLYCAN IDs or KCF data, instead of cutting and pasting into the text area.

15.3.2 ProfilePSTMM Output

As a result of training on the default glycan structures using shuffle number 5, the resulting glycan profiles will be displayed, as in Fig. 15.6. Looking closely at these results, one can see that there were basically two distinct profiles that

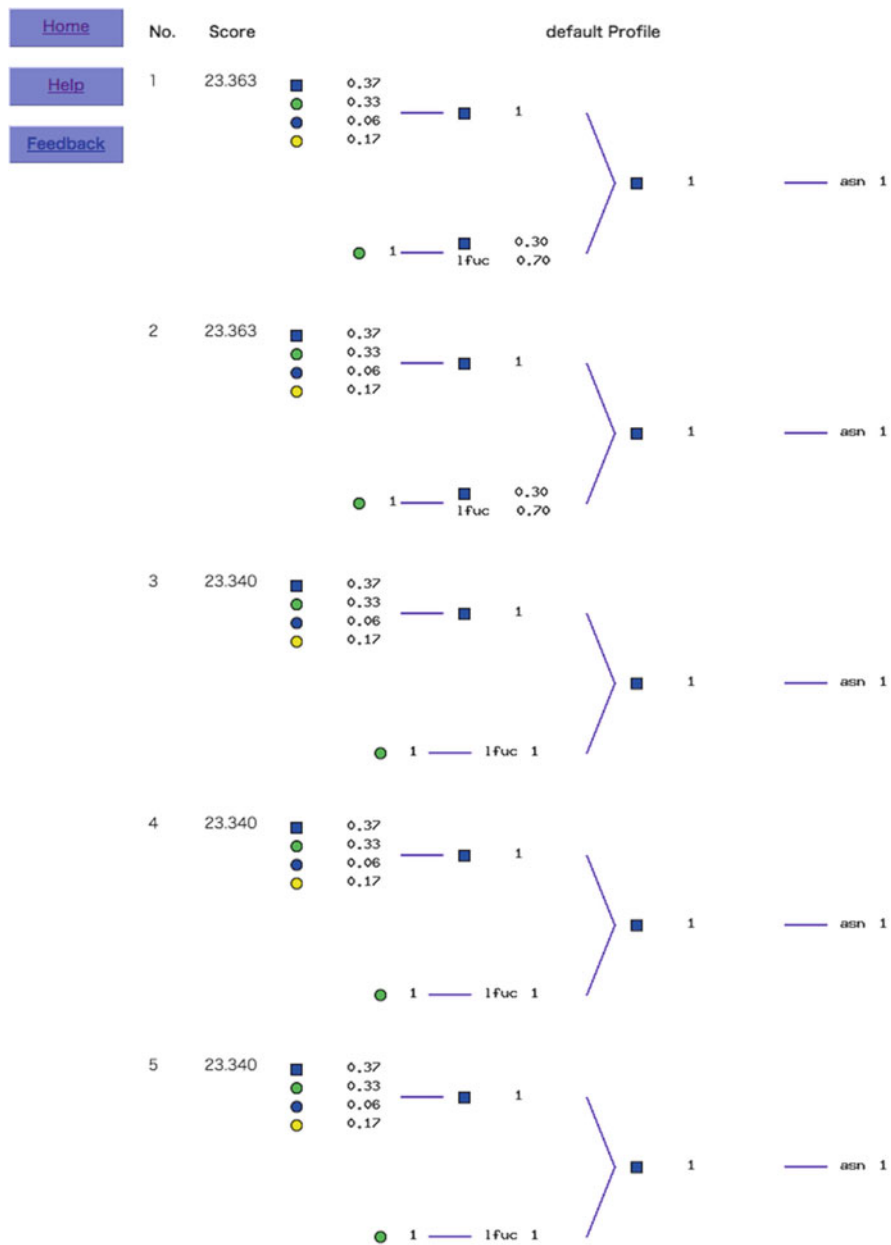


Fig. 15.6 Result screen after running ProfilePSTMM

were trained based on the data set, one with score 23.363 and one with 23.340. In both profiles, the biantennary pattern learned contained the same distribution of monosaccharides at the terminating ends. Mannose appeared 100 % on one end, and

GlcNAc, mannose, glucose, and galactose appeared with a ratio of 37:33:6:17. The single difference between these profiles appears at the neighboring position, where fucose either appears alone (100%) or may be replaced with a 30% probability with GlcNAc. The computed score is actually computed based on the *likelihood* score computed when learning the patterns in the glycan data set. Future work will entail assessing these scores to obtain significant values that can also be displayed to the user.

Note that the shape, or topology, of the profile is currently based on the concept of *maximum common subtree* of the input data, meaning that the shape of the profile that is output will be based on the smallest glycan structure that exists in the input data set. This causes the learning process to combine monosaccharide distributions into a smaller space, causing the results to be misleading. This is a currently known issue that is being updated and will be made more flexible in future versions, when it will be easier to ascertain the results.

15.3.3 ProfilePSTMM Tips

- At the time of this writing, the ProfilePSTMM Tool can be used for glycan data sets that do not contain very small glycan structures, but rather contain glycans that are of comparable size.
- In order to specify stronger binding glycans compared to others in the data set, weights can be implied by specifying multiple copies of heavily binding glycans in the input.
- The computation time may increase with larger shuffling number and larger number of glycans in the input.

15.4 Glycan Miner

The Glycan Miner Tool is based on a mining algorithm called α -closed frequent subtree mining (Hashimoto et al. 2008). This tool takes as input a large data set of glycans and attempts to find significant subtrees among the data. In a sense, it can find unique overly expressed glycan substructures within a set of glycan structures. It is suited for the analysis of glycan profiling data by mass spectrometry or glycan-binding data such as that from glycan arrays.

15.4.1 Glycan Miner Input

The input screen for the Glycan Miner Tool is shown in Fig. 15.7. The default values allow users to test the tool to see what kind of results can be obtained. The data



Glycan Miner

If you are mining CFG glycan array data, a step-by-step manual is available [here](#).

Data set name

Home

Help

Feedback

Please enter glycan structures in KCF format.

ENTRY	G00015		Glycan
NODE	8		
1	Asn	20	0
2	GlcNAc	12	0
3	GlcNAc	3	0
4	Man	-5	0
5	Man	-12	5
6	Man	-12	-5
7	GlcNAc	-20	5
8	GlcNAc	-20	-5
EDGE	7		
1	2:b1	1	
2	3:b1	2:4	
3	4:b1	3:4	
4	5:a1	4:6	

Or load it from disk 選択されていません

alpha (value between 0 and 1):

minimum support:

Fig. 15.7 Input screen for Glycan Miner with default values

should be specified in KCF format and can be inputted into the text area or specified with a filename containing the data. There are two options available: *alpha* and *minimum support*. These parameters refer to how a candidate subtree (extracted from the data set) should be handled. The *support* for any candidate subtree refers to the number of glycans in the input data set in which it appears. For example, if a data set of 80 structures, 50 *N*-glycans and 30 *O*-glycans, is given, the subtree consisting of the *Man*₃*GlcNAc*₂ core structure of *N*-glycans would have a support of 50. Candidate subtrees are computed from the input data set; they consist of all possible subtrees among the data.

The second parameter, *alpha*, essentially specifies the uniqueness of the subtrees to be output. An example would best illustrate this parameter, so two output examples are provided in Sect. 15.4.2. In simple terms, a larger value of *alpha*, closer to 1, will output a larger range of glycan subtrees, whereas a smaller value, closer to 0, will filter out redundant glycan subtrees.

15.4.2 Glycan Miner Output

Figures 15.8 and 15.9 are snapshots of the Glycan Miner results using the same support = 1 but different values for α . As illustrated, the results of $\alpha = .9$ contain

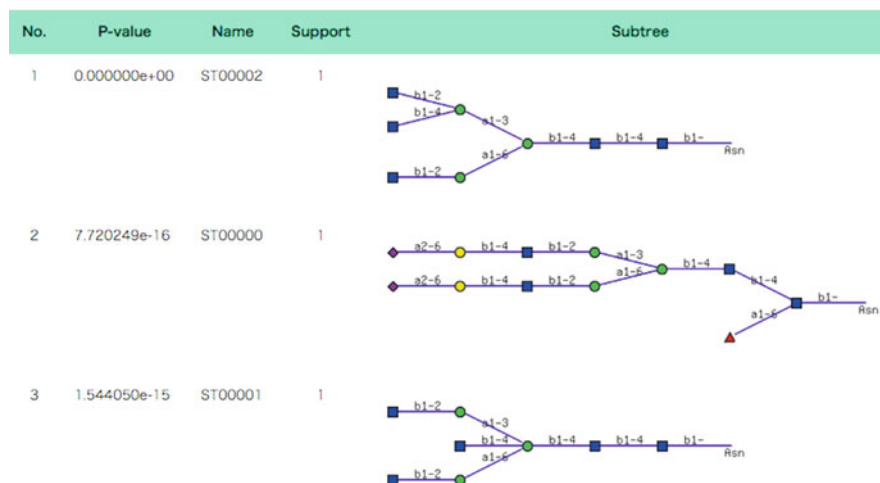


Fig. 15.8 Results of running Glycan Miner on the default data set, using a value of 1 for support and .3 for α

seven glycans, whereas those of $\alpha = .3$ contain just three. The difference lies in the fact that several of the seven glycans are rather similar to one another. More specifically, their support values are rather close, whereas the three selected using $\alpha = .3$ are less similar to one another.

The α -closed frequent mining algorithm compares the *support* between every glycan subtree pair in which one contains the other. Let us assume that there are two subtrees, the *N*-glycan core $Man_3GlcNAc_2$ which we call *largerT* and a subtree of this structure $Man_2GlcNAc_2$ which we call *smallerT*. The α -closed frequent mining algorithm will compare each of their respective supports (denoted as $supp(largerT)$ and $supp(smallerT)$) and will discard the larger if its support does not satisfy the equation $supp(largerT) < \max(\alpha \times supp(smallerT), minsup)$, where *minsup* is the *minimum support* specified by the user. Since it can be assumed that subtrees with larger values of *support* will be less overlapping, α can be used to make this differentiation.

The Glycan Miner Tool also sorts the resulting subtrees based on *p*-value, which is generated based on the data set inputted. That is, a random data set of subtrees are generated, with which the support of the subtrees in the results are compared. The smaller the *p*-value, the fewer the chance that the same results are obtained randomly, and so the subtrees with smaller *p*-value are listed first.

15.4.3 Glycan Miner Tips

- *p*-values are generated based on the input, so the *p*-value on smaller data sets will not be very meaningful. Moreover, because it is generated for each run,

No.	P-value	Name	Support	Subtree
1	0.000000e+00	ST00004	1	
2	7.720249e-16	ST00001	6	
3	1.544050e-15	ST00003	1	
4	2.316075e-15	ST00000	3	
5	3.088100e-15	ST00005	1	
6	3.860124e-15	ST00006	2	
7	4.632149e-15	ST00002	2	

Fig. 15.9 Results of running Glycan Miner on the default data set, using a value of 1 for support and .9 for α

the order of the results may change, especially for small data sets. Such results would imply that there is no significant difference between the results and that they should all be treated equally.

- Glycosidic bond information is very specific: unspecified linkage information will not be matched with structures whose linkages are specified.
- IUPAC-formatted data can be used for Glycan Miner by following the link at the top of the input page for using CFG data.

15.5 MCAW

MCAW, or multiple carbohydrate alignment with weights, was originally developed as a temporary tool to determine the shape of profiles for ProfilePSTMM. However, it was decided that MCAW could be used on its own to align multiple glycans to find consensus sequence patterns across a set of glycans, so it was made available as a web tool. MCAW generates easy-to-understand glycan profile images as results. Moreover, detailed parameters are available for advanced users who wish to fine-tune the results.

Similar to the ProfilePSTMM Tool, MCAW can be used to find patterns across a set of glycans, such as from glycan-array or lectin-array data. Note, however, that it is not a probabilistic model, meaning that all the input data will be used at face value, meaning that all the data will be used in the results. In contrast, a probabilistic model will account for noise in the data, and so discard data that may be outliers.

15.5.1 MCAW Input

The MCAW input screen is shown in Fig. 15.10. Multiple KCF data can be pasted into the text area, or a file containing such data can be selected. The advanced parameters are described in Table 15.2. The default values can be used normally, but these options allow users to fine-tune the results if specific characteristics of the data inputted are known.

15.5.2 MCAW Output

The output of MCAW is a figure, as in Fig. 15.11, of the glycans aligned into a profile, but the results in PKCF (profile KCF format) can also be downloaded by the link provided. A legend for the figure is also provided as a link.

In the figure, similar to ProfilePSTMM, distributions of glycans and their linkages are given at various positions of glycans. Highly consistent portions of the alignment will have higher distributions of a particular monosaccharide at those positions. Figure 15.11 illustrates an example. One can see that at the positions numbered 7 through 10, highly *conserved* patterns of monosaccharides are found. This is a sialyl-Lewis X structure, which is a common motif among glycans in mammals. Moreover, position 18, which is attached to position 7, also contains a 6-sulfate approximately 16 % of the time, indicating that 6-sulfated sialyl-Lewis X structures can also be found in the data.

Data set name

Enter the glycan structures in KCF format:

```

ENTRY      G04845          Glycan
COMPOSITION (Gal)3 (Glc)1 (GlcNAc)2 (LFuc)2 (Neu5Ac)1
MASS       1656.5
DBLINKS    CCSD: 23949
           GlycomeDB: 20420
           JCGGDB: JCGG-STR011245
NODE       9
           1 Glc      0  0
           2 Gal     -10  0
           3 GlcNAc  -20  10
           4 GlcNAc  -20 -10
           5 Gal     -30  15
           6 LFuc    -30  5
           7 LFuc    -30 -5
           8 Gal     -30 -15
           9 Neu5Ac  -40  15
EDGE       8
           1  2:b1  1:4
  
```

Or load KCF from a file: 選択されていません

Advanced weighting options

Gap penalty:

Monosaccharide:

- Linkage information -

Anomer:

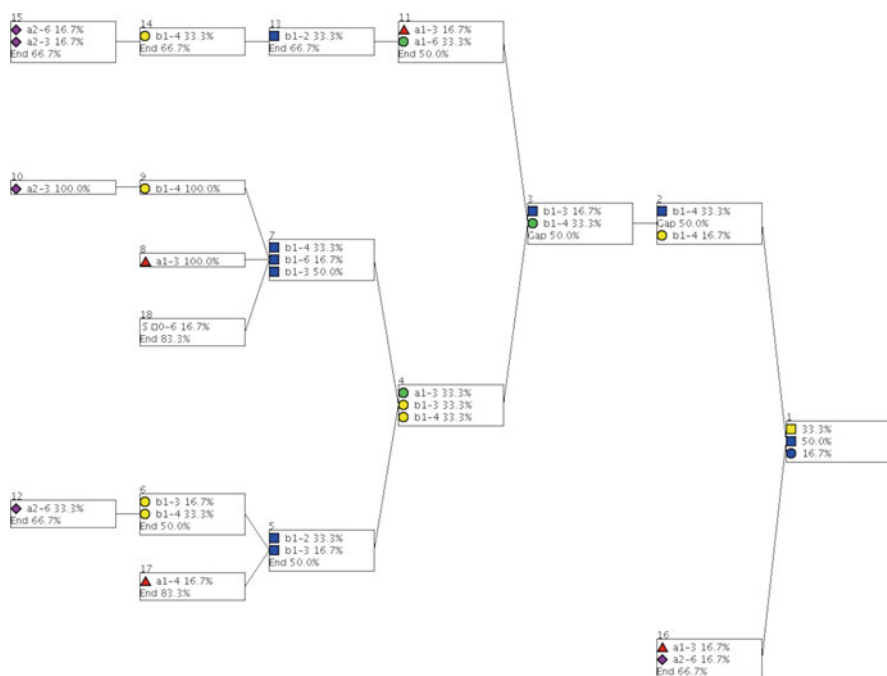
Non reducing side carbon number:

Reducing side carbon number:

Fig. 15.10 The MCAW input screen, where multiple KCF data can be specified, and a variety of parameters are available for fine-tuning the results

Table 15.2 The advanced parameters in MCAW

Parameter	Description
Gap penalty	The penalty value to use when inserting a gap where a monosaccharide should align. Used to shift matching subportions of glycans when necessary to align the most similar parts
Monosaccharide	The score for matching monosaccharides
Anomer	The score for matching anomers
Nonreducing side carbon number	The score for matching the carbon number of monosaccharide on the nonreducing side of glycosidic bonds
Reducing side carbon number	The score for matching the carbon number of monosaccharides on the reducing side of glycosidic bonds (usually 1 or 2)

**Fig. 15.11** A snapshot of the MCAW output screen, where distributions of monosaccharides and their glycosidic linkages are given at various positions

15.5.2.1 PKCF Format

The output of MCAW can also be downloaded as a PKCF (Profile KCF) file from the result page. The PKCF for the result in Fig. 15.11 is as follows.

```

ENTRY G04804-G04183-G04845-G05121-G05108-G04329 GlycanProfile
NODE 18
1 1=GlcNAc 2=GlcNAc 3=Glc 4=GalNAc 5=GalNAc 6=GlcNAc 0 0
2 1=GlcNAc 2=GlcNAc 3=- 4=- 5=- 6=Gal -8 6
3 1=Man 2=Man 3=- 4=- 5=- 6=GlcNAc -16 6
4 1=Man 2=Man 3=Gal 4=Gal 5=Gal 6=Gal -24 1
5 1=GlcNAc 2=GlcNAc 3=GlcNAc 4=GlcNAc 5=GlcNAc 6=GlcNAc -32 5
6 1=LFuc 2=LFuc 3=LFuc 4=LFuc 5=LFuc 6=LFuc -40 5
7 1=Gal 2=Gal 3=Gal 4=Gal 5=Gal 6=Gal -40 7
8 1=Neu5Ac 2=Neu5Ac 3=Neu5Ac 4=Neu5Ac 5=Neu5Ac 6=Neu5Ac -48 7
9 1=GlcNAc 2=GlcNAc 3=GlcNAc 4=0 5=0 6=0 -32 -3
10 1=Gal 2=Gal 3=Gal 4=0 5=0 6=0 -40 -2
11 1=Man 2=Man 3=0 4=0 5=0 6=LFuc -24 11
12 1=Neu5Ac 2=Neu5Ac 3=0 4=0 5=0 6=0 -48 -2
13 1=GlcNAc 2=GlcNAc 3=0 4=0 5=0 6=0 -32 11
14 1=Gal 2=Gal 3=0 4=0 5=0 6=0 -40 11
15 1=Neu5Ac 2=Neu5Ac 3=0 4=0 5=0 6=0 -48 11
16 1=0 2=0 3=0 4=Neu5Ac 5=0 6=LFuc -8 -6
17 1=0 2=0 3=LFuc 4=0 5=0 6=0 -40 -4
18 1=0 2=0 3=0 4=0 5=S 6=0 -40 3
EDGE 17
1 2-1:b1 1-1:4
1 2-2:b1 1-2:4
1 2-3 1-3
1 2-4 1-4
1 2-5 1-5
1 2-6:b1 1-6:4
2 3-1:b1 2-1:4
2 3-2:b1 2-2:4
2 3-3 2-3
2 3-4 2-4
2 3-5 2-5
2 3-6:b1 2-6:3
.
.
17 18-3 5-3
17 18-4 5-4
17 18-5 5-5:6
17 18-6 5-6
///

```

The general format of PKCF follows that of KCF, in that there are three major sections in order: ENTRY, NODE, and EDGE. The first section lists the glycan names as inputted into the MCAW tool, separated by hyphens, read in from the KCF files. The order of the names indicates the order of the items listed in the NODE and EDGE sections. The NODE and EDGE labels are followed by numbers, indicating

the number of positions and connected positions, respectively. Usually, the number of edges is one less than the number of positions. The lines following the NODE heading are the details regarding each position. They are formatted as follows:

```
<pos#> <glycan#1>=<residue> <glycan#2>=<residue> ...<glycan#n>=
<residue> <x-coord> <y-coord>
```

where `pos#` indicates the position number within the tree topology and `x-coord` and `y-coord` represent their *x*- and *y*-coordinates, respectively, as drawn. Thus at each position, the residue names derived from each glycan are listed in order of the glycans listed in the ENTRY line. This is followed by *x*- and *y*-coordinates, similar to the KCF format.

In contrast, the EDGE section consists of lines corresponding to each glycosidic bond in each glycan. Thus the number of lines in this section equals the number of connected positions times the number of glycans. In the example above, there are 17 edges in the tree topology and 5 glycans, so there is a total of 85 lines, which have been cut to save space. Each line in this section is formatted as follows:

```
<link#> <pos#>-<glycan#>:<anomer><carbon#> <pos#>-<glycan#>:
<carbon#>
```

where `link#` is a unique number given to each edge. However, if the anomer and/or carbon number information is unavailable, these are left out. If both are unavailable, then the colon “:” is omitted. This is, again, a similar format to the KCF representation, except with additional position information included.

15.5.3 MCAW Tips

- The input data should be in KCF format, but with the extra restriction that a name must be included in the ENTRY line and that the name should not have a hyphen in it.
- MCAW is still in alpha-stage and may not give results for complex data sets, such as those having a large variety of structures. It may work better on data sets with similar glycan structures.
- Glycans with no glycosidic bonds (i.e., single monosaccharides) cannot be included in the input data.
- As of May 2016, new functionality has been made available such that weights can be specified with IUPAC-formatted data: follow the link for using CFG array data, available at the top of the input screen of MCAW.

15.6 Kernel Tool

The Glycan Kernel Tool in RINGS is an implementation of the biochemically weighted Kernel (Jiang et al. 2011). This tool implements a kernel that utilizes a score matrix for weighting similar glycosidic linkages and monosaccharides

more highly in order to extract biochemically meaningful features from the data. Combined with a support vector machine (SVM), this tool can be used to detect biologically significant motifs across two data sets of glycans. It has been shown to extract the same motifs as those that were manually curated, which were consistent with the literature. This detection procedure consists of using the SVM to classify the two input data sets based on the kernel, and then feature extraction is performed based on the classification to detect the most significant features, or glycan substructures, that most distinguish the two data sets. In layman's terms, given two data sets of glycan structures to compare, this tool will find those glycan substructures that distinguish one data set from the other. That is, it will find substructures that are unique to the target data set, compared to the control.

15.6.1 Kernel Tool Input

Two data sets of glycans are required as input for comparison. The *target* data set of glycans is compared against the *control* data set to find unique substructures of glycans, or motifs. Figure 15.12 is a snapshot of the input screen for the Glycan

Home

Help

Feedback

Please input your kernel calculation ID to retrieve your previous result.

Data set name

Glycan Data
Enter KCF data consecutively.

Target data

```
ENTRY 1024_3_1 Glycan
NODE 4
1 gal -8.0 -2.0
2 glnac -8.0 2.0
3 gainac 0.0 -0.0
4 gln -14.0 2.0
EDGE 3
1 1:b1 3:3
2 2:b1 3:6
3 4:b1 2:4
///
ENTRY 1140_4_2 Glycan
NODE 4
1 neuac -16.0
-2.0
```

Or load KCF from a file: 選択されていません

Control data

```
ENTRY 1099_6 Glycan
NODE 4
1 gal -17.0 9.0
2 glnac -8.0 9.0
3 man 1.0 9.0
4 neuac -25.0
9.0
EDGE 3
1 2:b1 3:2
2 1:b1 2:4
3 4:a2 1:3
///
ENTRY 1273_3 Glycan
NODE 5
1 gal -17.0 9.0
```

Or load KCF from a file: 選択されていません

option

Subtree size min <= size <= max

Specify the range of subtree size (the number of monosaccharides) from one to nine for breaking down glycan structures.

Fig. 15.12 A snapshot of the input screen for the Glycan Kernel Tool. Two text areas are provided for the user to provide KCF-formatted data of two data sets of glycans to compare. Because of the potentially long computation time, previously executed runs of this tool are assigned computation IDs, which can be retrieved from the text field on the left-hand side

Kernel Tool. Two text areas are provided for the user to provide KCF-formatted data of two data sets of glycans to compare. However, large numbers of glycans in the data sets will take longer to compute. Therefore, computation IDs are assigned to each run of this tool. Thus, users can close this browser window and retrieve the results later once they are complete by returning to this page and inputting the computation ID in the text field on the left-hand side of the page.

15.6.2 Kernel Tool Output

The output of this tool is a list of the glycan features and their *layers*, along with their scores, as shown in Fig. 15.13. Details are described in the original

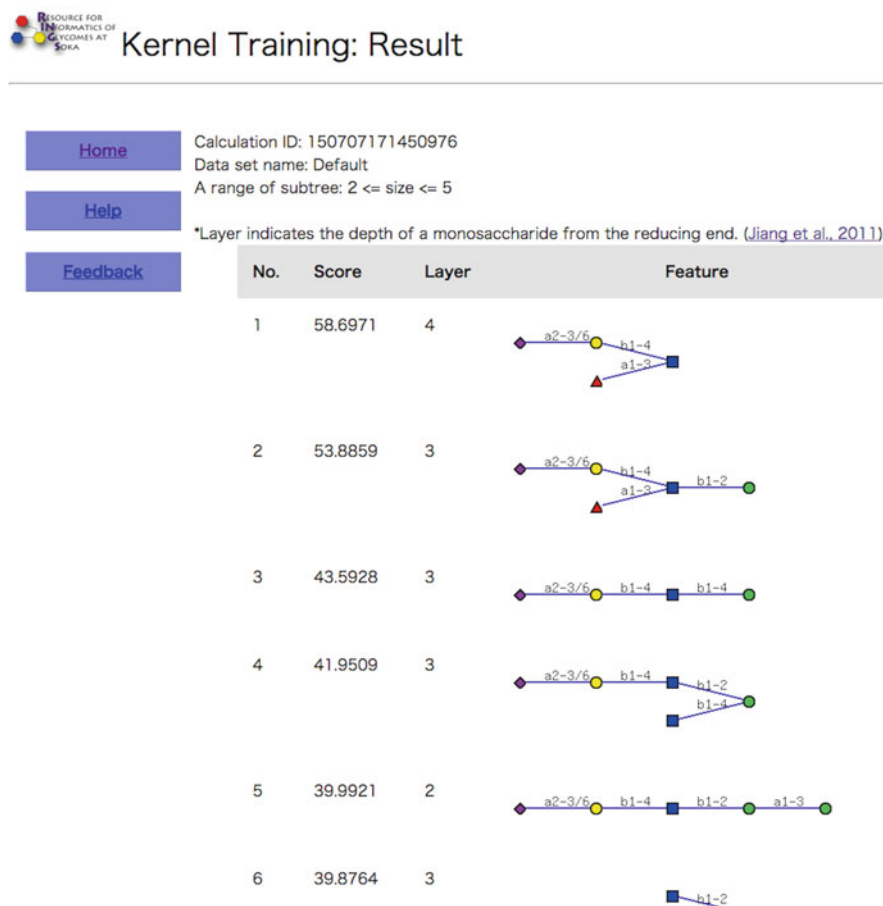


Fig. 15.13 A snapshot of the output screen for the Glycan Kernel Tool. The list of glycans, their scores, and their layers are listed in order of significance

manuscript (Jiang et al. 2011), but in general terms, the *layer* of the glycan feature is the distance of the given feature from the root node, or reducing end, of the original glycan within which it was found. That is, the layer indicates the number of glycosidic linkages between the feature and the reducing end. Thus, features with larger layers are those that are found closer to the nonreducing end of the input glycan(s). In Fig. 15.13, the top feature is actually a subtree of the second highest feature, which has an additional mannose (green circle) at its reducing end. Accordingly, their layers differ by one. Other features are also subsets of these top two features, indicating that this top feature is probably the single motif that most distinguishes the two data sets.

15.6.3 Kernel Tool Tips

- Data sets of over 40 structures may take over an hour to compute. Please be sure to save the computation ID for larger data sets.
- Since substructures unique to the target and not in the control will be extracted by this tool, swapping target and control data sets may result in interesting results, where structures unique to the control and not the target can be found.

15.7 GPP: Glycan Pathway Predictor

The glycan pathway predictor (GPP) is a tool that computes the *N*-glycans that could be theoretically biosynthesized given a set of glyco-enzymes involved with *N*-glycan biosynthesis. This generation procedure is based on the mathematical model proposed by Krambeck et al. in (2009). In short, the substrate specificities of each enzyme are specified using a mathematical model, and it is assumed that there are no limitations in the amount of each enzyme, sugar donors, or substrates. Then, GPP computes all possible glycan that can be synthesized using the input glycan structure and selected enzymes. The output results in a network of the synthesized glycans.

15.7.1 GPP Input

The GPP input screen is shown in Fig. 15.14. A single *N*-glycan formatted in KCF should be specified in the text area on the left. Then one or more enzymes should be selected from the list on the right. Finally, because this mathematical model could potentially continue indefinitely, a limit for the largest glycan structure to generate, indicated by molecular mass, should be specified. Table 15.3 gives a description of each enzyme that is available in GPP.

Data set name

Enter a glycan structure in KCF format:

```

ENTRY      G01234      Glycan
COMPOSITION (Gal)4 (GlcNAc)6 (LFuc)1 (Man)3
(Asn)1
MASS       2500.3 (Asn)
CLASS      Glycoprotein; N-Glycan
DBLINKS    CCSD: 2082 2992 2994 5199
13317
NODE       15
1  Asn      29  -3
2  GlcNAc   20  -3
3  GlcNAc   10  -3
4  Man      0   -3
5  Man     -10   4
6  Man     -10  -10
7  GlcNAc  -20   9
8  GlcNAc  -20  -1
9  GlcNAc  -20  -6
10 GlcNAc  -20 -14

```

Select one or more enzymes:

ManI
ManII
a6FucT
GnTI
GnTII
GnTIII
GnTIV
GnTV
iGnT
b4GalT
a3SiaT
iGnT
a6SiaT
b3GalT
FucTLe
FucTII
a3FucT
GalNAcT-A
GalT-B

Use the "Ctrl" key to select multiple enzymes. The "Shift" key can also select consecutive enzymes.

Or load KCF from a file:

選択されていません

Max of Molecular Mass:

Fig. 15.14 A snapshot of the input screen for the GPP Tool. A single *N*-glycan formatted in KCF should be specified. Then one or more enzymes should be selected from the list on the right. Finally, the mass limit for the generated glycans should be entered in order to prevent the computation from continuing indefinitely

15.7.2 GPP Output

The output screen for GPP is a Java applet. Therefore, similar to the DrawRINGS applet, Java security permissions of the web browser used will need to be adjusted accordingly in order to obtain the results. An open-source applet called ZGRViewer (<http://zvtm.sourceforge.net/zgrviewer.html>) is used to display the resulting pathway. A snapshot of the output screen is shown in Fig. 15.15. The functionality available in ZGRViewer includes zoom, swipe, and double-clicking of nodes. Figure 15.16 is a snapshot of the zoomed in screen. Nodes represent each glycan structure synthesized, and they are connected by edges representing the enzyme involved in the biosynthesis reaction. Nodes can be double-clicked to open a browser window displaying the detailed information of the glycan, as in Fig. 15.17.

15.7.3 GPP Tips

- The glycan structure used as the first structure must be an *N*-linked glycan and a substrate for at least one of the enzymes selected.
- The resulting pathway map may take some time to load. Clicking on the canvas once may refresh the map and display it.

Table 15.3 The glyco genes available in GPP

Abbreviation	Name	EC number	Corresponding genes
ManI	α 2-mannosidase I	3.2.1.113	MAN1A1, MAN1A2, MAN1B1, MAN1C1
ManII	α 3/6-mannosidase II	3.2.1.114	MAN2A1, MAN2A2
GnTI	β 2- <i>N</i> -acetylglucosaminyltransferase I	2.4.1.101	MGAT1
GnTII	β 2- <i>N</i> -acetylglucosaminyltransferase II	2.4.1.143	MGAT2
GnTIII	β 4- <i>N</i> -acetylglucosaminyltransferase III	2.4.1.144	MGAT3
GnTIV	β 4- <i>N</i> -acetylglucosaminyltransferase IV	2.4.1.145	MGAT4A, MGAT4B
GnTV	β 6- <i>N</i> -acetylglucosaminyltransferase V	2.4.1.155	MGAT5
b3GalT	β 3-galactosyltransferase		B3GALT1, B3GALT2, B3GALT5
b4GalT	β 4-galactosyltransferase	2.4.1.38	B3GALT1, B4GALT2, B4GALT3
a3SiaT	α 3-sialyltransferase	2.4.99.6	ST3GAL3
a6SiaT	α 6-sialyltransferase	2.4.99.1	ST6GAL1
a6FucT	α 6-fucosyltransferase	2.4.1.68	FUT8
FucTLe	α 3/4-fucosyltransferase III	2.4.1.65	FUT3, FUT5, FUT6
FucTH	α 2-fucosyltransferase Se, H	2.4.1.69	FUT1, FUT2
a3FucT	α 3-fucosyltransferase	2.4.1.152	FUT4, FUT7, FUT9
iGnT	Blood group i β 3- <i>N</i> -acetylglucosaminyltransferase	2.4.1.149	B3GNT1, B3GNT6
IGnT	Blood group I β 6- <i>N</i> -acetylglucosaminyltransferase	2.4.1.150	GCNT2
GalNAcT-A	Blood group A α 3- <i>N</i> -acetylglucosaminyltransferase	2.4.1.40	ABO
GalT-B	Blood group B α 3-galactosyltransferase	2.4.1.37	ABO

- The detail screen may not display in newer browsers due to Java applet security issues. GPP will be updated to include more enzymes in the near future, and the web display will also be updated accordingly.

15.8 Converter Utility

The Convert Tool in RINGS is a tool that aims to combine and automate the various glycan text format conversion utilities currently available in RINGS. It takes as input any of the major glycan text formats and provides options to convert the structures in the supported output format.



GPP Results

It may take some time to display large graphs.

Input Enzyme : ManI/ ManII/ a6FucT/ GnTI/ GnTII/ GnTIII/ GnTIV/ GnTV/ iGnT/ b4GalT/ a3SiaT/ IGnT/ a6SiaT/ b3GalT/ FucTLe/ FucTH/ a3FucT/ GalNAcT-A/ GalT-B/

Click [here](#) for a text file of these results.

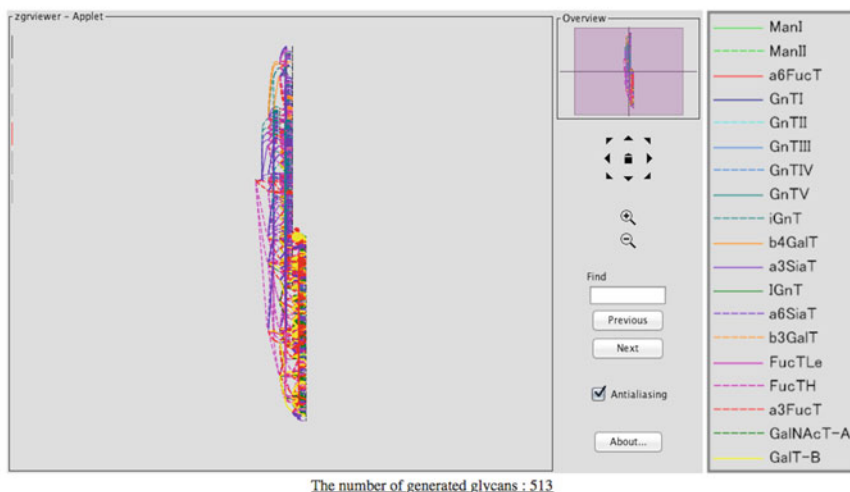


Fig. 15.15 A snapshot of the output screen of the GPP Tool

15.8.1 Convert Utility Input

The Convert Tool takes as input glycan structures in GLYDE-II, GlycoCT_{condensed}, KCF, IUPAC, LinearCode, and LINUCS formats. If multiple glycans are inputted, multiline formats such as GlycoCT and KCF need to be inputted with newlines in between. After clicking the “Submit” button, the representation format of the input is automatically recognized, and a list of possible output formats is provided as a pull-down menu. Figure 15.18 is a snapshot of the input screen for the Convert Tool, using GLYDE-II as an example. Note that only one structure in GLYDE-II can be converted at a time due to formatting issues in XML. Figure 15.19 shows how the input format of the text in Fig. 15.18 was recognized and that the possible representation formats to which it can be converted are KCF, GlycoCT_{condensed}, IUPAC, LinearCode, and MDL Mol. Moreover, the output file format can be selected from HTML, Text, JSON, and DL. HTML will display the results graphically, with images of the glycan structures that were converted. Text will display the results in plain text format, which is easier to process by computer and can be copied as text. JSON is a special format for computers to process the results, and DL refers to download, allowing users to download the results as a text file instead of displaying all the results in the browser, saving network bandwidth.

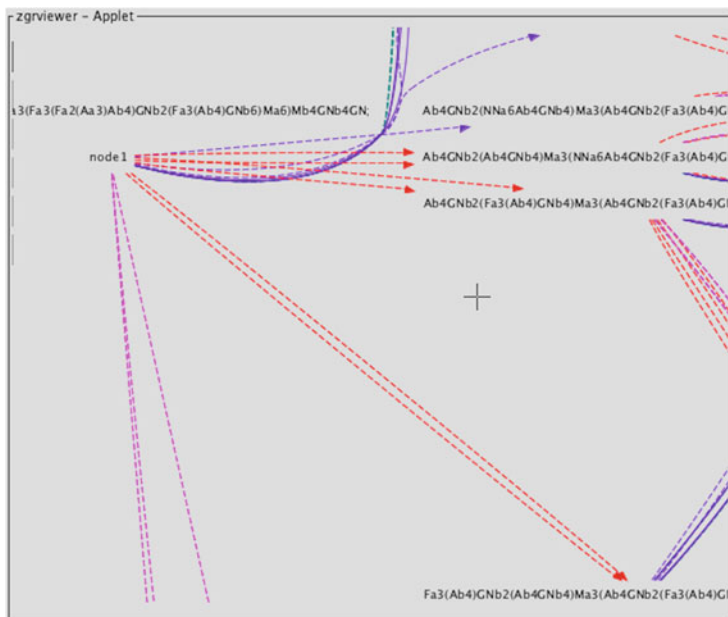


Fig. 15.16 A snapshot of the output screen zoomed in

15.8.2 Convert Utility Output

The output screen depends on the output formatted selected from the input page. Figure 15.20 is a snapshot of when the HTML format is selected. If either KCF or GlycoCT is selected as the output representation format, then a figure of each converted structure is also provided. This allows the user to visually confirm that the converted structure is correct.

15.8.3 Convert Utility Tips

- Depending on the conversion path via which input formats are converted to the selected output formats, some information may be lost in the process of conversion. This may further cause errors during the conversion. In that case, it is suggested that a format that contains maximal structural information, such as GLYDE-II or GlycoCT, be used as the input format. To do so, the existing input format can be translated to GLYDE-II or GlycoCT, and any missing information should be added manually before being translated to the target format.
- GLYDE-II-formatted structures can only be converted one structure at a time, due to formatting issues of XML.

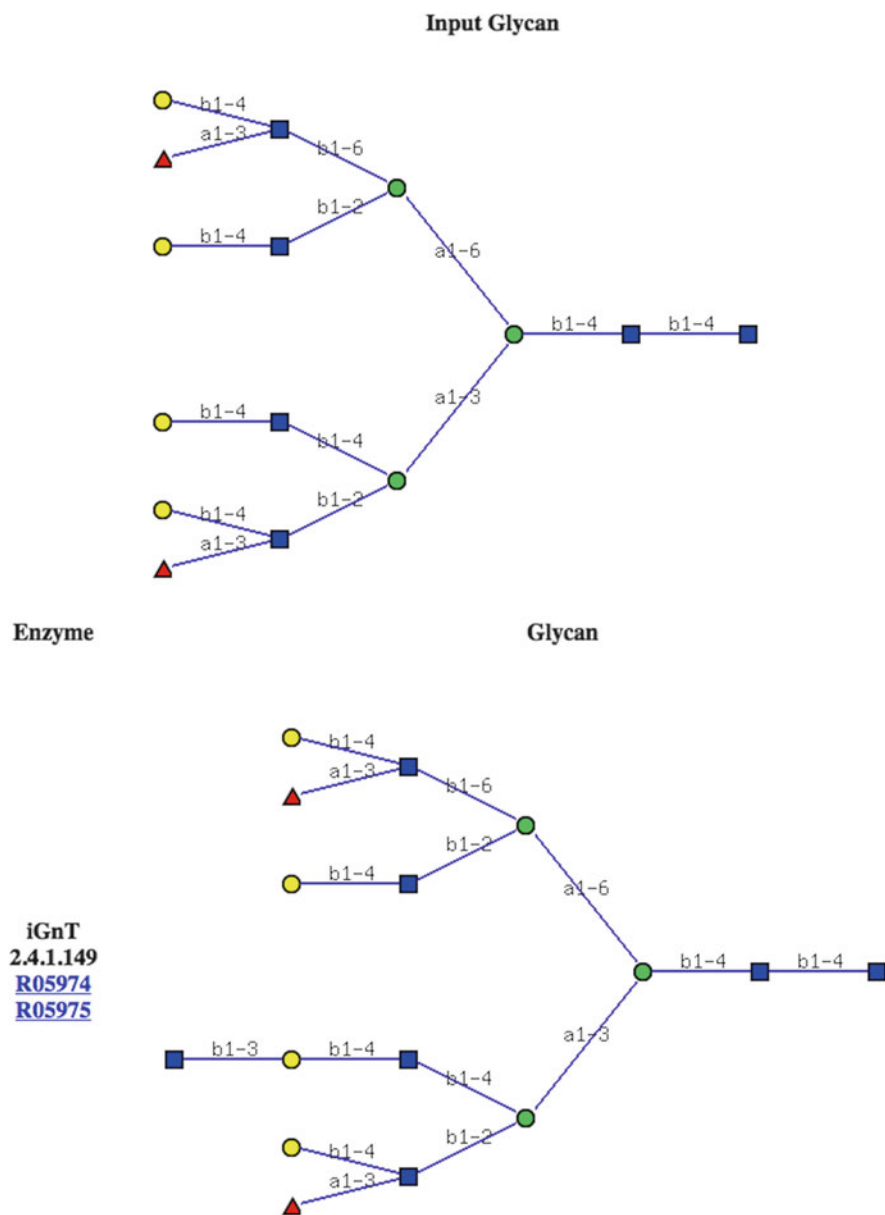


Fig. 15.17 A snapshot of the detail screen for a glycan selected from the output of the GPP Tool



Convert Tool

[Home](#)
[Help](#)
[Feedback](#)

Data set name

Please enter glycan structures in any format.

```
<?xml version="1.0" encoding="UTF-8"?>
<sugar version="1.0">
<residues>
  <basetype id="1" anomer="x" superclass="hex" ringStart="-1" ringEnd="-1" name="x-dman-HEX-x:x">
    <stemtype id="1" type="dman" />
  </basetype>
  <basetype id="2" anomer="a" superclass="hex" ringStart="1" ringEnd="5" name="a-dman-HEX-1:5">
    <stemtype id="1" type="dman" />
  </basetype>
  <basetype id="3" anomer="b" superclass="hex" ringStart="1" ringEnd="5" name="b-dglic-HEX-1:5">
    <stemtype id="1" type="dglic" />
  </basetype>
</residues>
</sugar>
</?xml>
```

Or load it from disk

Fig. 15.18 A snapshot of the input screen for the Convert Tool. In this example, GLYDE-II is used as input. Note that only one structure at a time can be converted from GLYDE-II format



Convert Tool

[Home](#)
[Help](#)
[Feedback](#)

Data set name

Please enter glycan structures in any format.

```
<?xml version="1.0" encoding="UTF-8" ?>
<GlydelI>
<molecule subtype="glycan" id="From_GlycoCT_Translation">
<residue subtype="base_type" partid="1" ref="http://www.monosaccharideDB.org/GLYDE-II.jsp?G=b-dglic-HEX-1:5" />
<residue subtype="substituent" partid="2" ref="http://www.monosaccharideDB.org/GLYDE-II.jsp?G=n-acetyl" />
<residue subtype="base_type" partid="3" ref="http://www.monosaccharideDB.org/GLYDE-II.jsp?G=a-lgal-HEX-1:5[6:d]" />
<residue subtype="base_type" partid="4" ref="http://www.monosaccharideDB.org/GLYDE-II.jsp?G=b-dgal-HEX-1:5" />
<residue subtype="base_type" partid="5" ref="http://www.monosaccharideDB.org/GLYDE-II.jsp?G=a-dgro-dgal-NON-2:6[1:a]2:keto[3:d]" />
</molecule>
</GlydelI>
```

Please select output format.

KCF
 GlycoCT(condensed)
 IUPAC
 LinearCode
 MDL Mol
 HTML TEXT JSON DL

Fig. 15.19 The output format selection screen of the Convert Tool after a structure(s) has been inputted. The input format is automatically recognized, and a pull-down menu of the available output formats is displayed



Convert Results

QueryNumber	GlycoCT(XML)	KCF	Image(KCF)
1		ENTRY Glycan NODE 7 1 man 0 0 2 man -8 1 3 glcnac -16 1 4 gal -24 1 5 man -8 -1 6 glcnac -16 -1 7 gal -24 -1 EDGE 6 1 2:a1 1:3 2 3:b1 2:2 3 4:b1 3:4 4 5:a1 1:6 5 6:b1 5:2 6 7:b1 6:4 ///	

Fig. 15.20 A snapshot of the Convert Tool results page, displayed in HTML format. Each converted structure is displayed along with its figure so that users can visually confirm the structure

15.9 Data Management System

A data management system is available for users who have registered an account on RINGS. Registration is free of charge, and all data is kept private on the RINGS server. The purpose of this system is to allow users to store their past analysis results online without having to download the results of each tool individually. Moreover, data and results stored on RINGS from one tool can be used as input to other tools when the data formats are compatible. Thus, this system enables users to use the RINGS tools more efficiently.

15.9.1 Data Management System Usage

Users can create an account with their email address by clicking the “User registration form” link located at the top of the main RINGS page (Fig. 15.21). After successfully creating an account, users can then log-in with their registered email address and password via the Login form at the top right of this page.

Once users have logged in with their registered email address and password, the main page displays a data tree at the left, as in Fig. 15.22. As shown in this figure, by clicking on the name of the data set itself, the name of the data set and comments regarding the data set can be inputted.

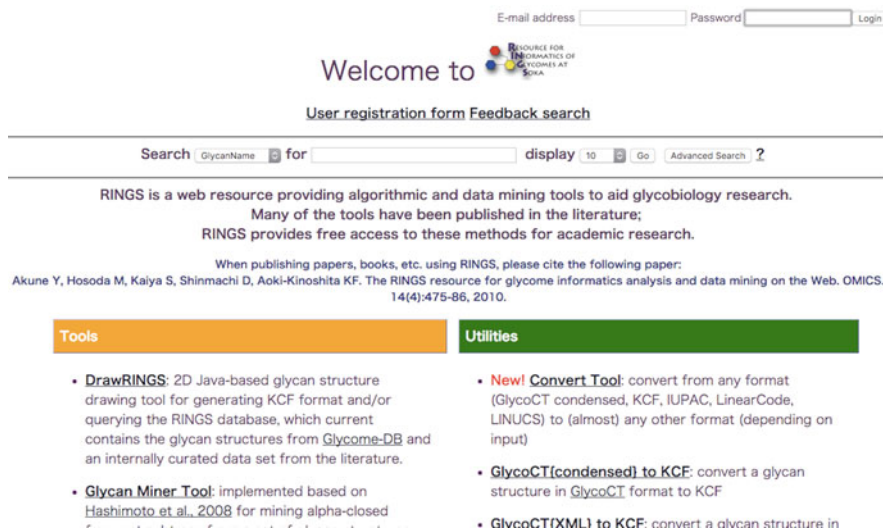


Fig. 15.21 Top of RINGS home page, where links to create a user account are available via the “User registration form” link. Registered users can then log-in with their registered email address and password, located at the top right of this page

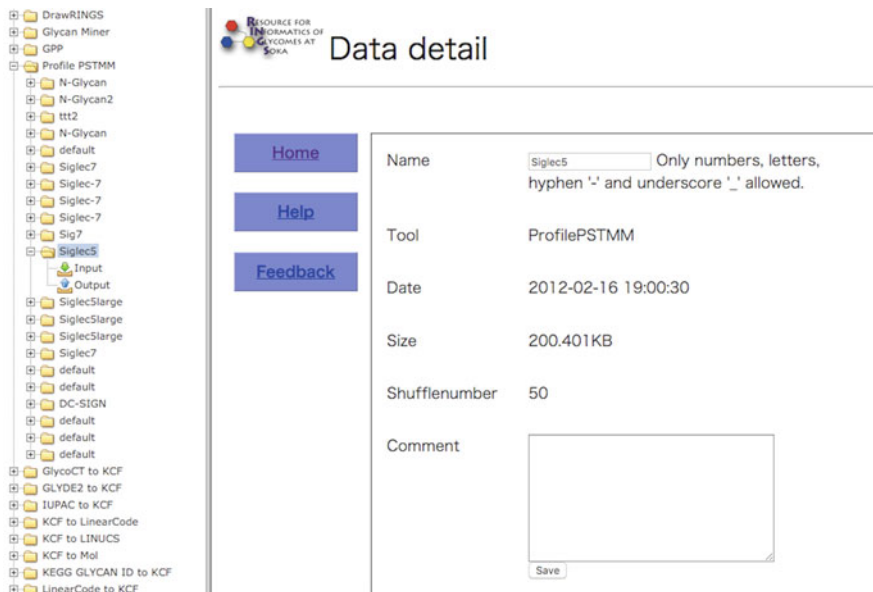


Fig. 15.22 Details regarding a data set can be inputted by clicking the name of the data set. The name can be modified and comments can be stored with each data set

Figure 15.23 is a view of the input data of the data set in Fig. 15.22. Having clicked the “Input” folder for this data set, the glycan structures used for input to ProfilePSTMM on this data set is shown on the right. A pull-down menu is available at the top left for other tools that can also take this data as input. The data set itself can also be downloaded via the “Download” link at the top right.

By clicking on the “Output” folder of a data set, the results of the analysis using the input data are shown. In Fig. 15.24, the output view for ProfilePSTMM is shown using the Profile Viewer. An option to view the results in text format is also provided on this page.

Finally, it is highly recommended that users log out using the “Logout” button at the bottom of the data tree when analysis is complete. This will ensure that the data is kept private and cannot be viewed by others. Note that the size of all the data is also indicated at the bottom. Although there is currently no quota, users can check their data usage and delete any unnecessary data as needed. Data can be deleted by clicking on a tool name in the data tree and then checking the data they want to delete. The “Delete” button is shown at the bottom of the page.

The screenshot displays the RINGS web interface. On the left, a hierarchical tree view shows folders such as 'DrawRINGS', 'Glycan Miner', 'GPP', 'Profile PSTMM', and 'Input'. The 'Input' folder is selected, revealing three glycan structures: 'gla46n1', 'gla46n2', and 'gla46n3'. Each structure is represented by a diagram with nodes and branches, labeled with 'S', 'a2-3', 'b1-4', and '-F'. A pull-down menu is open over the 'Input' folder, listing tools: 'GlycanMiner', 'Score Matrix Tool', 'ProfilePSTMM', 'KCF to LINUCS', and 'KCF to LinearCode'. A 'GO' button is next to the menu. At the top right, there is a 'Download' link. The main title of the page is 'Profile PSTMM input data'.

Fig. 15.23 Snapshot of RINGS after user has logged in and selected a data set. Here, the input data used for the analysis of Siglec 5 data is selected and can be viewed graphically on the right. Options to rerun this data using a tool in the pull-down menu are shown. The data itself can also be downloaded using the “Download” link at the right

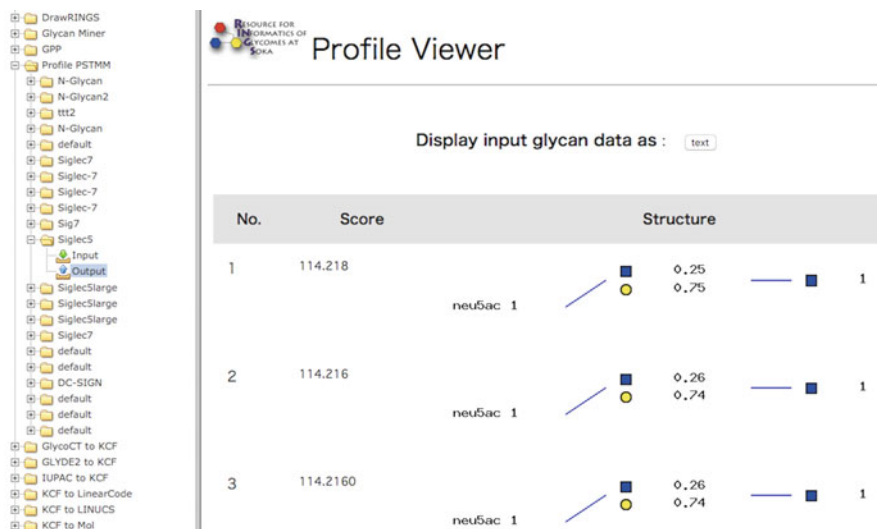


Fig. 15.24 The output view of the analytical results using ProfilePSTMM on the data in Fig. 15.22. The results can also be viewed as text by clicking the button at the top of this page

15.9.2 Data Management System Tips

- Sometimes the data tree does not display after logging in, due to cached cookies and having not logged out a previous time. This can usually be solved by logging in again, and the tree should be displayed.
- It is highly recommended that a data set name be specified whenever running a tool. The default name can be used of course, but managing the data can become difficult as more data is accumulated.
- If for some reason the data tree is not updated with the latest results, the “Reload” button located at the bottom of the data tree can be used to refresh the data tree.

15.10 Feedback System

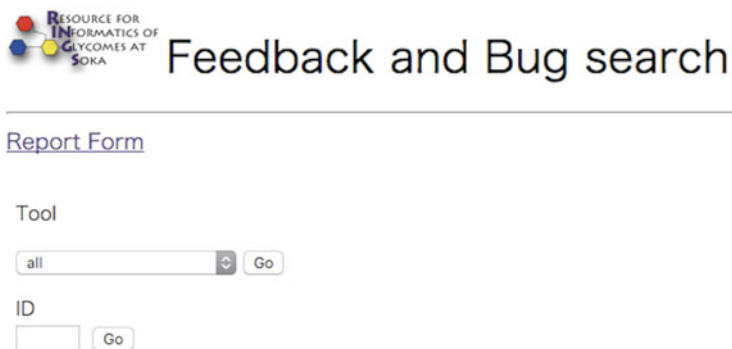
A feedback and bug report system is also available via the “Feedback search” link at the top of the main RINGS page, as shown in Fig. 15.21. Although anyone can view others’ feedback without logging in, users are required to register and log in order to report any new bugs or feedback.

15.10.1 Usage

Figure 15.25 shows the top part of the Feedback Form. The “Report Form” link will open a new form for users to enter feedback after logging in. This form provides three fields: the tool regarding which feedback is being entered, a title to summarize the problem or suggestion, and the details of how the problem occurred or regarding the suggestion. All feedback can be viewed by selecting “all” and clicking the “Go” button using the main Feedback Form (Fig. 15.25). Feedback regarding a specific tool can also be selected in this form. The results of this form show the list of feedback, including Title, Report Data, Status (open, closed, or duplicate), the tool name, the number of votes for the feedback, any comments from developers, and a priority indicator. The bottom of the Feedback Form also shows the latest feedback. Similar to the search results, specific feedback can be seen in detail by clicking the title of the feedback. The feedback detail page, shown in Fig. 15.26, provides users an option to vote on particular feedback. Accumulated votes will then be considered to be prioritized more highly.

15.10.2 Tips

- Voting on existing feedback will help motivate developers to prioritize future work.
- A clear description, including the data used, when encountering problems, is essential to be able to reenact the problem and help the debugging process.



RESOURCE FOR
INFORMATICS OF
GLYCOMES AT
SOKA

Feedback and Bug search

Report Form

Tool

all Go

ID

Go

Fig. 15.25 The top of the Feedback Form where users can report new feedback using the “Report Form” or view previous feedback

Feedback and Bug detail

Feedback or Bug ID	50
Title	Check input consistency
Report Date	2015-04-01 18:08:22
Reporter	Kiyoko Aoki-Kinoshita
Reported Against	0 <input type="button" value="Vote"/>
Priority	pending
Place	GlycanMiner
Feedback	The input data should be checked for inconsistencies, such as wrong number of edges/nodes in the KCF data.
User email	kkiyoko@soka.ac.jp

Fig. 15.26 The detailed page of a particular feedback. Users can vote for certain feedback if they agree and would like it to be prioritized

15.11 Summary

RINGS was developed to provide analytical tools for glycomics analysis, with an easy-to-use interface. However, without user feedback, the interfaces currently provided cannot be improved. Note that we are also under the process of rebuilding the RINGS interface using the latest technologies. Moreover, as mentioned in the Introduction, in the future, RINGS tools will be designed to allow any format to be used as input. We are also considering allowing GlyTouCan IDs to be used as well. Therefore, RINGS will continue to be improved as user demands are met.

Acknowledgements This work was supported in part by Grant-In-Aid for Young Scientists (A); KAKENHI (20016025); the Japan Society for the Promotion of Science (JSPS); the Ministry of Education, Culture, Sports, Science and Technology (MEXT); and the Soka University.

Reference

- Akune Y, Hosoda M, Kaiya S, Shinmachi D, Aoki-Kinoshita K (2010) The RINGS resource for glycome informatics analysis and data mining on the web. *OMICS* 14(4):475–486
- Aoki K, Yamaguchi A, Ueda N, Akutsu T, Mamitsuka H, Goto S, Kanehisa M (2004) KCaM (KEGG carbohydrate matcher): a software tool for analyzing the structures of carbohydrate sugar chains. *Nucleic Acids Res* 32(Web Server issue):W267–W272

- Aoki-Kinoshita K, Ueda N, Mamitsuka H, Kanehisa M (2006) Profilepstm: capturing tree-structure motifs in carbohydrate sugar chains. *Bioinformatics* 22(14):e25–e34
- Aoki-Kinoshita KF, Kanehisa M (2015) Glycomic analysis using KEGG glycan. *Methods Mol Biol* 1273:97–107
- Hashimoto K, Takigawa I, Shiga M, Kanehisa M, Mamitsuka H (2008) Mining significant tree patterns in carbohydrate sugar chains. *Bioinformatics* 24(16):i167–i173
- Jiang H, Aoki-Kinoshita KF, Ching WK (2011) Extracting glycan motifs using a biochemically-weighted kernel. *Bioinformatics* 7(8):405–412
- Krambeck FJ, Bennun SV, Narang S, Choi S, Yarema KJ, Betenbaugh MJ (2009) A mathematical model to derive n-glycan structures and cellular enzyme activities from mass spectrometric data. *Glycobiology* 19(11):1163–1175
- Ranzinger R, Herget S, von der Lieth CW, Frank M (2011) Glycomedb – a unified database for carbohydrate structures. *Nucleic Acids Res* 39(Database issue):D373–D376
- Ueda N, Aoki KF, Mamitsuka H (2004) A general probabilistic framework for mining labeled ordered trees. In: *SIAM international conference on data mining*. SIAM, Philadelphia