

RINGS マニュアル

2012 年 9 月 15 日

目次

第1章 RINGS について	5
第2章 DrawRINGS	9
第3章 GlycanMiner	11
第4章 Glycan Pathway Predictor (GPP)	13
第5章 GlycomeAtlas	17
第6章 Glycan Kernel Tool	23
第7章 MCAW	29
第8章 ProfilePSTMM	33
第9章 Utilities	35
第10章 KCF to Mol	37
第11章 ユーザー及びデータ管理	41

第1章 RINGSについて

RINGS (Resource for Informatics of Glycomes at Soka) とは糖鎖 (構造) 解析のためのリソースである。RINGS では、糖鎖を認識するタンパク質などが糖鎖のどの部分を認識しているかを予測するためのツールや、糖鎖間の共通部分木を探すツールなど、糖鎖解析に用いられる様々なツールが提供されている。主に、質量分析などの糖鎖構造プロファイルデータが手元にある仮定で開発されてきたが、今後はある糖鎖構造情報から関連する遺伝子やタンパク質、修飾部位などの情報も調べられるようなツール開発も計画している。

Welcome to  RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

User registration form Feedback search

Search for display Go [Advanced Search](#) ?

RINGS is a web resource providing algorithmic and data mining tools to aid glycomics research.
Many of the tools have been published in the literature;
RINGS provides free access to these methods for practical research.
Web services are also available via [BioMOBY](#) for programmers interested in incorporating these tools in custom programs.

Tools

- **DrawRINGS**: 2D Java-based glycan structure drawing tool for generating KCF format and/or querying the RINGS database, which current contains the glycan structures from [Glycome-DB](#) and an internally curated data set from the literature.
- **Glycan Miner Tool**: implemented based on [Hashimoto et al., 2008](#) for mining alpha-closed frequent subtrees from a set of glycan structures
- **Glycan Pathway Predictor (GPP) Tool**: implemented based on [Krambeck et al., 2005, which was later improved in Krambeck et al., 2009](#), for dynamically computing the *N*-glycan biosynthesis pathway from a given glycan structure
- **GlycomeAtlas**: Visualization of glycome profiling data on human and mouse tissue samples.
- **Glycan Kernel Tool**: Glycan structure classification tool for finding distinguishing glycan substructures in a target data set compared to a control.
- **MCAW Tool**: glycan structure multiple alignment tool. Align a set of glycan structures and visualize the result as a glycan profile. This tool can be used to find commonalities among a group of glycan structures.
- **Profile PSTMM Tool**: implemented based on [Aoki-Kinoshita et al., 2006](#), generate glycan profiles from glycan structure data which can be entered together with binding affinity data, for example, for a particular biological sample

Utilities

- **GlycoCT to KCF**: convert a glycan structure in [GlycoCT](#) format to KCF
- **GLYDE2 to KCF**: convert a glycan structure in GLYDE2 to KCF
- **IUPAC to KCF**: convert a glycan structure in IUPAC format to KCF
- **KCF to image**: retrieve the image given a KCF
- **KCF to LinearCode**: retrieve the LinearCode format given a KCF
- **KCF to LINUCS**: retrieve the LINUCS format given a KCF
- **KCF to Mol**: retrieve the chemical structure in MOL format given a glycan structure in KCF
- **KEGG GLYCAN ID to KCF**: retrieve the KCF for a given KEGG GLYCAN ID
- **LinearCode to KCF**: retrieve the KCF format given a LinearCode
- **LINUCS to KCF**: convert a glycan structure in LINUCS format to KCF

Documentation

- **Help**: brief users manual for the tools and utilities provided
- **What's new?**: latest updates

Downloads

図 1.1: RINGS のトップ画面。糖鎖解析に用いられる様々なツールや糖鎖構造を表す形式を変換するユーティリティを公開している。

RINGS には糖鎖構造のデータベースも含まれ、GlycomeDB と呼ぶ網羅的な糖鎖構造データ

ベースと、RINGS独自の糖鎖構造も含まれる。RINGSの構造は1996年以降に文献から手動で抽出した糖鎖構造である。また、糖鎖関連の脂質やレクチンの情報も以下のデータベースから参照し、データベースに格納している。

- 脂質

- LipidMaps - <http://www.lipidmaps.org/>
- LipidBank - <http://lipidbank.jp/>

- レクチン

- CNRS の Lectines - <http://www.cermav.cnrs.fr/lectines/>
- Animal Lectins DB - <http://www.imperial.ac.uk/research/animalletins/>

RINGSで糖鎖構造を表示する際、米国の Consortium for Functional Glycomics (CFG)[14] の提案されている単糖の標準シンボルを用いて糖鎖構造を表す。図 1.2 に CFG の標識を示す。

● Galp	● Manp	◇ NeupGc
■ GalpNAc	■ ManpNAc	◆ KDN
■ GalpN	■ ManpN	◇ GlcpA
● Glcp	▲ Fucp	◇ Idop
■ GlcpNAc	☆ Xylp	◇ GalpA
■ GlcpN	◆ NeupAc	◇ ManpA
		□ Not defined

図 1.2: RINGS で使用する各単糖のシンボル。CFG[14] によって推薦されているシンボルである。

RINGSには二種類のプログラムを用意している。図 1.1 の画面で示すように、左側には「ツール」であり、右側に「ユーティリティ」を提供している。ツールは主にデータマイニングの手法や数理的モデルを用いたツールや、糖鎖構造を描画し、検索できるツールである。ユーティリティは、様々な糖鎖構造の表記法を変換できるツールのことで、まだ未完成だが、RINGS以外で使われている表記法をRINGSのツールで解析できるよう、KCFへの変換ツールはほとんどカバーされている。

また、マニュアル以外に、ダウンロードできるソフトとして、糖鎖構造を比較するツール (KCaM) 及び糖転移酵素の発現データから糖鎖構造を予測するプログラム (GlycoXSP) が提供されている。KCaM[2] も GlycoXSP[13] も過去に発表されている手法を用いて開発されている。

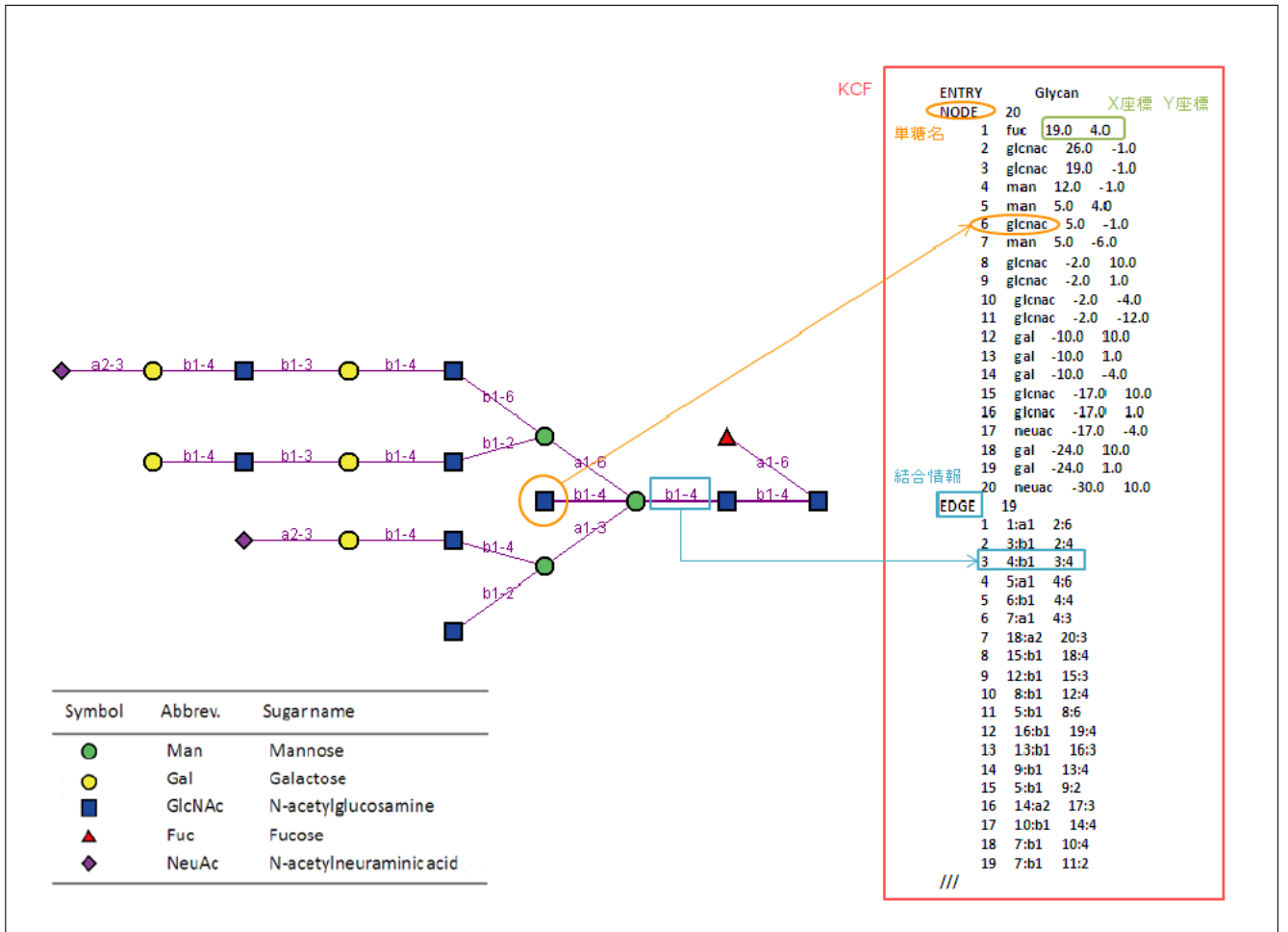


図 1.3: 糖鎖の画像と KCF。糖鎖の単糖情報を NODE で、糖結合情報を EDGE で表示している。NODE の項目は、NODE 番号、単糖名、単糖の x 座標、y 座標を含み、EDGE の項目は、EDGE 番号、結合している NODE 番号、配座情報を含んでいる。EDGE での、結合している NODE 番号と配座情報は「:」で区切って表示されている。

第2章 DrawRINGS

DrawRINGS はユーザーが自由に糖鎖構造を描画でき、その構造に似た構造をデータベースから検索することができるツールである。データベースは、GlycomeDB、RINGS の一方または両方を選択できる。

利用目的

DrawRINGS は主に糖鎖構造の描画、KCF からの編集、そして検索に用いられる。また、検索結果の糖鎖構造の詳細からさらなる関連した情報への探索も可能である。

利用方法

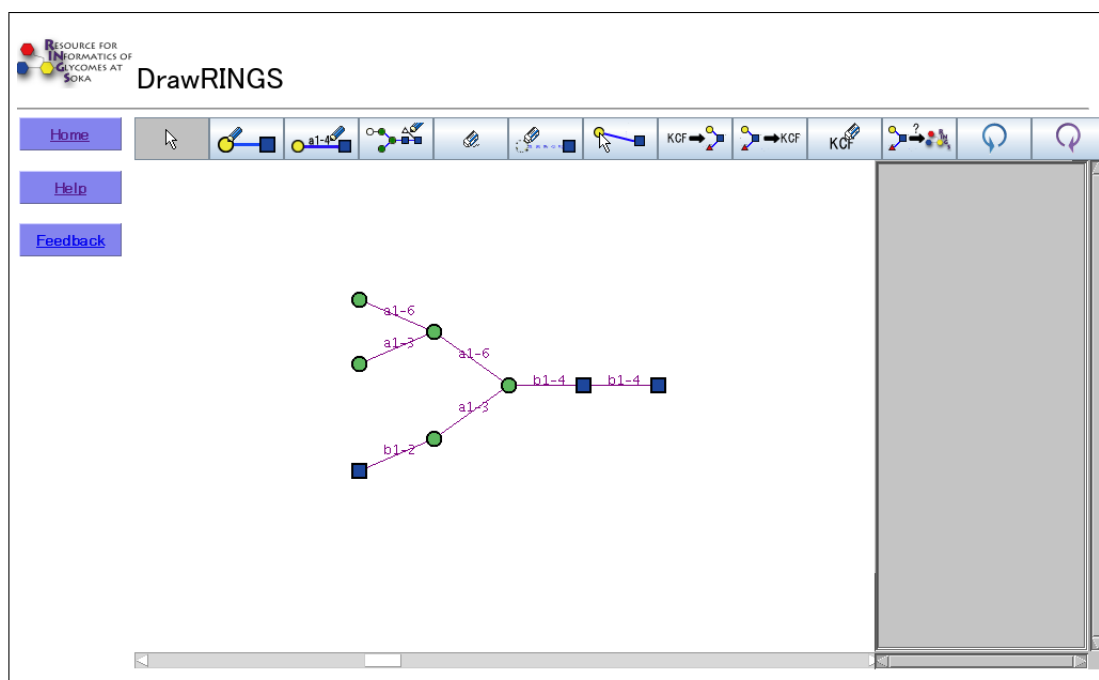


図 2.1: DrawRINGS のトップ画面。キャンバス部分に、各種ボタンを用いて糖鎖構造を描画することができる。テキストエリアでは KCF 形式のテキストの入出力が可能である。また、描画した構造を検索することが可能である。

DrawRINGS の各種ボタンの機能は次の通りである。

- Draw Node - 単糖を描画。キャンバス部分のクリックした箇所に、代表的な単糖が表示され、そのリストから選択した単糖が描画される。その他の単糖は「Other」を選択し、テキストを入力することで表示できる。

- Draw Edge - 糖結合を描画。キャンバスに描画された二つの単糖を線で結び、再度線上をクリックすると、一般的な糖結合の配座のリストが表示される。その他の配座は「Other」を選択し、テキストを入力することで表示できる。
- Core structure - N型糖鎖のコア構造や、O型糖鎖、グリコサミノグリカン、そしてLewis構造などの、よく見られる糖鎖構造が用意されている。このボタンをクリックし、キャンバス内をクリックすると、これらのコア構造メニューが表示され、クリックした場所に表示される。
- Clear All - キャンバス内の構造を全て削除する。
- Erase Node - 単糖を削除。単糖をクリックすると、削除される。
- Move Node - 単糖を移動。単糖をドラッグすると、移動することができる。
- Draw KCF - 右側のテキストエリアに KCF 形式の情報があれば、その構造をキャンバス内に表示する。
- KCF Text out - キャンバス内に表示している構造を右側のテキストエリアに KCF 形式で表示する。
- Text Clear - 右側のテキストエリアの情報を削除する。
- Run query - 描画した構造をクエリとして、RINGS のデータベースを検索し、類似した順に糖鎖構造を図 2.2 のように、別ウィンドウで表示する。
- Undo - 最後の動作を取り消す。
- Redo - 最後に取り消した動作を再度実行する。

RESOURCE FOR
INFORMATICS OF
GLYCOMES AT
SOKA

DrawRINGS Result

NEXT 1 2 3 4 5 / 105 Go Text file

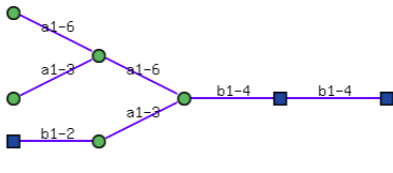
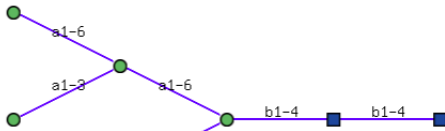
No.	Score	Glycan id	Reaction id	Structure ?
1	100.0	G01101		
2	98.6	G02252		

図 2.2: DrawRINGS からクエリ検索した画面。検索した構造のスコアと ID、構造の画像が表示されている。ID か画像をクリックすることでその構造の詳細な情報を閲覧できる。

第3章 GlycanMiner

GlycanMiner とは、大量の糖鎖構造の中から「 α -closed frequent subtree」[8] と呼ぶ、頻繁に出現する糖鎖の部分構造を抽出するツールである。

α -closed frequent subtree を説明するために、まず「frequent subtree」を定義する。直訳すると、「頻繁に出現する部分木」のことであり、大量の木構造の中に、ある「特定の数」分、出現する部分構造のことである。この「特定の数」は通常「minsup」と呼び、minimum support の略である。support とは、部分構造が現れる木構造の数である。つまり、support は部分構造を含む木構造の数であり、最初から指定する必要がある。

しかし、frequent subtree を抽出するだけでは、非常に似た部分構造 (frequent subtree) がたくさん出力されることになる。これらの部分構造を区別し、「十分に異なる部分構造」を抽出するため、 α -closed frequent subtree と呼ぶ概念が考案された。

では、 α -closed frequent subtree とは、以下の式を満たす部分構造 T のことである。

$$support(P) < \max(\alpha * support(T), minsup)$$

P は T を含む木の部分構造を指す。つまり、 P も T も入力の木構造に含まれる部分構造であり、しかも T は P の部分構造でもある。それぞれの support 値を比較し、 P の support 値が T の support 値の α パーセント以下であれば、 T は α -closed frequent subtree である。ここで α というパラメータを指定する必要があり、0 ~ 1 の値を持つ。

利用目的

GlycanMiner は、糖鎖構造が大量に生成された場合に、共通かつ有意に見つかる糖鎖の部分構造を探すために用いられる。例えば、糖鎖アレイの場合、コントロールとターゲットとそれぞれに強く結合する糖鎖構造群が出力される。そうすると、コントロールとターゲットの糖鎖構造群を個別に実行し、 α -closed frequent subtree の結果を比較できる。一方、糖鎖の質量分析で得られたたくさんの糖鎖構造情報を解析し、特異的に現れる糖鎖の部分構造の分析にも用いることができる。

利用方法

1. ユーザーは、入力画面で KCF 形式の糖鎖構造情報を入力またはファイルからロードする。
2. α に 0 ~ 1 までの数字を、minsup support に数字をそれぞれ入力し、Go mine ボタンを押すと解析結果が表示される。

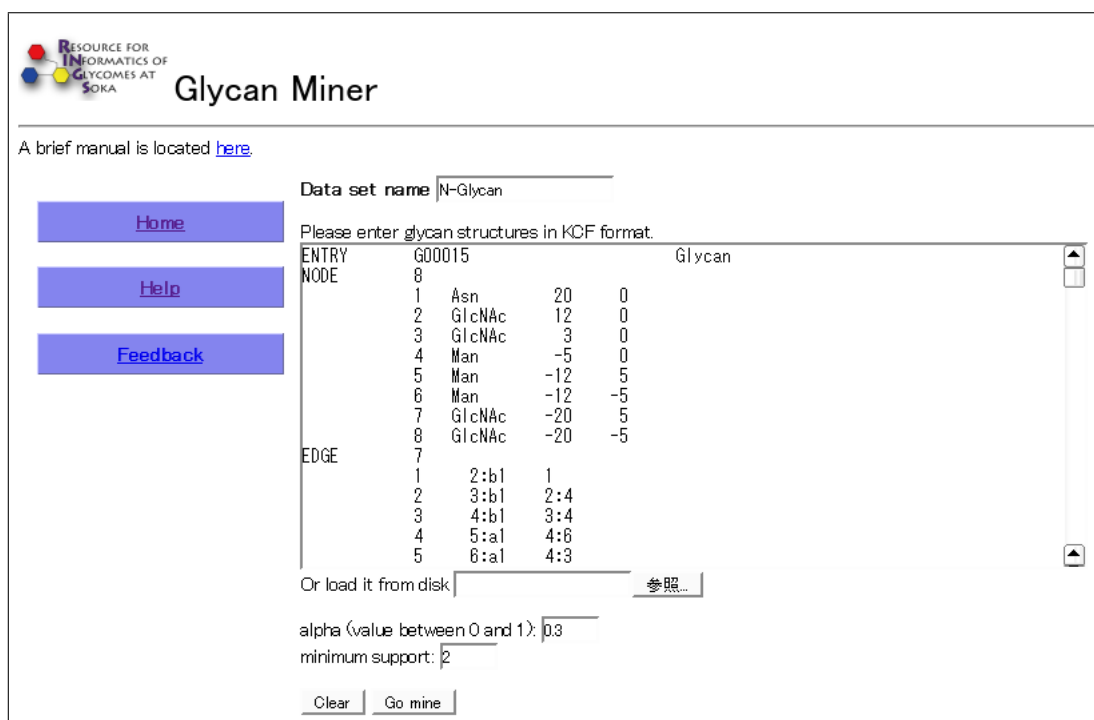


図 3.1: GlycanMiner の入力画面。KCF 形式の糖鎖構造情報を入力またはファイルからロードする。次に alpha に 0~1 まで数字を、minimum support に数字をそれぞれ入力し Go mine ボタンを押すと解析結果が表示される。

No.	P-value	Name	Support	Subtree
1	0.000000e+00	ST00000	3	
2	4.288171e-46	ST00003	2	

図 3.2: GlycanMiner の結果画面。P-value、部分構造に自動的に付けられた名前、Support、部分構造の画像が表示される。

第4章 Glycan Pathway Predictor (GPP)

GPP とは糖転移酵素による N 型糖鎖の生合成を予測するためのツールである。基質特異性は数理的モデル [11, 10] で表し、全経路を素早く計算できる。しかし、この数理的モデルの計算は無限に続く可能性があるため、それを制限するために糖鎖の分子量の最大値をユーザーからの入力によって設定する。設定した分子量より大きい糖鎖構造は生成しないようになる。また本モデルは、濃度や局在情報は考慮していないため、論理的な予測を行う。今後、改良する予定である。

利用目的

教育の面において、GPP を用いて理論的に可能な糖鎖構造の全体像がまず見える。また、ノックアウト実験などの予測のためにもパスウェイの変化を調べることも可能である。

利用方法

1. ユーザーは、入力画面で KCF 形式の糖鎖構造情報を入力またはファイルからロードし、酵素の選択と上限となる分子量の最大値を設定して Submit ボタンをクリックする。
2. 結果画面の上部に入力した糖鎖の図と酵素の情報を表示する。そして入力した酵素に対して選択した酵素が反応しない場合、"No reaction." という文字を表示する。反応が行われる場合は Java アプレットを用いて反応経路の全体図が表示される。この図は拡大縮小、検索などの機能が利用できる。
3. 各糖鎖には、ハイパーリンクが埋め込んであり一反応の結果を表示することができる。

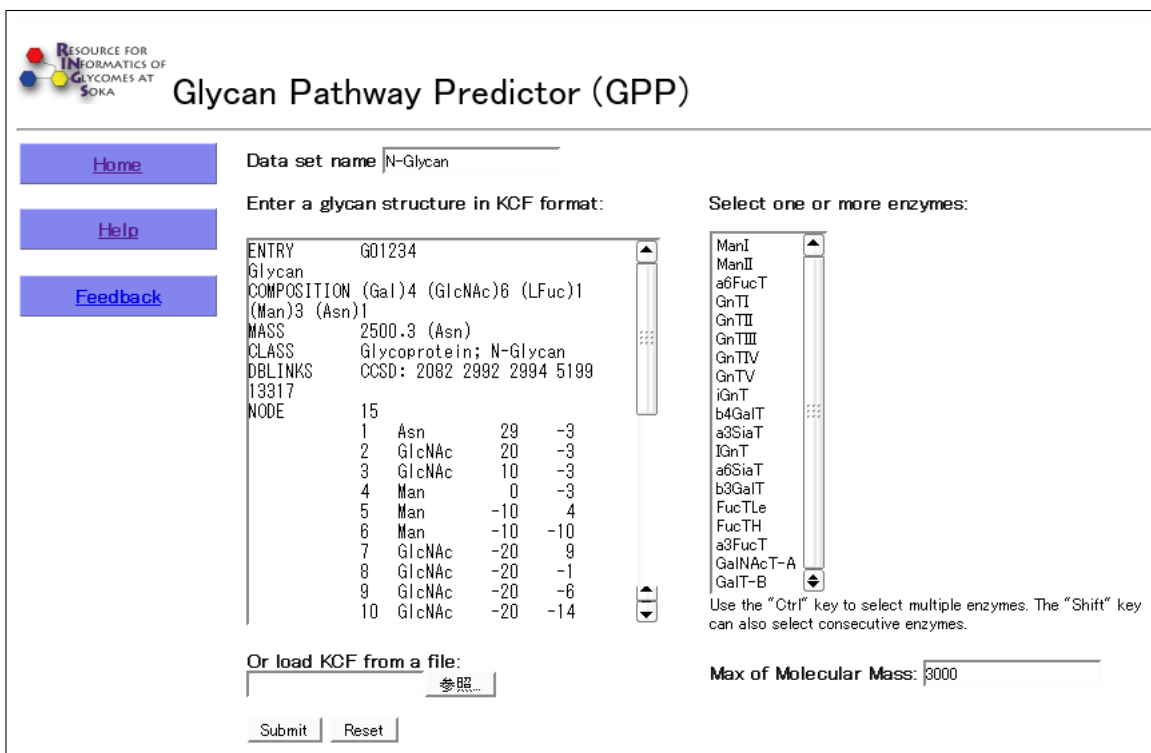


図 4.1: GPP の入力画面。ユーザーは糖鎖構造の情報を KCF 形式で入力またはファイルをロードし、酵素を選択する。また、計算に必要な糖鎖の分子量の最大値を設定して submit をクリックすることで結果画面を表示させる。

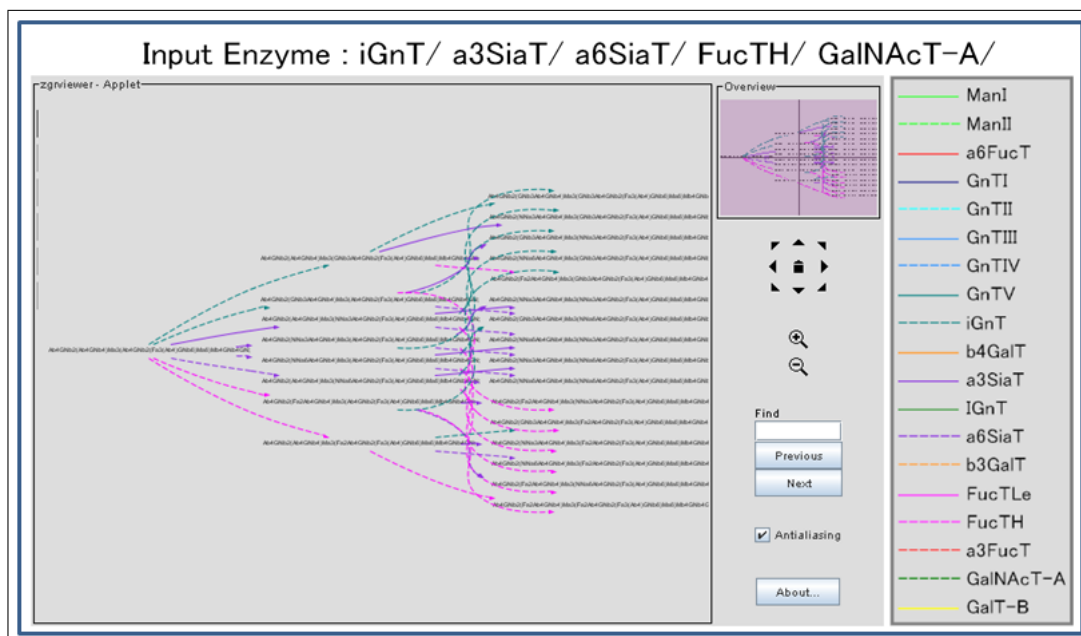


図 4.2: 計算された反応経路を示す結果画面。グラフ中の線は各酵素を表している。各糖鎖構造はリアコードで示され、図 4.3 のような一反応の結果画面にリンクしている。

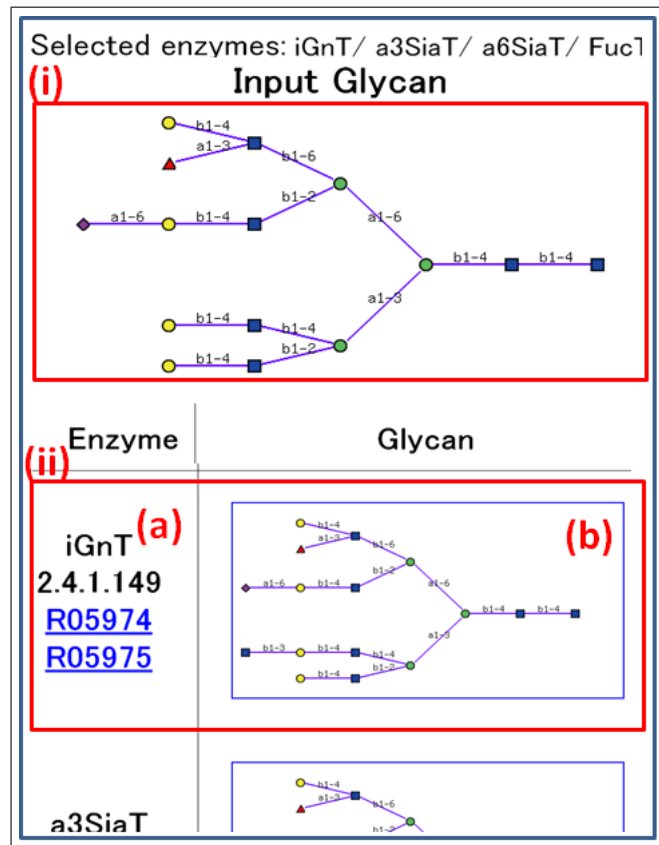


図 4.3: 一反応の結果を表示する画面。(i) ではクリックした糖鎖構造を画像で表示し、(ii) でその糖鎖に対する一段階の反応を表示している。(ii) の (a) には酵素名とその酵素の EC 番号及びリアクション ID が表示され、(b) にはその酵素の反応によって生成される糖鎖の構造が画像で表示されている。

第5章 GlycomeAtlas

GlycomeAtlas は CFG が提供されている糖鎖プロファイリングデータの可視化ツールである。CFG は ES-MS/MS や MALDI-TOF MS を用いてヒトとマウスの様々な組織の糖鎖構造を同定しているため、これらの情報を見やすく表示するツールである。糖鎖構造の検索も可能で、さらに個人的に収集された糖鎖プロファイリングデータの可視化も可能である。(しかし、現時点においてはヒトとマウスのみに対応している)

利用目的

糖鎖構造の分散や様々な組織の中から検索するためのツールであり、糖鎖プロファイリングの解析のために利用できる。

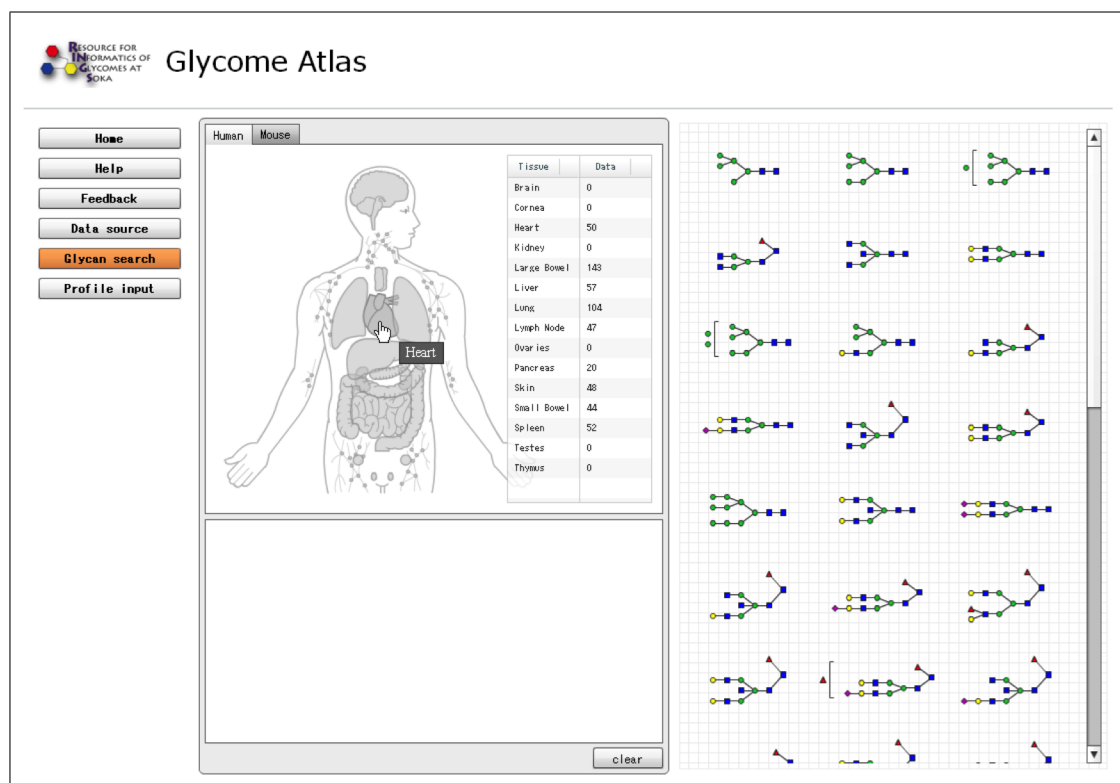


図 5.1: GlycomeAtlas の初期画面。心臓をクリックすると、ヒトの心臓において同定された糖鎖構造が右パネルに表示される。

利用方法

初期画面の操作 GlycomeAtlas の初期画面は図 7.3 のように表示される。上部の“Human”や“Mouse”のタブでそれぞれの生物種の画面に切り替えることができる。この画面で例えば、ヒトの心臓をクリックすると、糖鎖構造の一覧が右パネルに表示される。また、右パネルの糖鎖構造をクリックすると、図 7.1 のようにその糖鎖構造の拡大図が左下のパネルに追加され、そして存在する組織が強調表示される。なお、複数の糖鎖構造も次々と左下のパネルに追加することもできる。そのとき、既にクリックされている構造かを確認するため、重複して選択することはできないようになっている。

Tissue	Data
Brain	0
Cornea	0
Heart	50
Kidney	0
Large Bowel	143
Liver	57
Lung	104
Lymph Node	47
Ovaries	0
Pancreas	20
Skin	48
Small Bowel	44
Spleen	52
Testes	0
Thyruus	0

図 5.2: 右パネルの一覧の中の糖鎖構造をクリックすると、左下のパネルに追加され、存在する組織が強調表示される。

左下パネル：選択された糖鎖構造 左下パネルに表示されている糖鎖構造をクリックすると、図 5.3 のように、組織に加え、Data List も強調表示される。

生物種の切り替え 図 5.4 のように、糖鎖構造を選択した状態で他の生物種のタブを表示すると、その生物種の組織も強調表示された状態で表示される。

糖鎖構造の検索 LinearCode 形式 [7] で糖鎖構造を入力すると、その構造を検索することができる。図 5.5 が示すように、LinearCode 形式で糖鎖構造が入力された場合、その構造を確認するために糖鎖構造の図が表示される。

糖鎖構造の検索結果は図 5.6 のようになる。入力された糖鎖構造が左下の画面に追加され、クリックすると見つかった組織が強調表示される。

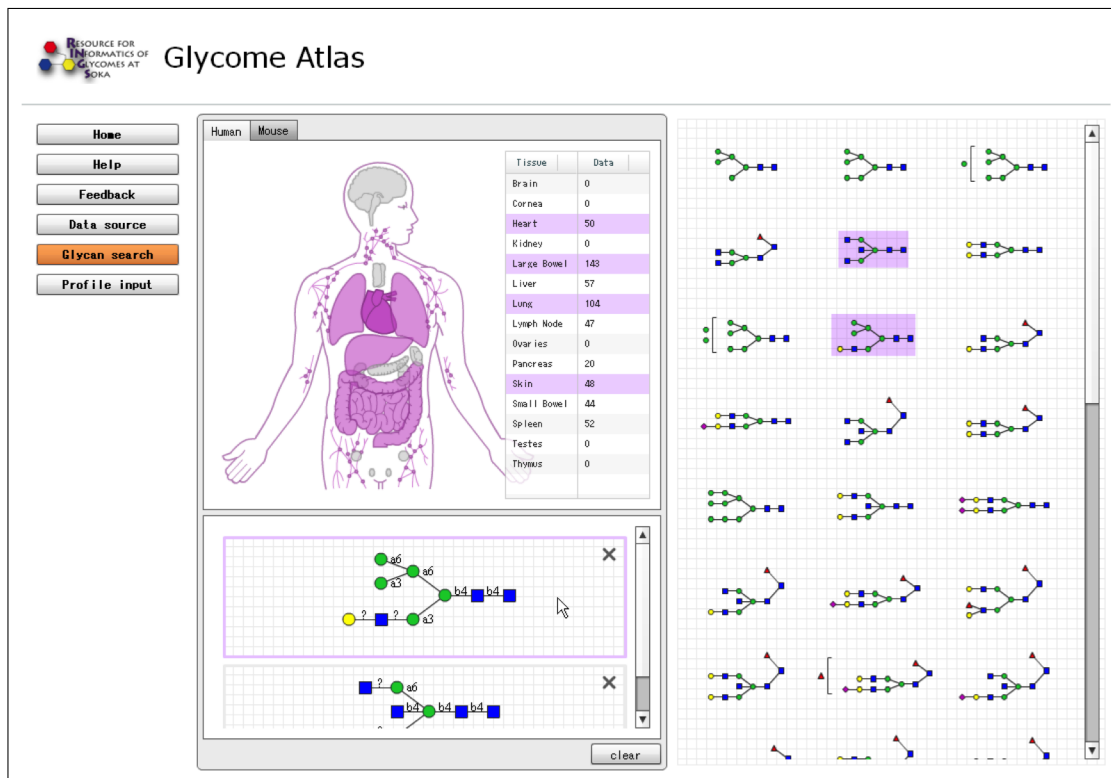


図 5.3: 糖鎖構造をクリックする度に、左下のパネルに追加される。なお、同じ構造が追加されないようになっている。

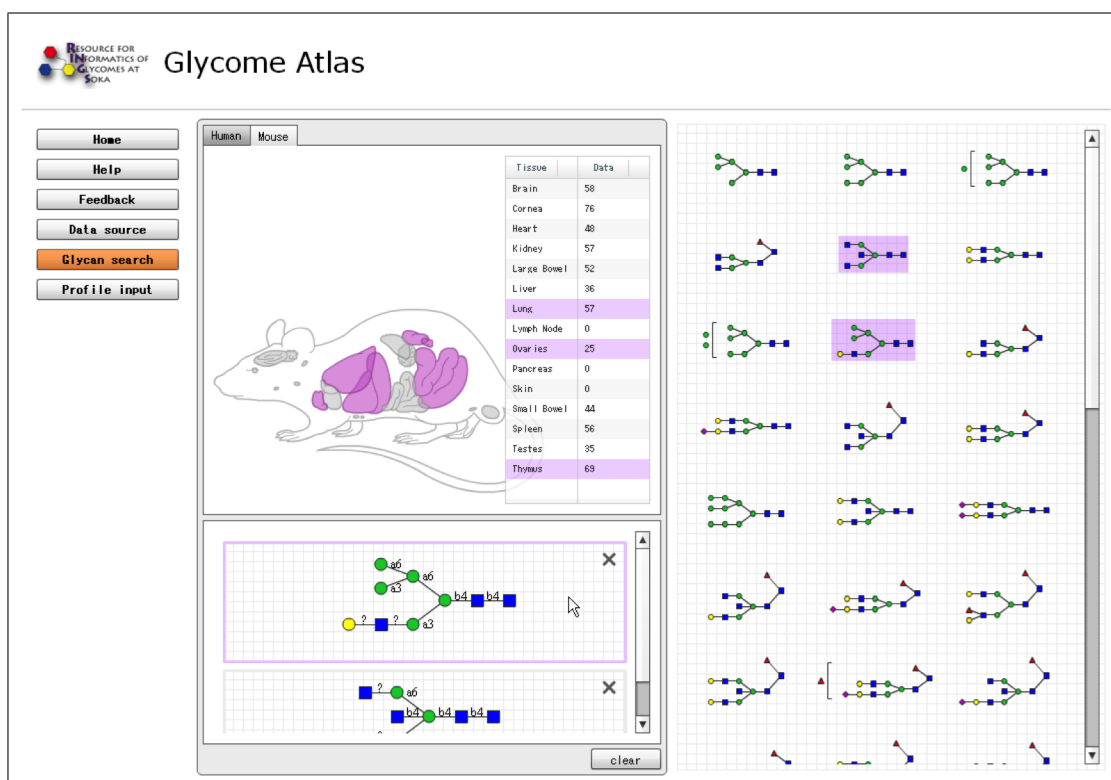


図 5.4: マウスを表示すると、選択された糖鎖構造がヒト同様に強調表示された状態で表示される。

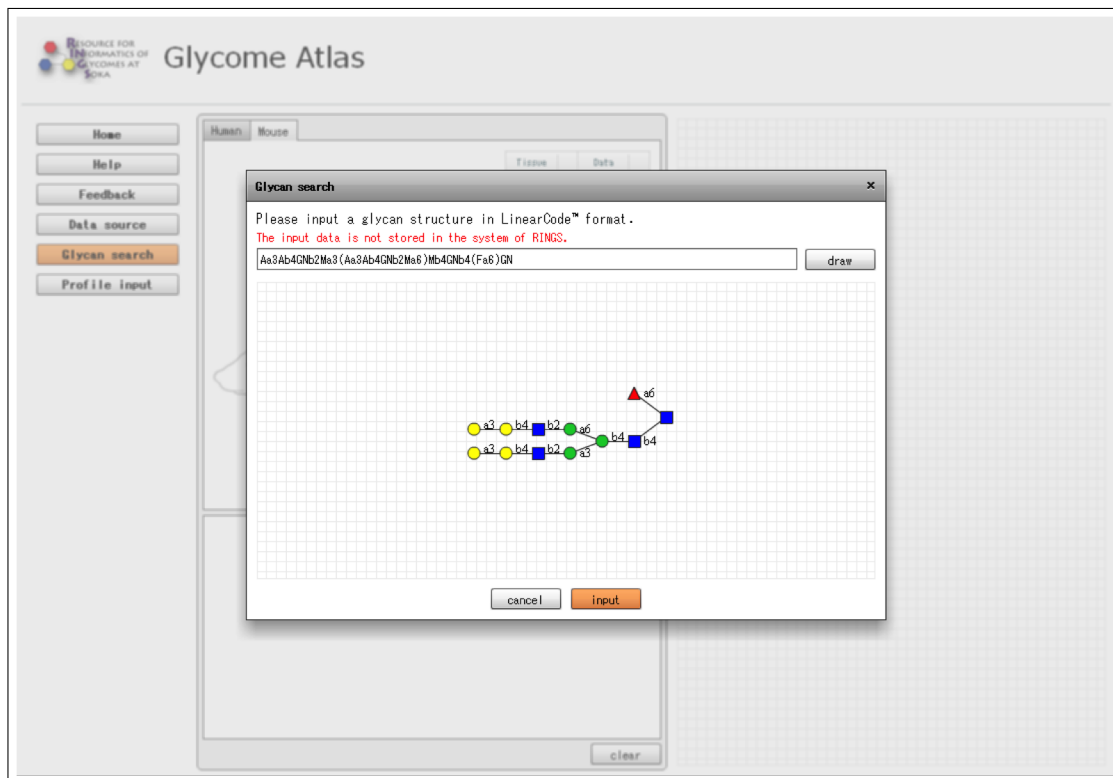


図 5.5: 糖鎖構造を検索する場合、LinearCode 形式で構造を指定し、その下に入力された糖鎖構造の図が表示される。

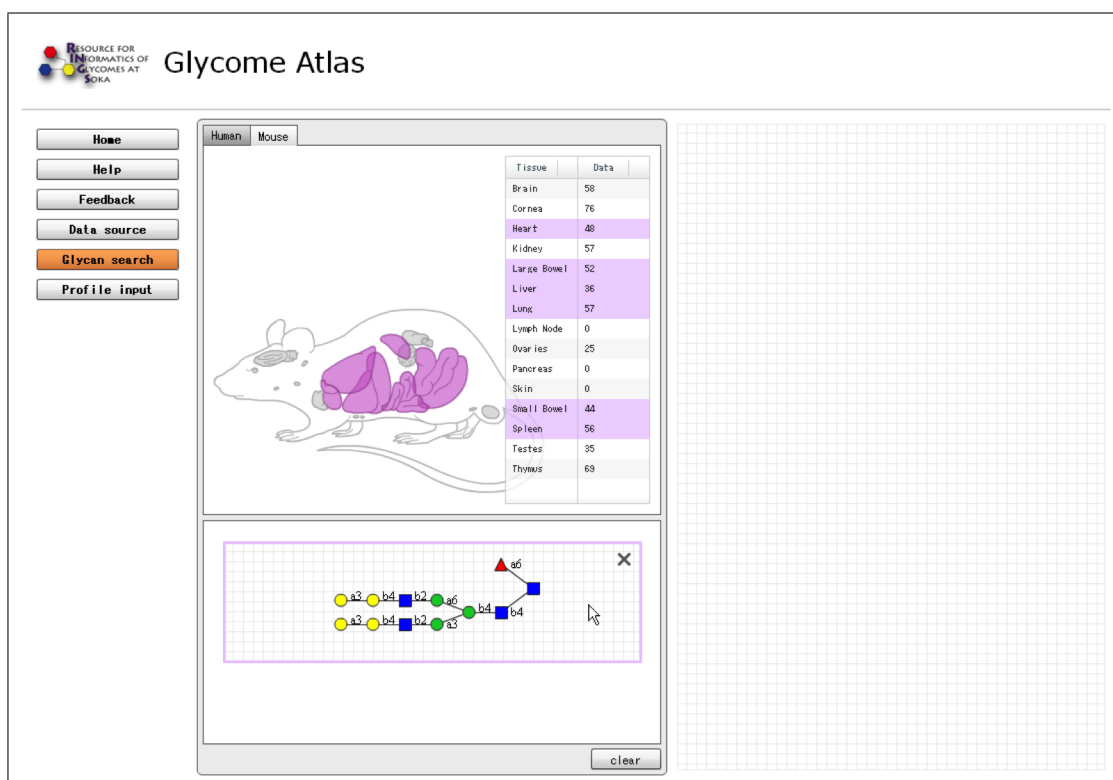


図 5.6: 糖鎖構造を検索した結果画面。問合せ構造が左下のパネルに追加され、クリックすると存在する組織が強調表示される。

個人データの可視化 CFG のデータ以外に、同様な組織別糖鎖プロフィールデータを表示するための機能も利用できる。左側のメニューの「Profile Input」ボタンをクリックすると、図 5.7 画面が表示される。サンプルファイルへのリンクもあるので、参考にしてファイルを作成できる。「Browse...」ボタンでローカルコンピュータからファイルを選択すると、下の表に読み取れた糖鎖構造の一覧が表示される。構造検索同様、LinearCode 形式 [7] で糖鎖構造を指定してください。

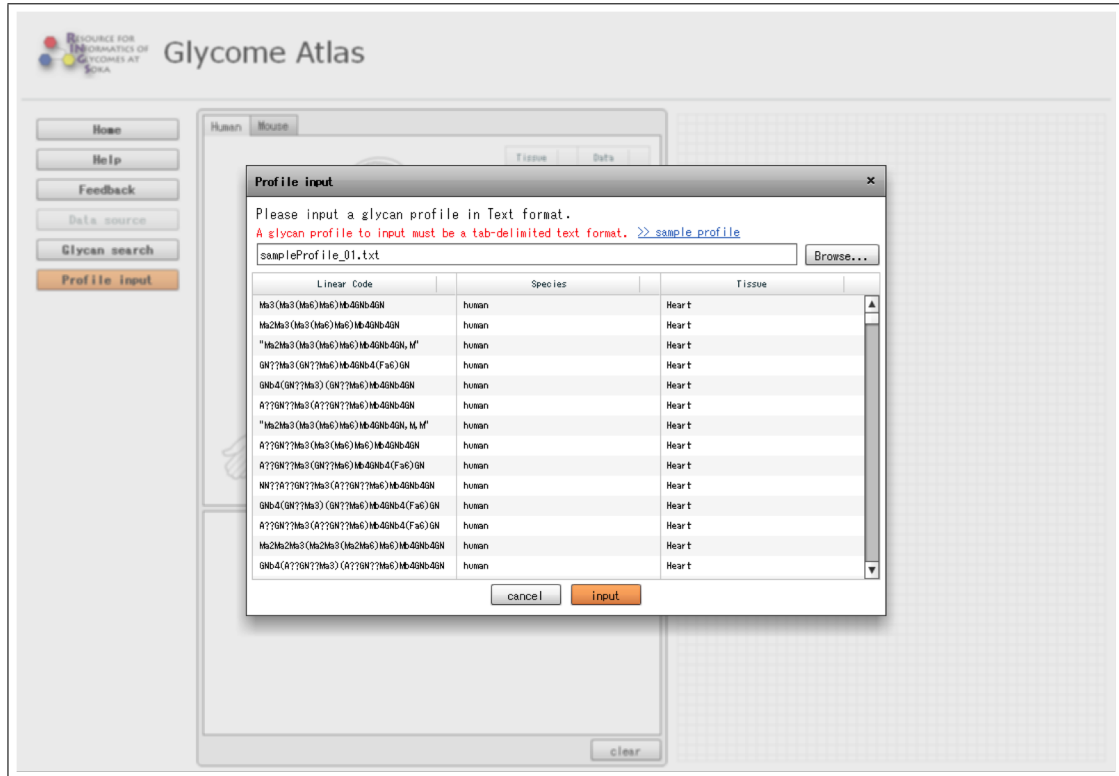


図 5.7: 糖鎖の組織別プロファイリングデータをファイルから読み取って表示することができる。この図はそのファイルを指定する画面である。

なお、読み取ったファイルに問題があった場合は、図 5.8 のような画面が表示される。そして問題になった行が通知される。

ファイルが問題なく読み取れた場合は、図 7.3 に似た画面が表示されるが、各組織に存在する糖鎖構造の数はファイルの内容を反映する。

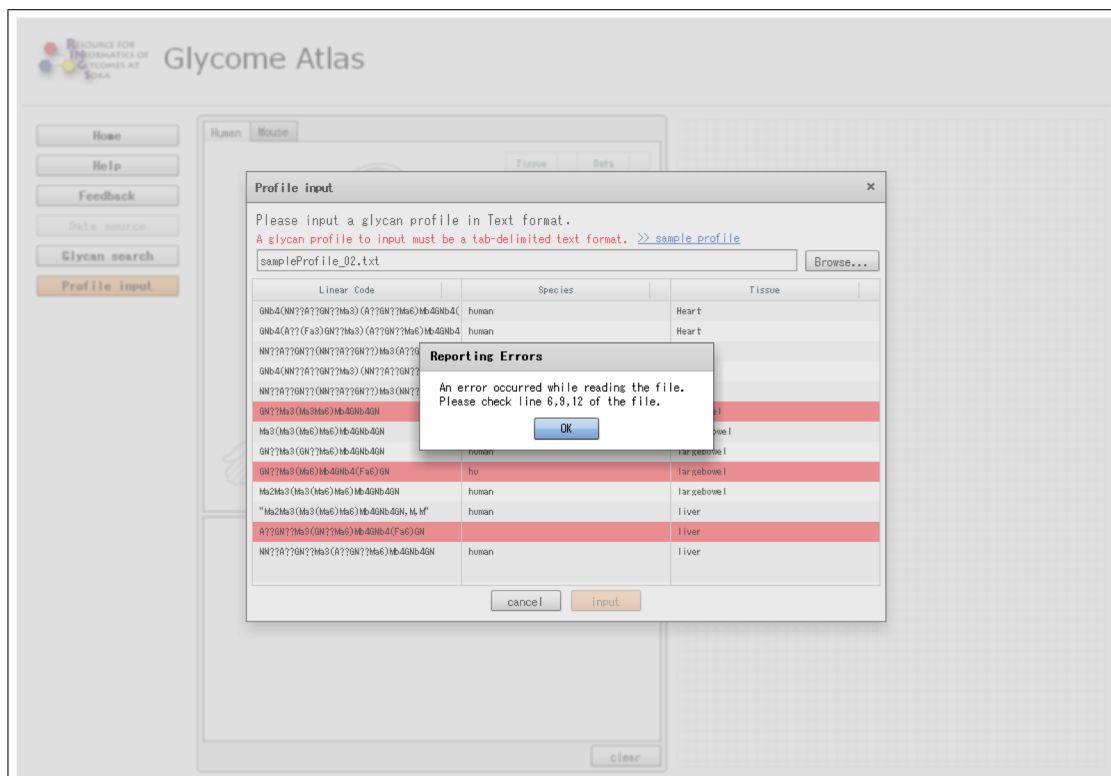


図 5.8: プロファイルデータのファイル形式に問題があって読み取れなかった場合、この図のように問題となった行が表示される。

第6章 Glycan Kernel Tool

質量分析 (MS) や MS/MS 解析を通じてサンプルから得た多量の糖鎖構造を、視覚的判断のみによって構造比較やプロファイルリングすることは困難である。例えば、コントロール及びターゲットサンプルから得たデータの中から、特徴的な部分糖鎖構造を抽出する作業には研究者の経験や技術、時間が多分に必要とされる。Glycan Kernel Tool では、カーネル分類法を応用した糖鎖構造解析アルゴリズムを用いることで、2種類の糖鎖構造データから短時間で特徴的な糖鎖部分構造を予測することができる。

Glycan Kernel Tool は、入力された2種類のデータセットを比較分類し、特徴的な糖鎖部分構造を予測するツールである。例えば、質量分析 (MS) や MS/MS 解析を通じてコントロールサンプル及びターゲットサンプルから得た糖鎖構造郡の中からどちらかのサンプルにのみ特異的に存在する糖鎖部分構造を予測することができる。

利用目的

2種類の糖鎖データを比較分類して、特徴的な糖鎖部分構造を抽出したい場合に利用することができる。

RESOURCE FOR INFORMATICS OF GLYCOMES AT SOHA

Glycan Kernel Tool

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

Data set name

Glycan Data
Enter KCF data consecutively.

Target data

```
ENTRY 1024_3_1Glycan
NODE 4
1 gal
-8.0 -2.0
2 glcnac
-8.0 2.0
3 galnac
0.0 -0.0
4 glcn
-14.0 2.0
EDGE 3
1 1:b1
3:3 2 2:b1
3:6
```

Control data

```
ENTRY 1099_6 Glycan
NODE 4
1 gal
-17.0 9.0
2 glcnac
-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
```

Or load KCF from a file:

Or load KCF from a file:

option

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

min <= size <= max

As an estimate of computation time, this table lists the minimum and maximum data sizes allowed and their corresponding computation times.

	data size	computation time
minimum	10 entries per data set	1 minute
maximum	400 entries per data set	4 hours

図 6.1: Glycan Kernel Tool の入力画面 Data set Name: ユーザは任意のデータ名を入力することができる。RINGS へユーザ登録している場合には、ユーザはこのデータ名から入力及び結果情報を参照することができる。Glycan Data: ターゲット及びコントロールデータとして、KCF 形式で記述された糖鎖構造情報を直接書き込むことができる。または、テキストファイルを使用することも可能である。subtree size: 糖鎖構造を部分構造へと分解する際のサイズ (単糖数) を指定することができる。部分構造のサイズには 1 から 9 までを指定できる。また、結果はここで指定されたサイズにもとづいて出力される。

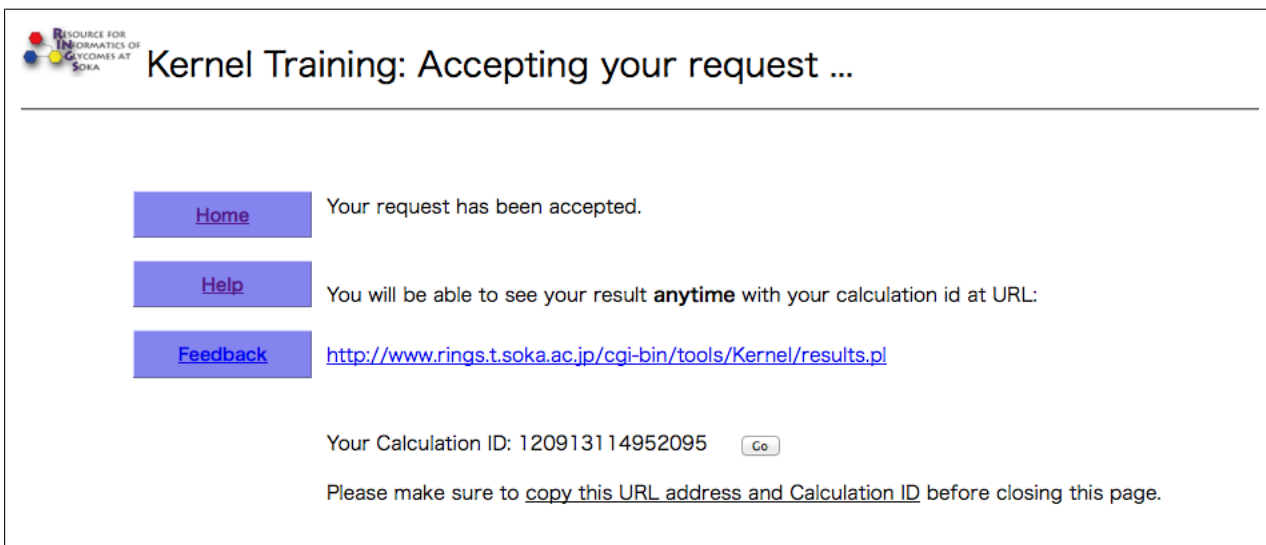


図 6.2: Calculation ID の付与入力画面の run ボタンをクリックした後、この画面が表示され Calculation ID が表示される。入力した糖鎖構造数が 200 個を超える場合、計算に 10 分以上必要となる。そのため、この ID を使用して計算結果を別途表示させる。

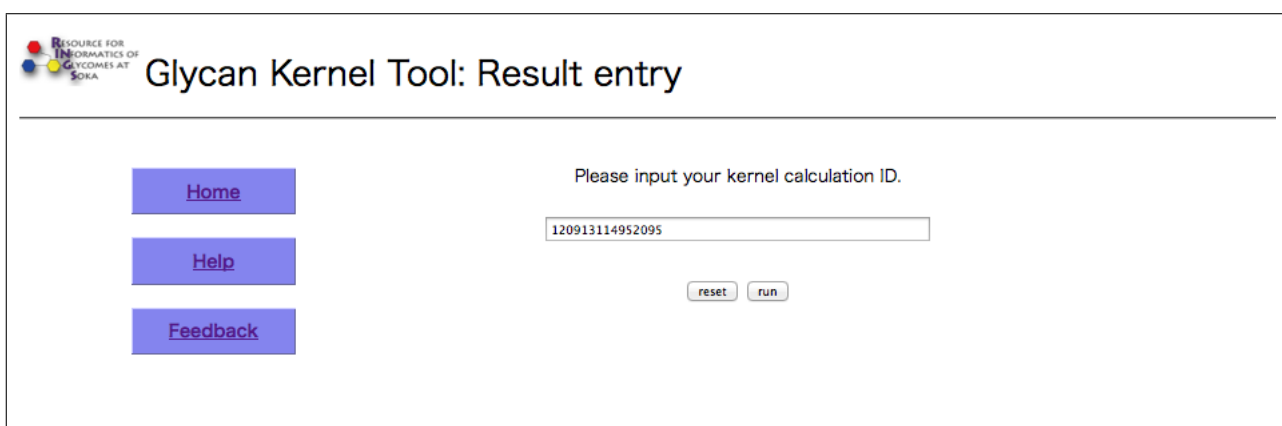


図 6.3: Calculation ID の入力画面先ほど得た Calculation ID を入力する。もし、ユーザが異なる Calculation ID を持っている場合、この画面からその ID を入力することもできる。

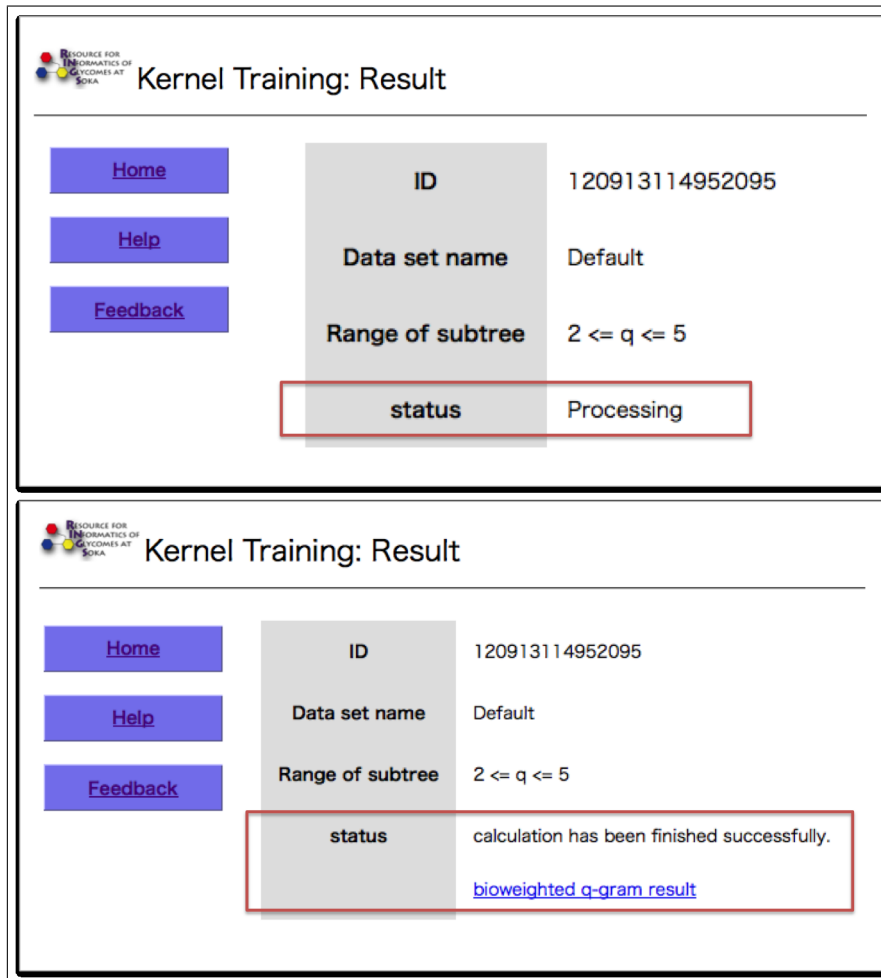


図 6.4: 計算処理状況の確認画面この画面では、ユーザは計算処理の状況を確認することができる。status が "processing" であれば、計算がまだ終わっていないことを示す (上図)。また、計算処理を終えていれば、status は "calculation has been finished successfully" となり、結果画面へのリンクが表示される (下図)。

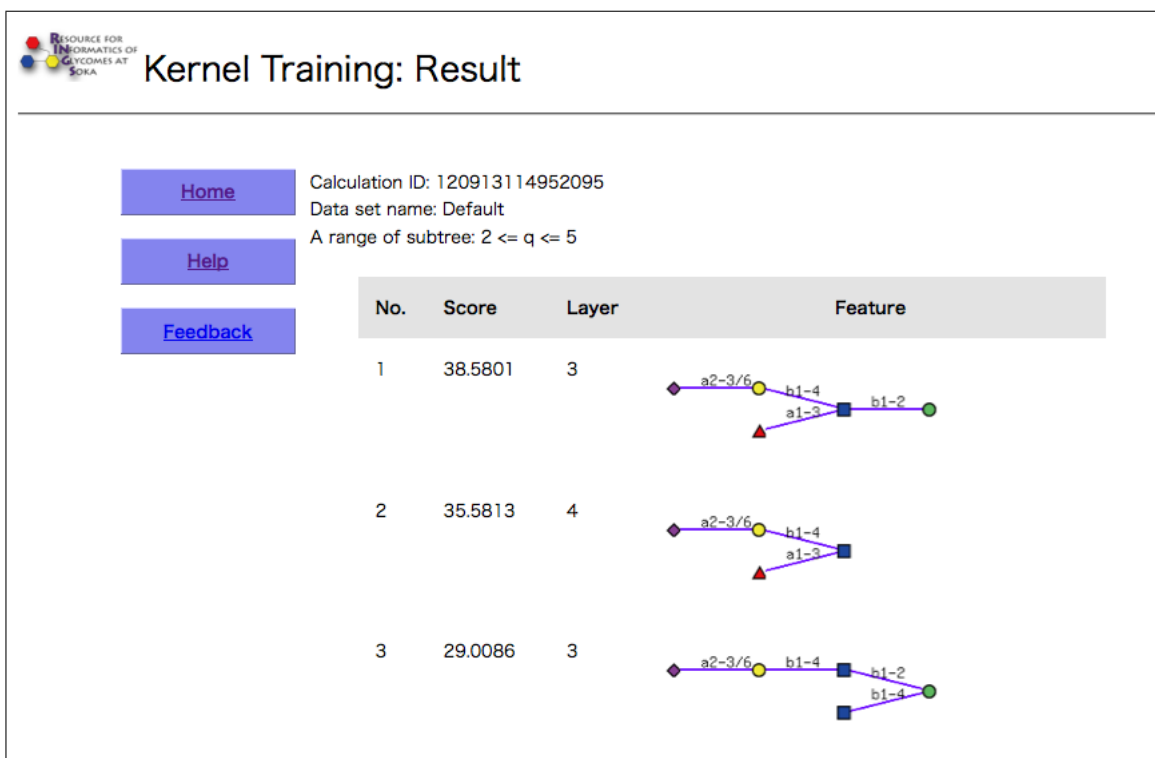


図 6.5: 結果画面画面上部には、Calculation ID、Data set name、糖鎖部分構造のサイズが表示される。その下に、本ツールの計算結果がスコアの高い順に表示される。結果には、順位、スコア、layer 及び糖鎖部分構造が出力される。スコアを比較することで、その構造が結果群の中でどの程度「特徴的糖鎖部分構造」であるか、という目安として使用できる。計算上、入力データ中の糖鎖構造が部分構造に分解されるとき、還元末端を「1」とした時の還元末端からの深さ (layer) がその部分構造情報中に付与される。

第7章 MCAW

MCAW は複数の糖鎖構造から共通する単糖・結合を揃えて整列させた構造を可視化するツールである。複数のアミノ酸配列から配列の保存された領域や類似パターンを比較する際によく使用されている ClustalW でアミノ酸配列の整列 (アラインメント) が行われているように、複数の糖鎖構造をアラインメントする。アラインメントした結果を糖鎖構造プロファイルとして、共通する部分構造の割合を表して構造の特徴を可視化する。これは、複数のアミノ酸配列のアラインメント結果に共通にみられる高度な保存領域を同定するアミノ酸配列プロファイルを参考にしている。

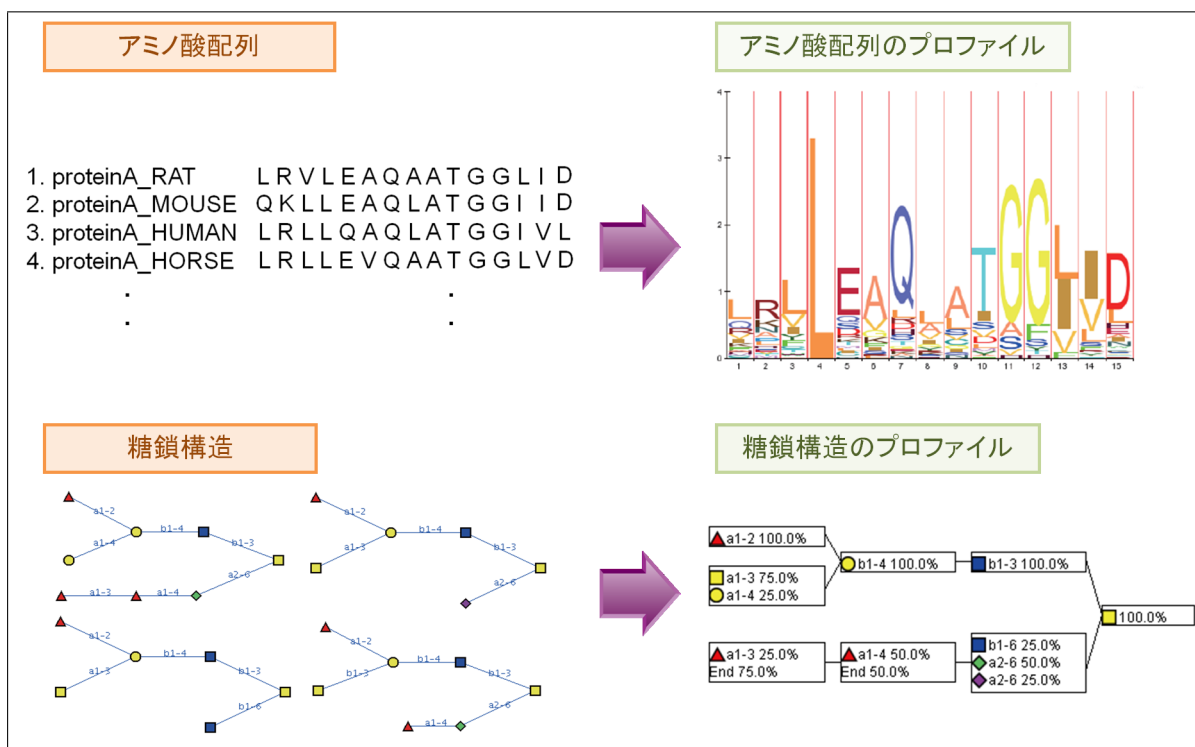


図 7.1: アラインメント結果をプロファイルで表した図

利用目的

複数の糖鎖構造が共通する構造パターンを見つけ出すためのツールであり、共通するパターンの割合をみる事ができる。糖鎖アレイデータからレクチンが認識する糖鎖の構造パターンを検索するのに用いたことがある。

利用方法

1. ユーザーは、入力画面で KCF 形式の糖鎖構造情報を複数入力またはファイルからロードし、"Advanced weighting options" のアライメント時のオプションスコアを設定して Submit ボタンをクリックする。

RESOURCE FOR
INFORMATICS OF
GLYCOMES AT
SOKA

MCAW (Multiple Carbohydrate Alignment with Weights)

Home | Help | Feedback

Data set name: default

Enter a glycan structure in KCF format:

```
ENTRY G04845 Glycan
COMPOSITION (Gal)3 (Glc)1 (GlcNAc)2 (LFuc)2 (Neu5Ac)1
MASS 1658.5
DBLINKS CCSID: 23349
GlycomeDB: 20420
JCGGDB: JCGG-STR011245
NODE
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

糖鎖構造の入力

複数の糖鎖構造をKCFで入力

入力はファイルを選択することも可能

ギャップが入る時のペナルティスコア

Advanced weighting options

単糖が一致したときのスコア

アノマーが一致したときのスコア

非還元末端側の炭素番号が一致したときのスコア

還元末端側の炭素番号が一致したときのスコア

実行

図 7.2: MCAWtool の入力画面。KCF 形式で糖鎖構造を複数入力し整列させる際のスコア計算を調節できる。大きいスコアを入力するとその項目が一致しやすくなる。

2. 結果画面には、入力した複数の糖鎖構造のアライメント結果を表示する。整列された単糖と結合の情報が、入力構造中でどのくらい一致したのかを表している。割合が高ければ入力構造中に共通して見られる構造パターンだということがわかる。
3. 同じ構造を入力してもオプションスコアの値を変えることによってアライメント結果が変わる。例えば、アノマーのオプションスコアを他のスコアより大幅に高く設定するとアノマーが揃うようなアライメント結果が得られる。

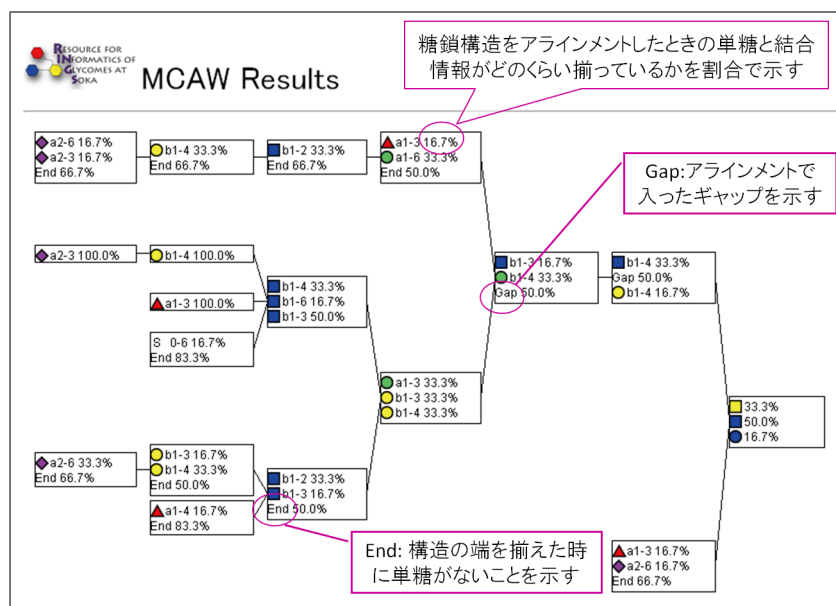


図 7.3: MCAWtool の結果画面。糖鎖構造をアラインメントした結果を糖鎖構造プロファイルとして表示している。揃えられた単糖と結合がどのくらい一致しているのかを割合で示している。結果に表示されている”Gap”はアラインメントで入ったギャップを示し、”End”は構造の橋を揃えたときに単糖がないことを示す。単糖は CFG で提案されている単糖の標準シンボルを用いて表している。

第8章 ProfilePSTMM

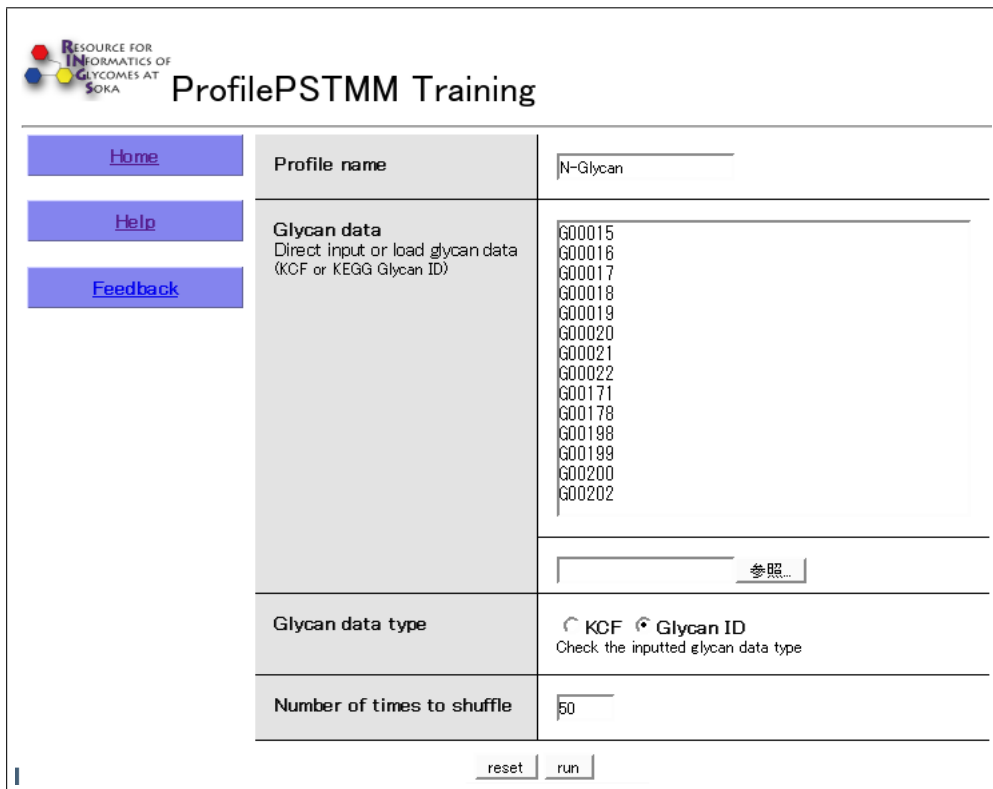
ProfilePSTMM とは、機械学習のモデル [5] を用いており、大量の糖鎖構造の情報から共通なプロフィールを抽出することができる。GlycanMiner と異なり、共通のプロフィールは部分木に限らず、直接つながっていない場合でも同時に現れる部分構造が見出される。

利用目的

ProfilePSTMM の応用例として、糖鎖を認識するタンパク質などが糖鎖のどの部分を認識しているかを予測することができる。過去に糖鎖アレイデータからレクチンが認識する糖鎖プロフィールを抽出するのに用いたことがある。

利用方法

1. ユーザーは、入力画面で KCF 形式の糖鎖構造情報か Glycan ID を入力またはファイルからロードし、その形式を選択する。GlycanID を使った場合、オプションとして各 ID の右側にタブで区切って数値を指定することもできる。例として、糖鎖アレイの結合親和性の値を用いることができる。
2. シャッフルする回数を入力する。確率モデルであるため、局所的な最適解が出力される可能性があるため、複数回実行し、最も高いスコアを結果として取り出す。
3. 最後に run ボタンを押すと解析結果が表示される。シャッフル回数が多ければ多いほど、時間がかかるが、最適解を探すためには多い数がおすすめである。



RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

ProfilePSTMM Training

Home Help Feedback

Profile name	N-Glycan
Glycan data Direct input or load glycan data (KCF or KEGG Glycan ID)	<input type="text"/> <input type="button" value="参照..."/>
Glycan data type	<input type="radio"/> KCF <input checked="" type="radio"/> Glycan ID Check the inputted glycan data type
Number of times to shuffle	50

reset run

図 8.1: ProfilePSTMM の入力画面。ユーザーは、入力画面で KCF 形式の糖鎖構造情報か Glycan ID を入力またはファイルからロードし、その形式を選択する。次に、シャッフルする回数を入力する。確率モデルであるため、局所的な最適解が出力される可能性があるため、複数回実行し、最も高いスコアを結果として取り出す。そして run ボタンを押すと解析結果が表示される。

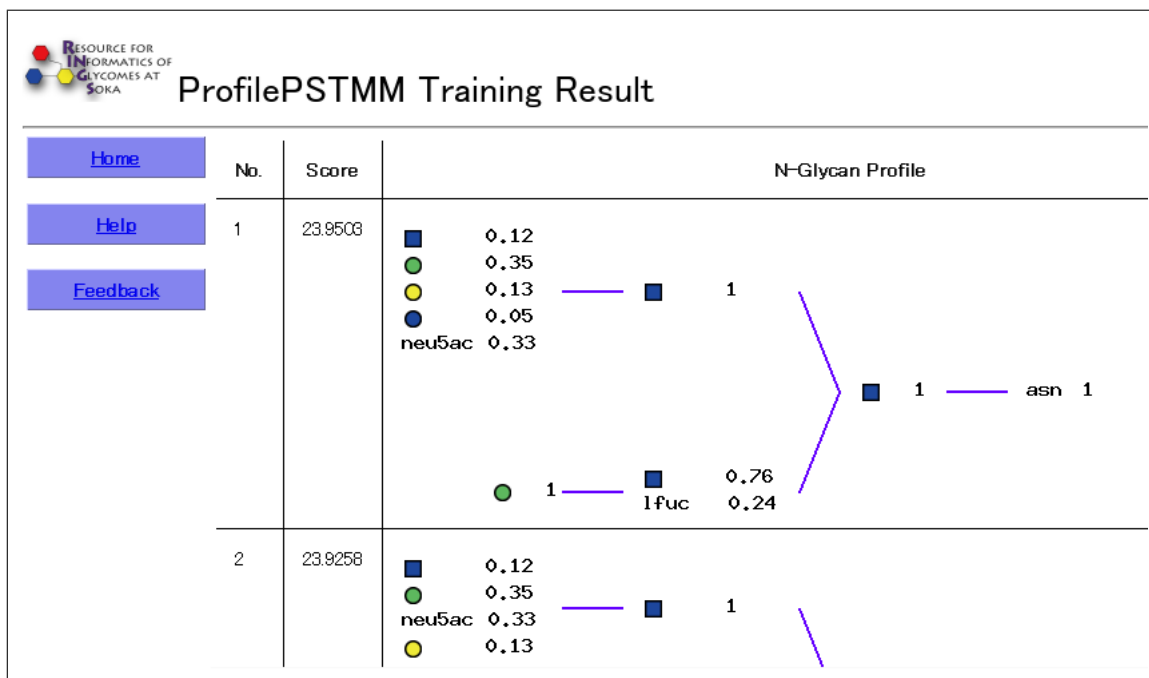


図 8.2: ProfilePSTMM の結果画面。スコアとプロファイリングの結果が表示される。

第9章 Utilities

RINGS には糖鎖の形式を変換するためのユーティリティがある。

- GlycoCT to KCF - GlycoCT 形式 [9] の糖鎖構造を KCF 形式に変換する。
- GLYDE-II to KCF - XML の Glyde-II 形式 [12] の糖鎖構造を KCF 形式に変換する。
- IUPAC to KCF - 糖鎖構造の IUPAC 形式 [3] を KCF 形式に変換する。
- KCF to image - KCF 形式の入力に対し、PNG 形式の画像を出力する。
- KCF to LinearCode - KCF 形式の糖鎖構造を Linear Code 形式 [7] に変換する。
- KCF to LINUCS - KCF 形式の糖鎖構造を LINUCS 形式 [6] に変換する。
- KCF to Mol- KCF 形式の糖鎖構造を分子レベルの Mol 形式に変換する。
- KEGG GLYCAN ID to KCF - KEGG の GLYCAN ID に対応する KCF 形式の情報を出力する。
- LinearCode to KCF - Linear Code 形式 [7] の糖鎖構造を KCF 形式に変換する。
- LINUCS to KCF - LINUCS 形式 [6] の糖鎖構造を KCF 形式に変換する。

利用方法

1. ユーザーは、入力画面で変換する形式の糖鎖構造情報を入力またはファイルからロードする。
2. 決定ボタンを押すと変換された形式で表示される。

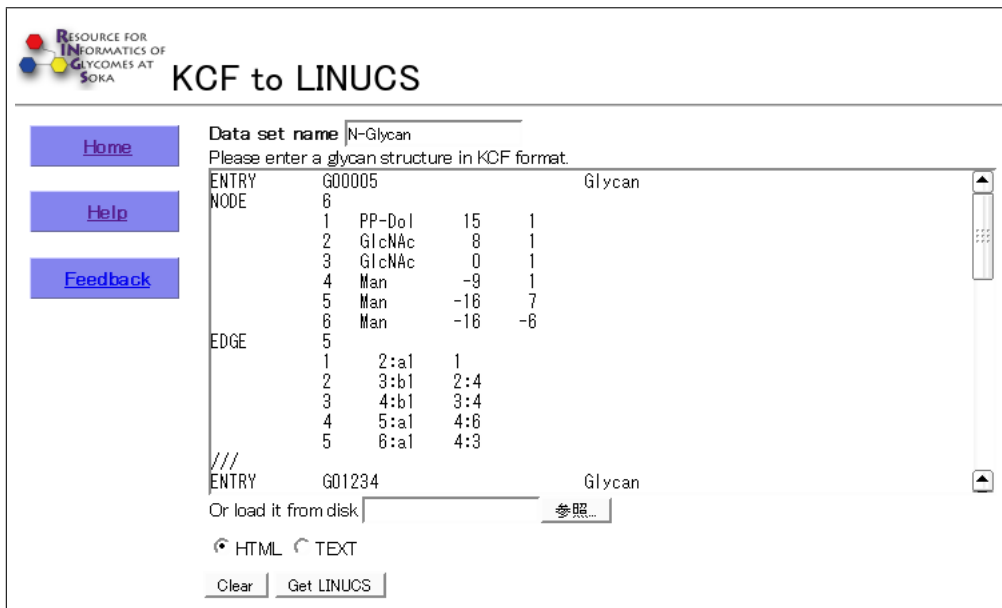


図 9.1: ユーティリティの入力画面の例。ユーザーは、入力画面で変換する形式の糖鎖構造情報を入力またはファイルからロードする。また、決定ボタンを押すと変換された形式で表示される。

No.	Glycan id	LINUCS
1	G00005	[PP-Dol]([0+])[a-GlcNAc]([4+])[b-GlcNAc]([4+])[b-Man]([6+])[a-Man]([3+])[a-Man]([0])]
2	G01234	[Asn]([0+])[b-GlcNAc]([4+])[b-GlcNAc]([4+])[b-Man]([6+])[a-Man]([6+])[b-GlcNAc]([4+])[b-Gal]([3+])[a-LFuc]([2+])[b-GlcNAc]([

図 9.2: ユーティリティの結果画面。変換後の形式の糖鎖構造が表示される。

第10章 KCF to Mol

利用方法

1. ユーザーは、入力画面で KCF 形式の糖鎖構造情報を入力またはファイルからロードする。
2. 表示方法を選択し、決定ボタンを押すと変換された Mol 形式の糖鎖構造を画像で表示する。
3. Get Molfile から Mol ファイルを取得できる。

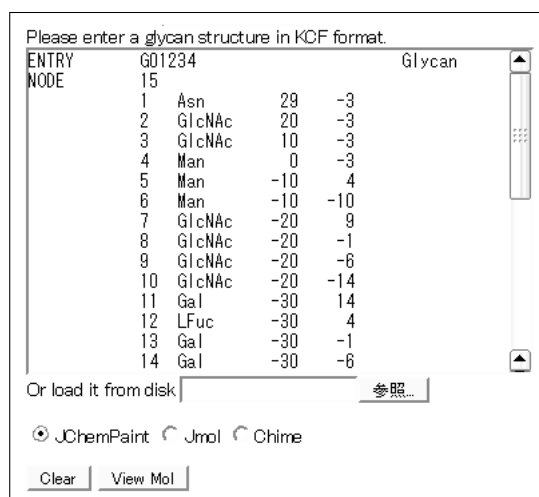


図 10.1: KCFtoMol の入力画面。ユーザーは、入力画面で KCF 形式の糖鎖構造情報を入力またはファイルからロードする。次に表示方法を選択し、決定ボタンを押すと変換された Mol 形式の糖鎖構造を画像で表示する。

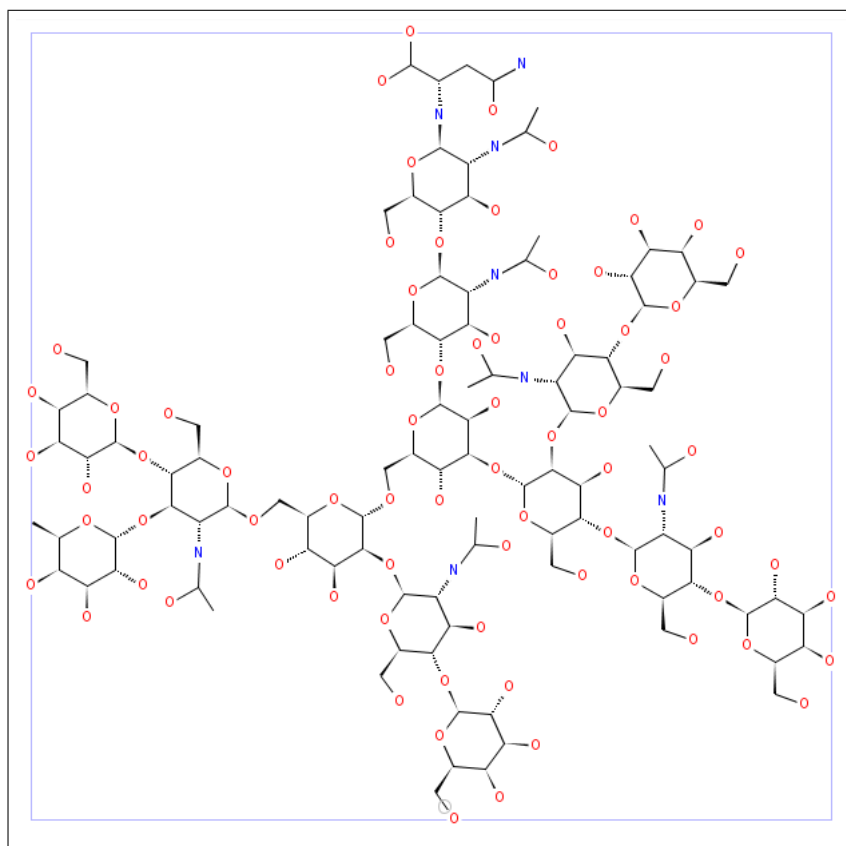


図 10.2: JChemPaint の結果画面。

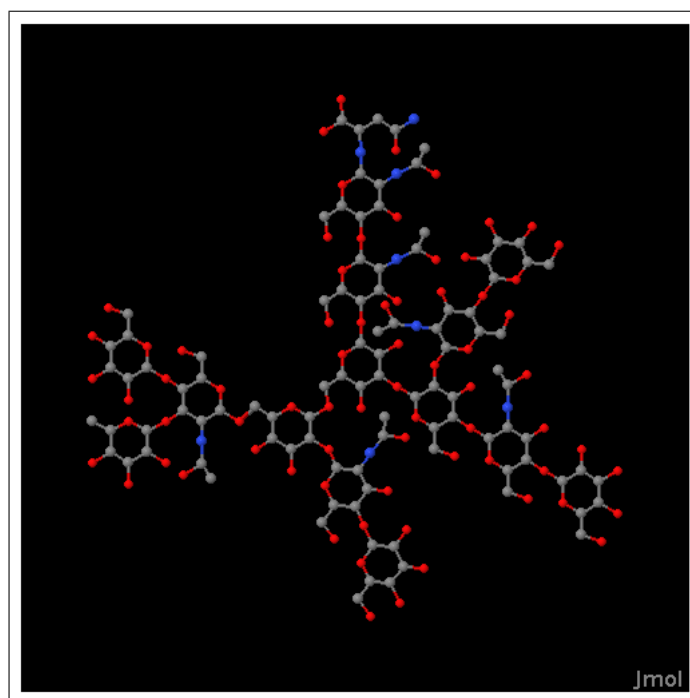


図 10.3: Jmol の結果画面。

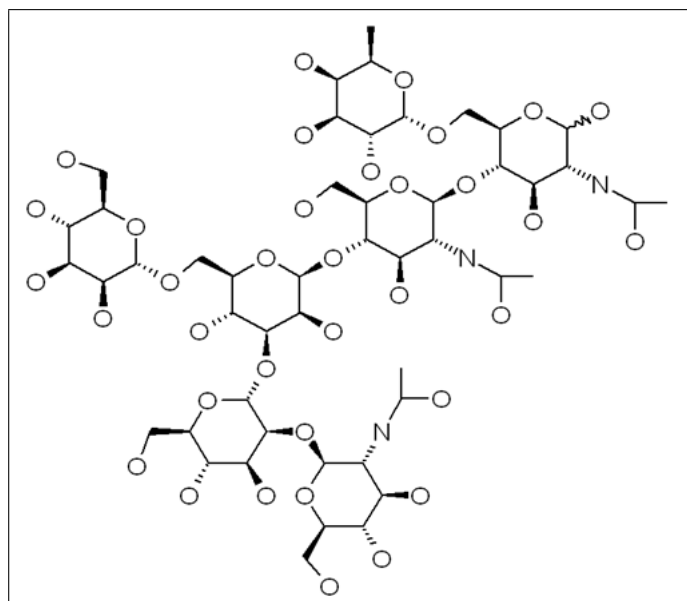


図 10.4: Chime の結果画面。

第11章 ユーザー及びデータ管理

RINGSでは、ユーザー登録を行うことにより、ツールを実行する際に入力したデータと解析結果を保存することができる。

ユーザー登録方法

1. RINGSのトップ画面の「User registration form」をクリックするとユーザー登録画面が表示される。
2. 名前、性別、Eメールアドレス、役職、所属、住所、国、電話番号、パスワードを入力する。必須項目は、名前、Eメールアドレス、パスワードである。入力が終わればSubmitボタンを押す。
3. 入力した情報を確認し、間違いがなければ「send」ボタンを押すことでユーザー登録が完了する。間違いがあれば「back」ボタンを押して修正することができる。

保存する方法

1. RINGSのトップ画面からEメールアドレスとパスワードを入力し「login」ボタンを押すことで左側にデータツリーが表示される。
2. 使用したいツールを選び、任意のデータセット名を入力してツールを実行すると解析結果が表示される。
3. 左側のデータツリーに入力したデータセット名が表示され、入力データと解析結果が保存されたこと示す。「Output」をクリックすると再度解析結果が表示される。
4. データツリーの「Input」をクリックすると保存された入力データが表示される。
5. 右上の「Download」をクリックするとテキスト形式で表示される。
6. プルダウンメニューからツールを選んで「Go」ボタンを押すとツールの入力画面が、保存した入力データを入力した状態で表示される。

情報の修正

1. ユーザーの情報の修正
 - (a) データツリーのユーザー名をクリックするとユーザーの詳細画面が表示される。

RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

User registration form

[Home](#) *First/Given name: RINGS *Last/Family name: Soka

[Help](#) Gender: male

[Feedback search](#) *E-mail: soka_rings@rings.ac.jp

Professional title or position in organization: student

Institution/Affiliation/Company: Soka University

Address: Soka University, 1-236 Tangi-cho, Hachioji City, Tokyo

*Country: Japan

Phone: 03-xxxxx-xxxxx

*Password (6 to 16 character):

*Please enter password again:

図 11.1: ユーザー登録画面。名前、性別、E-メール、役職、所属、住所、国、電話番号、パスワードを入力する。必須項目は、名前、E-メール、パスワードである。入力が済んだら Submit ボタンを押す。

RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

Check your data

Name	RINGS Soka
Gender	male
E-mail	soka_rings@rings.ac.jp
Position	student
Affiliation	Soka University
Address	Soka University, 1-236 Tangi-cho, Hachioji City, Tokyo
Country	Japan
Phone	03-xxxx-xxxx
Password	*****

図 11.2: ユーザー登録の確認画面。登録する情報に間違いがあれば「back」ボタンを押して修正することができる。間違いがなければ「send」ボタンを押すことでユーザー登録が完了する。

Welcome RINGS Soka!!

open all | close all

- RINGS Soka
 - DrawRINGS
 - Glycan Miner
 - GPP
 - Profile PSTMM
 - GlycoCT to KCF
 - GLYDE2 to KCF
 - IUPAC to KCF
 - KCF to LinearCode
 - KCF to LINUCS
 - KCF to Mol
 - KEGG GLYCAN ID to KCF
 - LinearCode to KCF
 - LINUCS to KCF

Total size : 0.00MB

Welcome to **User registration form** [Feedback search](#)

Search for display

Tools

- DrawRINGS**: 2D Java-based glycan structure drawing tool for generating KCF format and/or querying the RINGS database, which current contains the glycan structures from [Glycome-DB](#) and an internally curated data set from the literature.
- Glycan Miner Tool**: implemented based on [Hashimoto et al., 2008](#) for mining alpha-closed frequent subtrees from a set of glycan structures
- Glycan Pathway Predictor (GPP) Tool**: implemented based on [Krambeck et al., 2005](#), which was later improved in [Krambeck et al., 2009](#), for dynamically computing the *N*-glycan biosynthesis pathway from a given glycan structure
- Profile PSTMM Tool**: generate glycan profiles from glycan structure data which can be entered together with binding affinity data, for example, for a particular biological sample

Utilities

- GlycoCT to KCF**: convert a glycan structure in [GlycoCT](#) format to KCF
- GLYDE2 to KCF**: convert a glycan structure in GLYDE2 to KCF
- IUPAC to KCF**: convert a glycan structure in IUPAC format to KCF
- KCF to image**: retrieve the image given a KCF
- KCF to LinearCode**: retrieve the LinearCode format given a KCF
- KCF to LINUCS**: retrieve the LINUCS format given a KCF
- KCF to Mol**: retrieve the chemical structure in MOL format given a glycan structure in KCF
- KEGG GLYCAN ID to KCF**: retrieve the KCF for a given KEGG GLYCAN ID
- LinearCode to KCF**: retrieve the KCF format given a LinearCode
- LINUCS to KCF**: convert a glycan structure in LINUCS format to KCF

Documentation

- Help**: brief users manual for the tools and utilities provided
- What's new?**: latest updates

Downloads

図 11.3: ログイン後のトップ画面。ログインすると画面左側にデータツリーが表示される。この状態でツールを実行すると入力データと解析結果が保存される。

Welcome RINGS Soka!!

open all | close all

- RINGS Soka
 - DrawRINGS
 - Glycan Miner
 - GPP
 - Profile PSTMM
 - GlycoCT to KCF
 - GLYDE2 to KCF
 - IUPAC to KCF
 - KCF to LinearCode
 - KCF to LINUCS
 - KCF to Mol
 - KEGG GLYCAN ID to KCF
 - LinearCode to KCF
 - LINUCS to KCF

Total size : 0.00MB

Reload

Logout

RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

Glycan Miner

Home

Help

Feedback

Data set name

Please enter glycan structures in KCF format.

ENTRY	G00015				Glycan			
8	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:

Clear

図 11.4: ログイン状態の GlycanMiner の入力画面。まず、任意のデータセット名を入力する(この例では N-Glycan)。次に、KCF 形式の糖鎖情報、必要なパラメータを入力する。ツールを実行させると解析結果が表示され、入力データと解析結果が保存される。

Welcome RINGS Soka!!

open all | close all

- RINGS Soka
 - DrawRINGS
 - Glycan Miner
 - N-Glycan
 - Input
 - Output
 - GPP
 - Profile PSTMM
 - GlycoCT to KCF
 - GLYDE2 to KCF
 - IUPAC to KCF
 - KCF to LinearCode
 - KCF to LINUCS
 - KCF to Mol
 - KEGG GLYCAN ID to KCF
 - LinearCode to KCF
 - LINUCS to KCF

Total size : 0.01MB

Reload

Logout

RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

Miner Results

alpha: 0.5
Minsup: 3
Subtrees: 1

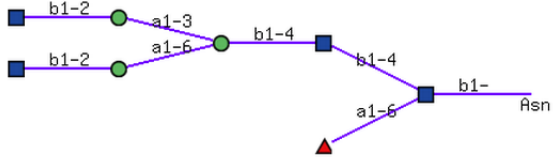
No.	P-value	Name	Support	Subtree
1	0.000000e+00	ST00000	3	

図 11.5: ログイン状態の GlycanMiner の解析結果画面。左側のデータツリーに入力したデータセット名が表示され、保存されたことを示す。保存された解析結果は、データツリーの Output をクリックすると表示することができる。

Welcome RINGS Soka!!

open all | close all

- RINGS Soka
 - DrawRINGS
 - Glycan Miner
 - N-Glycan
 - Input
 - Output
 - GPP
 - Profile PSTMM
 - GlycoCT to KCF
 - GLYDE2 to KCF
 - IUPAC to KCF
 - KCF to LinearCode
 - KCF to LINUCS
 - KCF to Mol
 - KEGG GLYCAN ID to KCF
 - LinearCode to KCF
 - LINUCS to KCF

Total size : 0.01 MB

Reload

Logout

RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

Glycan Miner input data

GlycanMiner [Download](#)

ProfilePSTMM
KCF to LINUCS
KCF to LinearCode

Input structures

[G00015](#)

[G00016](#)

[G00017](#)

図 11.6: 保存された入力データ。データツリーの Input をクリックすると保存された入力データが表示される。右上の Download をクリックするとテキスト形式で表示される。また、プルダウンメニューからツールを選んで「Go」ボタンを押すと選んだツールの入力画面に保存された入力データが表示される。

Welcome Hirota Yasuoka!!

open all | close all

- RINGS Soka
- Profile PSTMM
- Glycan Miner
 - N-Glycan
 - Input**
 - Output
- KEGG GLYCAN ID to KCF
- KCF to LINUCS
- LINUCS to KCF
- GlycoCT to KCF
- KCF to LinearCode
- LinearCode to KCF
- DrawRINGS
- GLYDE2 to KCF
- IUPAC to KCF
- GPP
- KCFtoMol

Total size : 1.29MB

RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

Glycan Miner input data

GlycanMiner [Download](#)

GlycanMiner
ProfilePSTMM
KCF to LINUCS
KCF to LinearCode

Input structures

[G00015](#)

[G00016](#)

[G00017](#)

Welcome Hirota Yasuoka!!

open all | close all

- RINGS Soka
- Profile PSTMM
- Glycan Miner
 - N-Glycan
 - Input**
 - Output
- KEGG GLYCAN ID to KCF
- KCF to LINUCS**
- LINUCS to KCF
- GlycoCT to KCF
- KCF to LinearCode
- LinearCode to KCF
- DrawRINGS
- GLYDE2 to KCF
- IUPAC to KCF
- GPP
- KCFtoMol

Total size : 1.29MB

RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

KCF to LINUCS

Home

Data set name

Please enter a glycan structure 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

HTML TEXT

図 11.7: 上図が GlycanMiner の保存した入力データ、下図が KCFtoLINUCS の入力画面。上図のプルダウンメニューからツールを選び「Go」ボタンをクリックすると下図のように上図の入力データが入力された状態の入力画面が表示される。

Welcome RINGS Soka!!

open all | close all

- RINGS Soka
 - DrawRINGS
 - Glycan Miner
 - N-Glycan
 - GPP
 - FucTH
 - a3FucT
 - ALL
 - Profile PSTMM
 - GlycoCT to KCF
 - GLYDE2 to KCF
 - IUPAC to KCF
 - KCF to LinearCode
 - KCF to LINUCS
 - KCF to Mol
 - KEGG GLYCAN ID to KCF
 - LinearCode to KCF
 - LINUCS to KCF

Total size : 1.09MB

Reload

Logout

RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

User detail

Home

Help

Feedback

ID	13
Name	RINGS Soka
Gender	male
E-mail	soka_rings@rings.ac.jp
Position	student
Affiliation	Soka University
Address	Soka University, 1-236 Tangi-cho, Hachioji City, Tokyo
Country	Japan
Phone	03-xxxx-xxxx
Last login time	2011-02-16 16:20:31
Entry time	2011-02-16 16:00:40

Edit

Delete

図 11.8: ユーザーの詳細画面。ユーザーの名前をクリックすると表示される。また、Edit ボタンを押すことで詳細を変更することができる。

(b) 「Edit」ボタンを押すことでユーザーの情報を変更することができる。

2. データの情報の修正

(a) データツリーのデータセット名をクリックするとデータの詳細画面が表示される。

(b) データの詳細画面では、データセット名の変更やコメントの保存を行うことができる。

The screenshot shows the RINGS Soka web interface. On the left is a file tree under 'RINGS Soka' with folders like 'DrawRINGS', 'Glycan Miner', and 'N-Glycan'. The 'N-Glycan' folder is selected. Below the tree are 'Reload' and 'Logout' buttons. The main area is titled 'Data detail' and contains a table of metadata for the selected dataset:

Name	N-Glycan
Tool	GlycanMiner
Date	2011-02-16 16:39:18
Size	5.819KB
alpha	0.5
minsup	3
Comment	<input type="text"/>

Navigation buttons 'Home', 'Help', and 'Feedback' are on the left. A 'Save' button is at the bottom of the comment field.

図 11.9: データの詳細画面。データツリーのデータセット名をクリックするとデータの詳細が表示される。ここでは、データセット名の変更やコメントを保存することができる。ただし、データセット名は英数字と「_」のみ使用可能である。

3. KCF データの取り扱い

RINGSでは、一つの糖鎖構造(KCF)を入力として実行されるツールと、複数の糖鎖構造(KCFs)を入力として実行されるツールを区別している。例えば、DrawRINGSは一つの糖鎖構造を入力として実行され、複数の糖鎖構造を結果として出力する。しかし、DrawRINGSを用いて複数の糖鎖構造を入力していくうちに、これらを集めて複数の糖鎖構造のデータセット(KCFs)として他のツールで実行できるようにしたい。そこで、データツリーのDrawRINGSをクリックすると、Data listが表示され、今まで入力してきた糖鎖構造が閲覧できる。左側のチェックボックスを用いて複数の構造を選択すると、複数の糖鎖構造を入力とするツールを実行できる。

- (a) データツリーの DrawRINGS をクリックするとデータリストが表示され、過去に入力した糖鎖構造が閲覧できる。
- (b) 他のツールに入力として用いたい糖鎖構造のチェックボックスを選択する。
- (c) 下のプルダウンメニューから、ツール名を選び、GO ボタンをクリックすると、選択したツールの入力画面が表示され、選んだ構造が入力された状態になる。

参考文献

- [1] Y. Akune, M. Hosoda, S. Kaiya, D. Shinmachi, and K.F. Aoki-Kinoshita. The RINGS Resource for Glycome Informatics Analysis and Data Mining on the Web. *OMICS: A Journal of Integrative Biology*, 14(4):475–486, 2010.
- [2] K.F. Aoki, A. Yamaguchi, N. Ueda, T. Akutsu, H. Mamitsuka, S. Goto, and M. Kanehisa. KCaM (KEGG Carbohydrate Matcher): a software tool for analyzing the structures of carbohydrate sugar chains. *Nucleic Acids Res.*, 32:W267–W272, 2004.
- [3] K. Aoki-Kinoshita. An introduction to bioinformatics for glycomics research. *PLoS Computational Biology*, 4(5):e1000075, 2008.
- [4] K.F. Aoki-Kinoshita. *Glycome Informatics: Methods and Applications*. Chapman & Hall/CRC, 2009.
- [5] K.F. Aoki-Kinoshita, N. Ueda, H. Mamitsuka, and M. Kanehisa. ProfilePSTMM: capturing tree-structure motifs in carbohydrate sugar chains. *Bioinformatics*, 22(14):e25–e34, 2006.
- [6] A. Bohne-Lang, E. Lang, T. Forster, and C.-W. von der Lieth. LINUCS: LInear Notation for Unique description of Carbohydrate Sequences. *Carbohydrate Research*, 336(1):1–11, 2001.
- [7] B. Ehud, N. Yael, A. Yaniv, H. Asaf, I. Ori, N. Dotan, and D. Avinoam. A Novel Linear Code Nomenclature for Complex Carbohydrates. *Trends Glycoscience Glycotechnology*, 14(77):127–137, 2002.
- [8] K. Hashimoto, I. Takigawa, M. Shiga, M. Kanehisa, and H. Mamitsuka. Mining significant tree patterns in carbohydrate sugar chains. *Bioinformatics*, 24(16):i167–i173, 2008.
- [9] S. Herget, R. Ranziner, K. Maass, and C.-W. von der Lieth. GlycoCT– a unifying sequence format for carbohydrates. *Carbohydrate Research*, 343:2162–2171, 2008.
- [10] F.J. Krambeck, S.V. Bennun, S. Narang, S. Choi, K.J. Yarema, and M.J. Betenbaugh. A mathematical model to derive N-glycan structures and cellular enzyme activities from mass spectrometric data. *Glycobiology*, 19(11):1163, 2009.
- [11] F.J. Krambeck and M.J. Betenbaugh. A mathematical model of N-linked glycosylation. *Biotechnology and Bioengineering*, 92(6):711–728, 2005.
- [12] N. Packer, C.-W. von der Lieth, K.F. Aoki-Kinoshita, et al. Frontiers in glycomics: bioinformatics and biomarkers in disease. An NIH white paper prepared from discussions by the

focus groups at a workshop on the NIH campus, Bethesda MD (September 11-13, 2006). *Proteomics*, 8(1):8–20, 2008.

- [13] A. Suga, Y. Yamanishi, K. Hashimoto, S. Goto, and M. Kanehisa. An improved scoring scheme for predicting glycan structures from gene expression data. *Genome Inform.*, 18:237–246, 2007.
- [14] A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, et al., editors. *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press, 2nd edition, 2009.