RINGS マニュアル

平成 23 年 2 月 22 日

目 次

1 RINGS について 2

2 DrawRINGS 5

3 GlycanMiner 7

4 Glycan Pathway Predictor (GPP) 9

5 ProfilePSTMM 12

6 Utilities 14

7 ユーザー及びデータ管理 16

1　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。

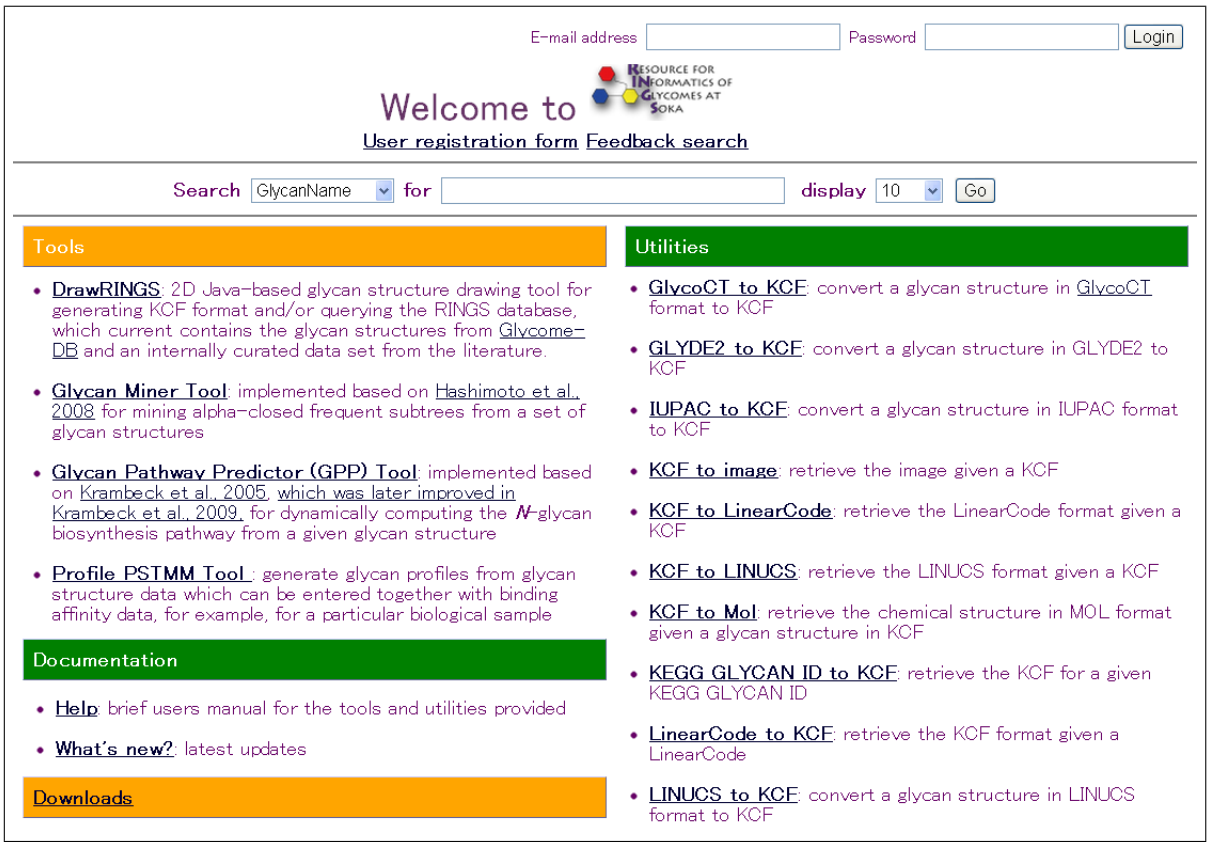

図 1: RINGSのトップ画面。糖鎖解析に用いられる様々なツールや糖鎖構造を表す形式を変換するユーティリティを公開している。

RINGSには糖鎖構造のデータベースに加え、脂質やレクチンの情報も含まれている。糖鎖構造情報はGlycomeDBと呼ばれる網羅的な糖鎖構造データベースと、RINGS独自の糖鎖構造を含む。後者は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、RINGSの構造とリンクされている。

- 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/animallectins/

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 2 に CFG の標識を示す。

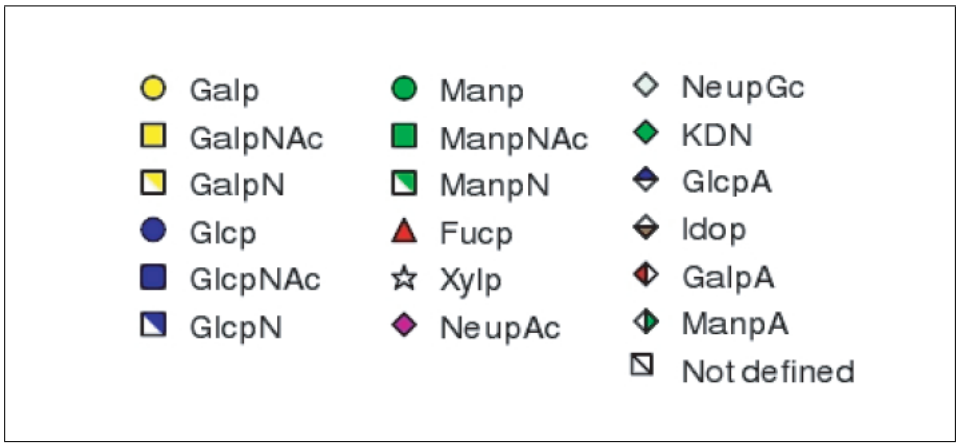

図 2: RINGS で使用する各単糖のシンボル。CFG[14] によって推薦されているシンボルである。

RINGS には二種類のプログラムを用意している。図 1 の画面から見ると、左側には「ツール」で右側に「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 形式への変換ツールはほとんどカバーされている。KCF 形式の例は図 3 に示す。

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。

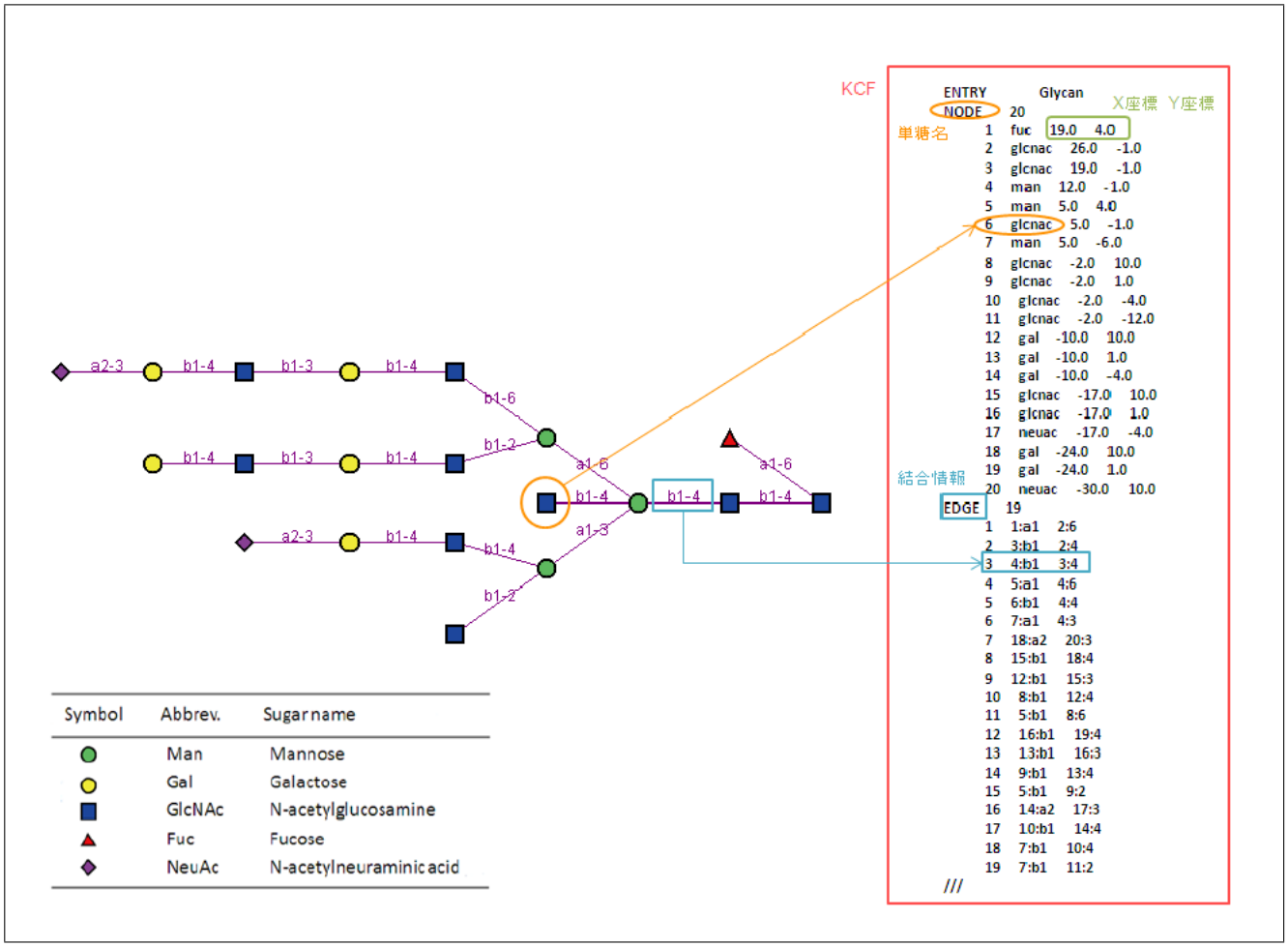

図 3: 糖鎖の画像と KCF。糖鎖の単糖情報を NODE で、糖結合情報を EDGE で表示している。NODE の項目は、NODE 番号、単糖名、単糖の x 座標、y 座標を含み、EDGE の項目は、EDGE 番号、結合している NODE 番号、配座情報を含んでいる。EDGE での、結合している NODE 番号と配座情報は「：」で区切って表示されている。

2　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

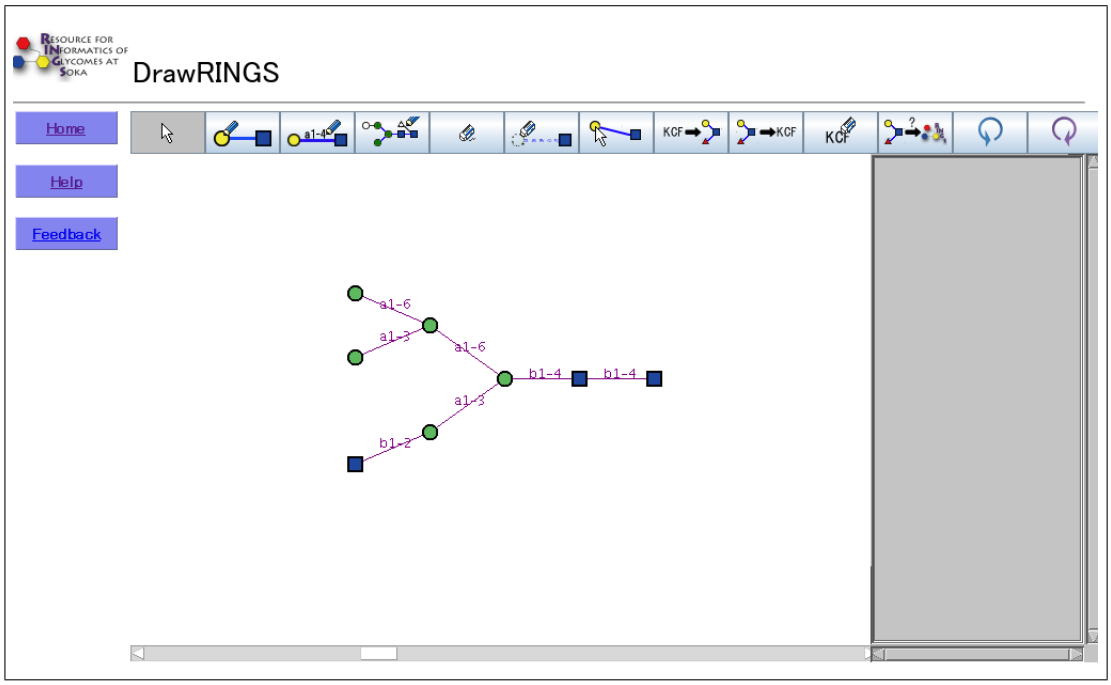

図 4: DrawRINGS のトップ画面。キャンバス部分に、各種ボタンを用いて糖鎖構造を描画することができる。テキストエリアでは KCF 形式のテキストの入出力が可能である。また、描画した構造を検索することが可能である。

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、類似性を示すスコアの高い順に糖鎖構造を図 5 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。

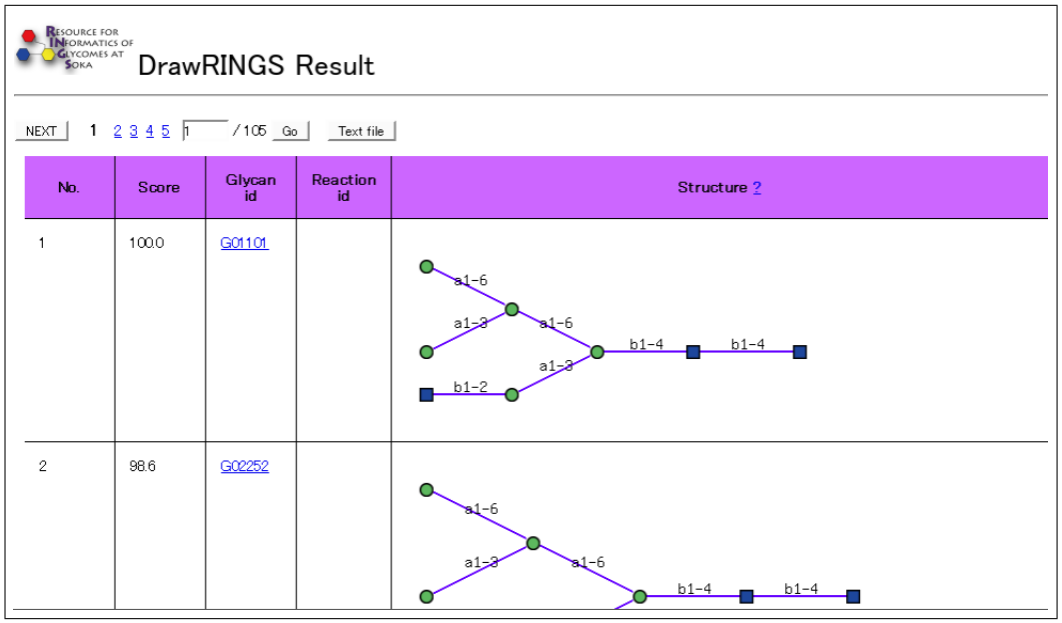

図 5: DrawRINGS からクエリ検索した画面。検索した構造のスコアと ID、構造の画像が表示されている。ID か画像をクリックすることでその構造の詳細な情報を閲覧できる。

3 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」のことである。出現する部分木の support をまず計算する。support とは、部分構造を含む木構造（つまり入力された糖鎖構造）の数である。この support の数が、「ある程度」大きければ、その部分木は frequent subtree と定義する。つまり、たくさんの木構造に含まれる部分木は frequent subtree と定義し、この「たくさん」の数を次に定義する必要がある。この「たくさんの数」は minimum support と呼び、通常「minsup」と略す。GlycanMiner では、ユーザーが minsup を指定する必要がある。

しかし、frequent subtree を抽出すると、その frequent subtree の中に含まれる全ての部分木も frequent subtree として出力されてしまう。このように、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ってしまう。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。

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。α が小さければ小さいほど、制限が厳しく、出力される frequent subtree の数が制限される。

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. alpha 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。

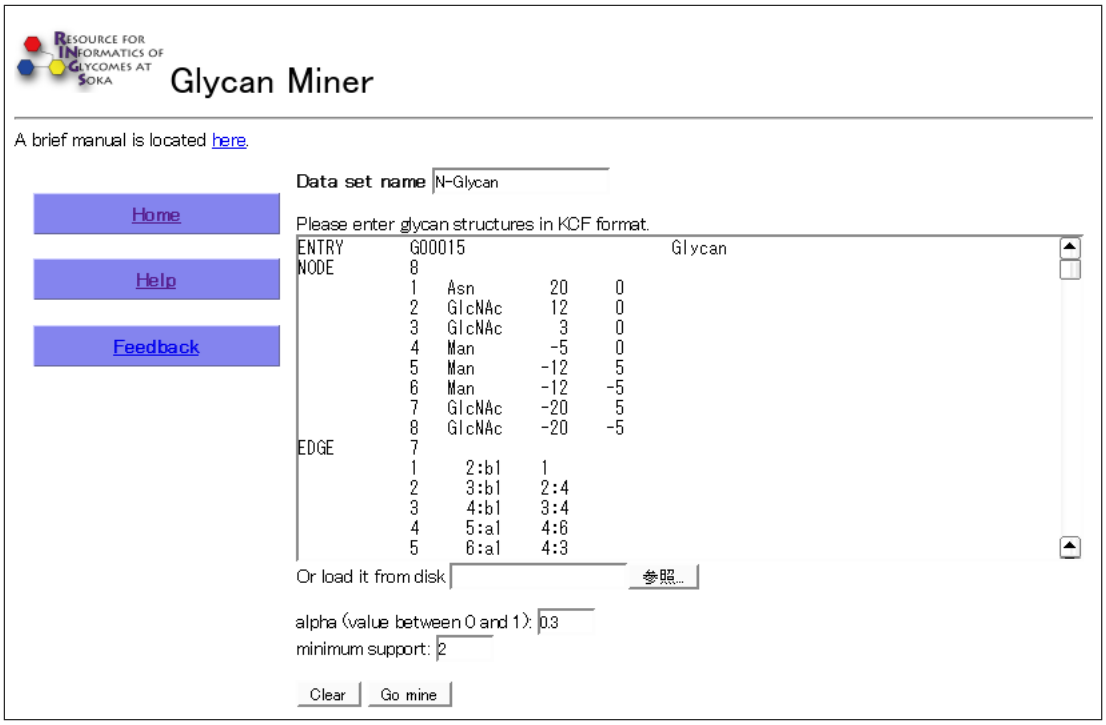

図 6: GlycanMiner の入力画面。KCF 形式の糖鎖構造情報を入力またはファイルからロードする。次に alpha に 0〜1 まで数字を、minimum support に数字をそれぞれ入力し Go mine ボタンを押すと解析結果が表示される。

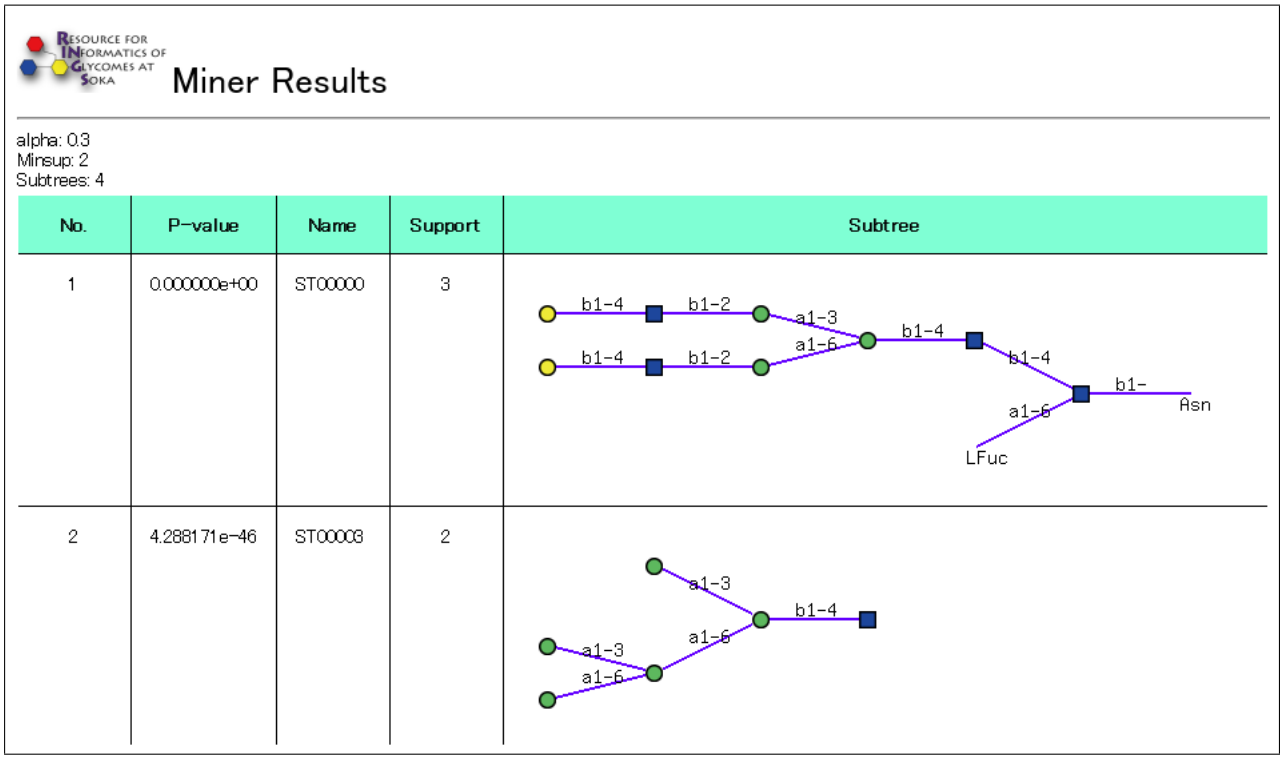

図 7: GlycanMiner の結果画面。P-value、部分構造に自動的に付けられた名前、Support、部分構造の画像が表示される。

4 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

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。

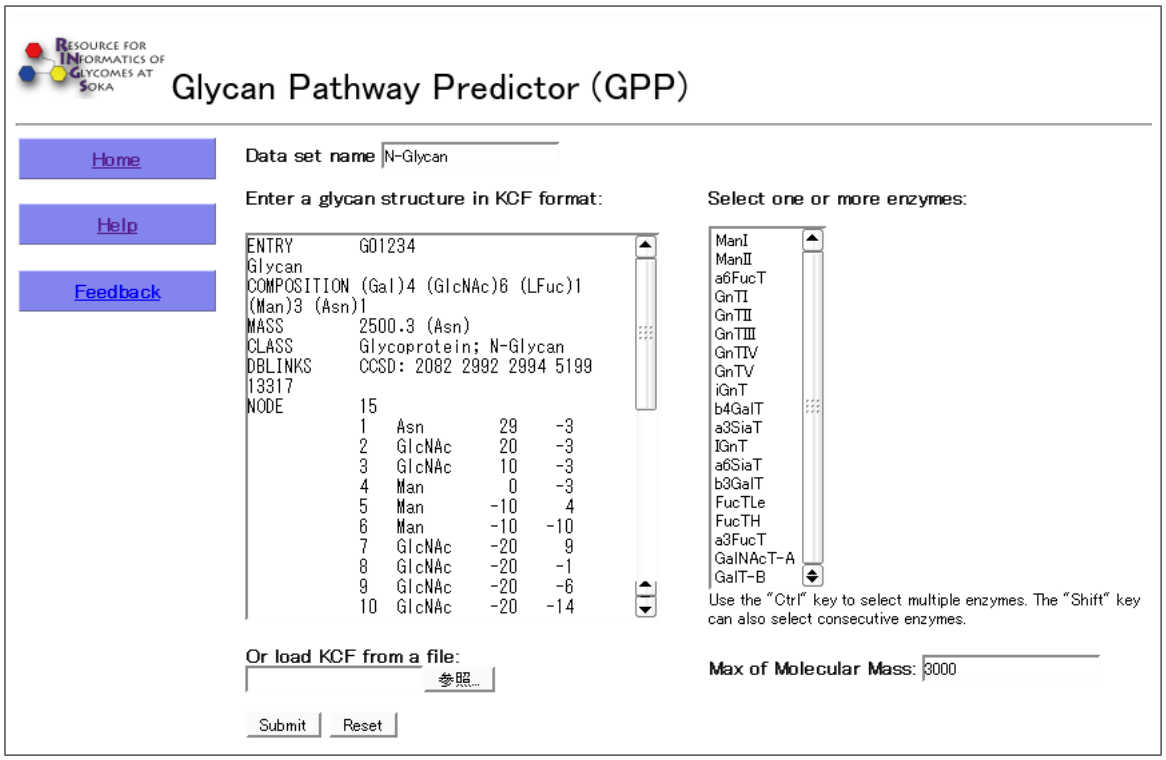

図 8: GPP の入力画面。ユーザーは糖鎖構造の情報を KCF 形式で入力またはファイルをロードし、酵素を選択する。また、計算に必要な糖鎖の分子量の最大値を設定して submit をクリックすることで結果画面を表示させる。

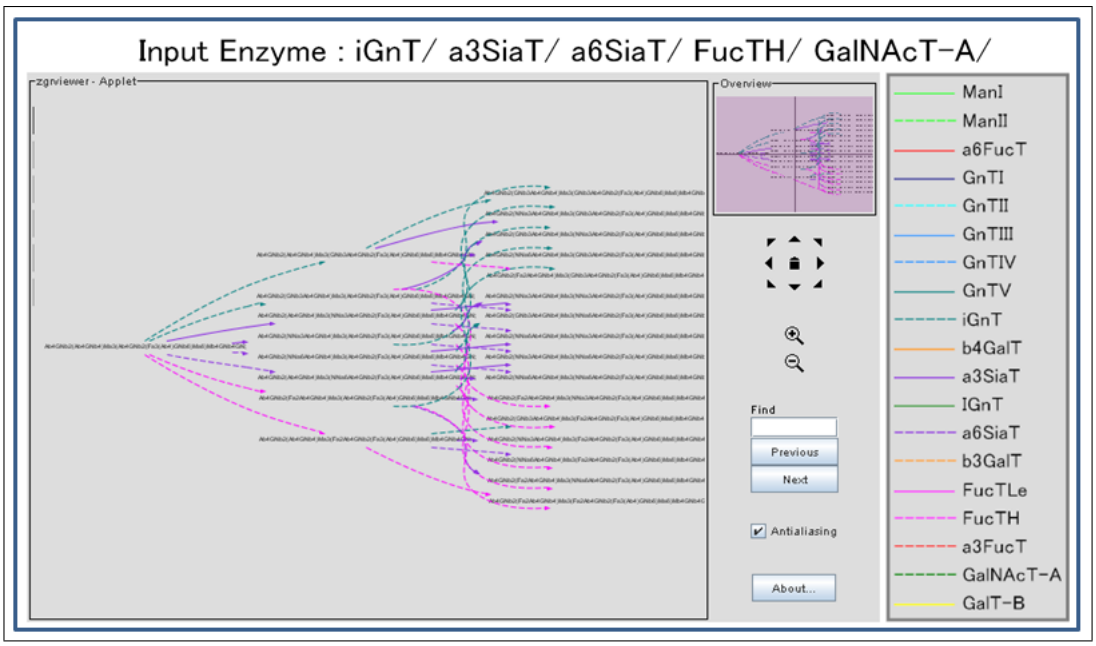

図 9: 計算された反応経路を示す結果画面。グラフ中の線は各酵素を表している。各糖鎖構造はリニアコードで示され、図 10 のような一反応の結果画面にリンクしている。

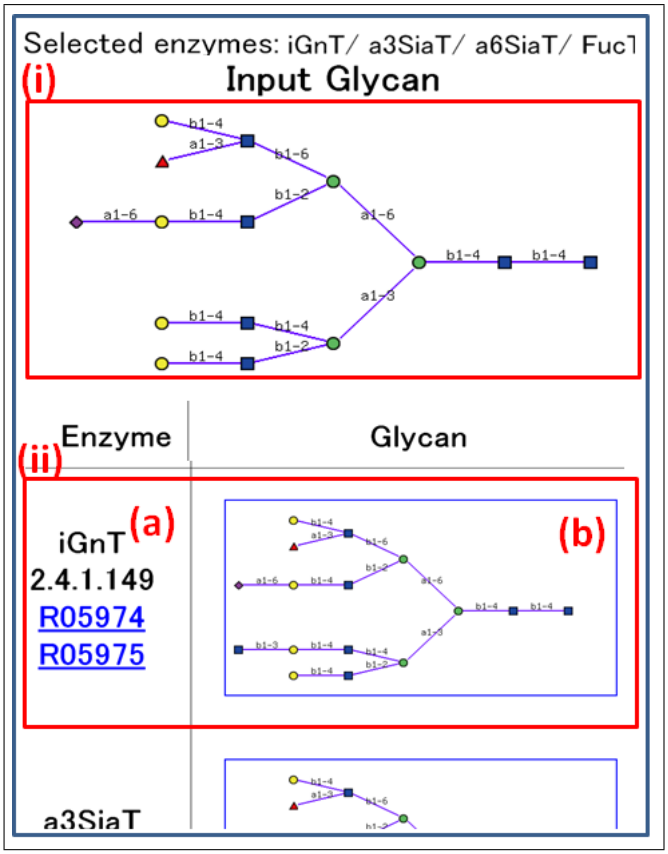

図 10: 一反応の結果を表示する画面。(i) ではクリックした糖鎖構造を画像で表示し、(ii) でその糖鎖に対する一段階の反応を表示している。(ii) の (a) には酵素名とその酵素の EC 番号及びリアクション ID が表示され、(b) にはその酵素の反応によって生成される糖鎖の構造が画像で表示されている。

5 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ァイルを抽出することができる。GlycanMiner と異なり、共通のプロファイルは部分木に限らず、直接つながっていなくて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ァイ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。ProfilePSTMM は確率モデルの特徴である局所的な最適解を出力することがあるため、複数回実行し、最も高いスコアを最適解として出力する。

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有効である。

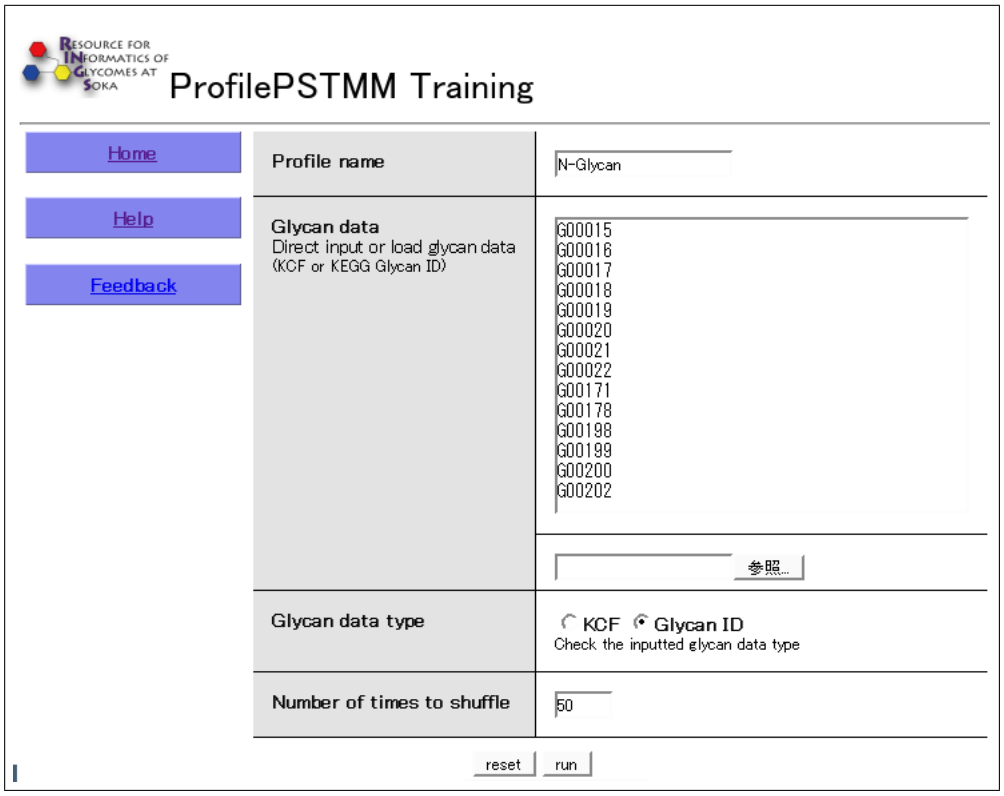

図 11: ProfilePSTMM の入力画面。ユーザーは、入力画面で KCF 形式の糖鎖構造情報か Clycan ID を入力またはファイルからロードし、その形式を選択する。次に、シャッフルする回数を入力する。そして run ボタンを押すと解析結果が表示される。

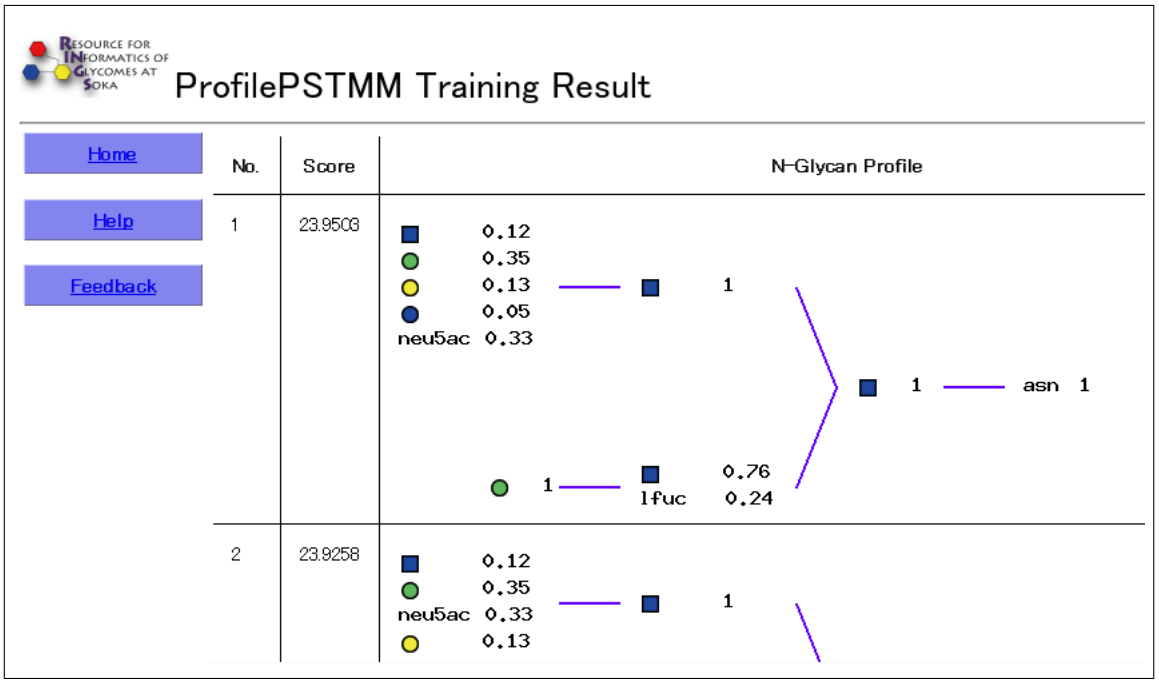

図 12: ProfilePSTMM の結果画面。スコアとプロファイリングの結果が表示される。

6 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 形式の糖鎖構造を原子レベルの MDL 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. 糖鎖構造情報を変換する形式で入力またはファイルからロードする。
2. 結果の表示法を選択する（HTML かテキスト）。HTML の場合は、結果が表で表示され、テキストの場合は、生のデータが出力される。
3. 決定ボタンを押すと変換されたデータが選択された表示形式で出力される。

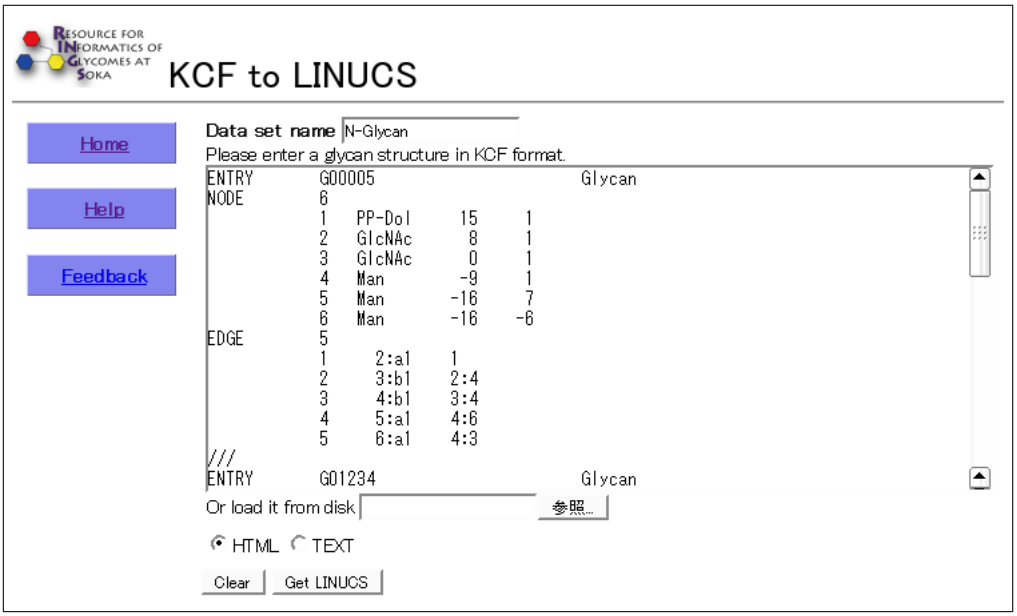

図 13: ユーティリティの入力画面の例。ユーザーは、入力画面で変換する形式の糖鎖構造情報を入力またはファイルからロードする。また、決定ボタンを押すと変換された形式で表示される。

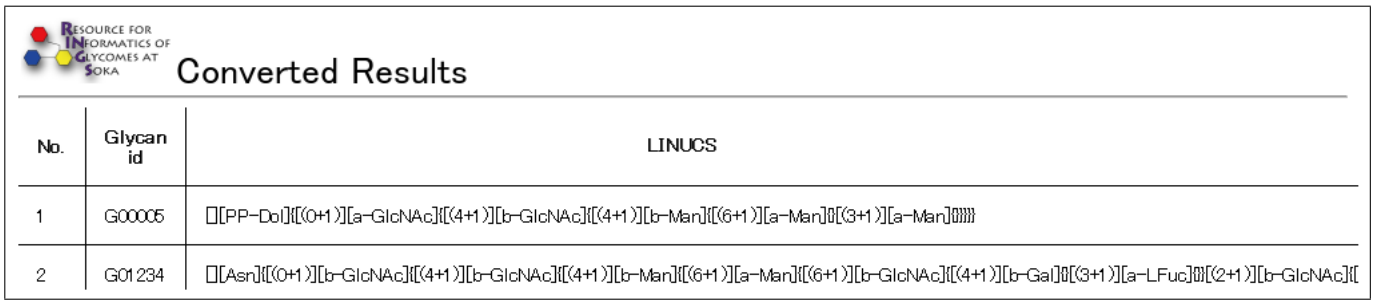

図 14: ユーティリティの結果画面。変換後の形式の糖鎖構造が HTML 形式で表として表示される。

7　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。

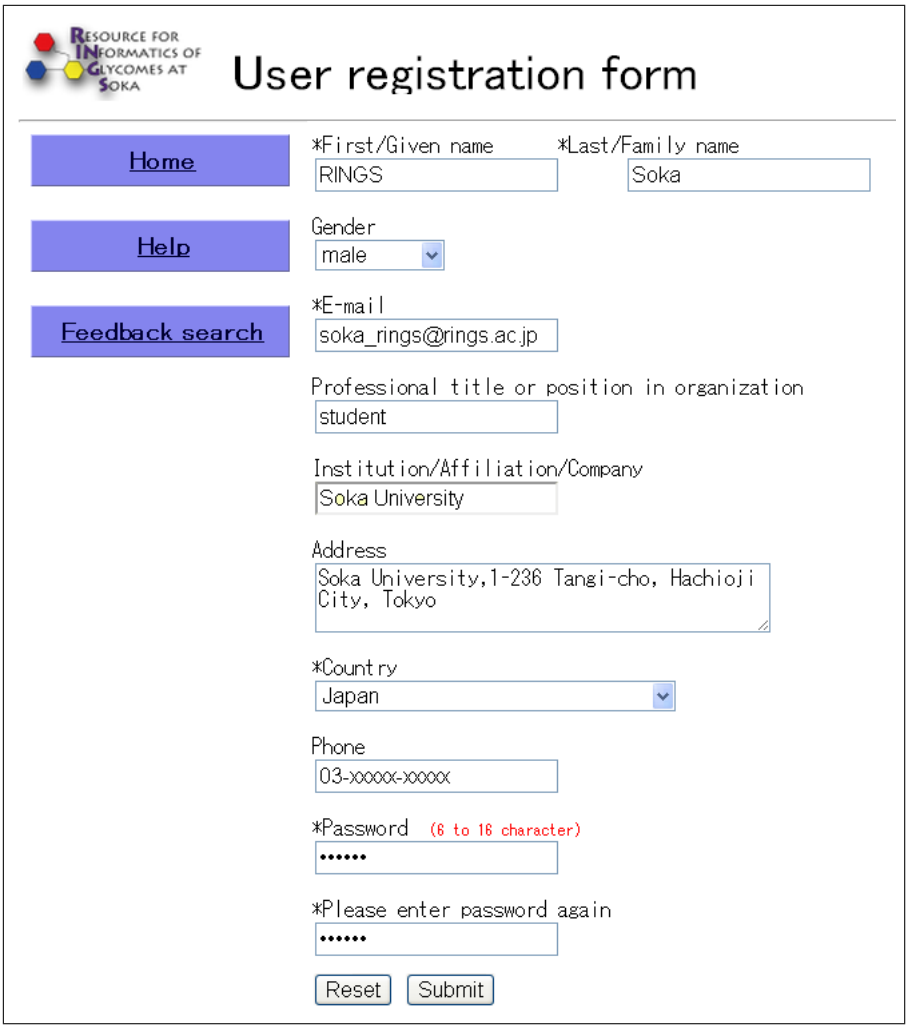

図 15: ユーザー登録画面。名前、性別、E-メール、役職、所属、住所、国、電話番号、パスワードを入力する。必須項目は、名前、E-メール、パスワードである。入力が済んだら Submit ボタンを押す。

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。

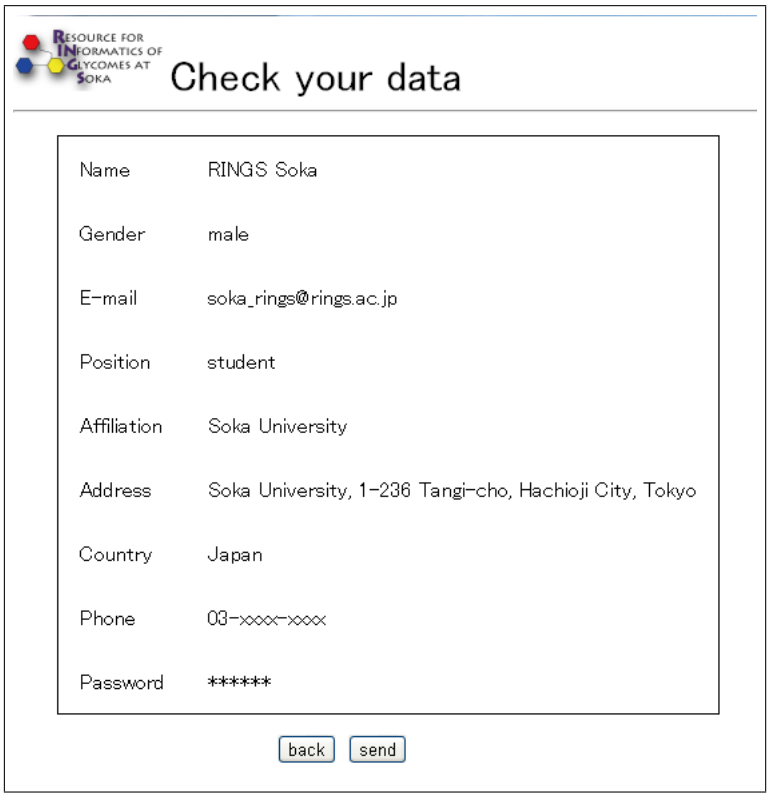

図 16: ユーザー登録の確認画面。登録する情報に間違いあれば「back」ボタンを押して修正することができる。間違いがなければ「send」ボタンを押すことでユーザー登録が完了する。

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。
   (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。

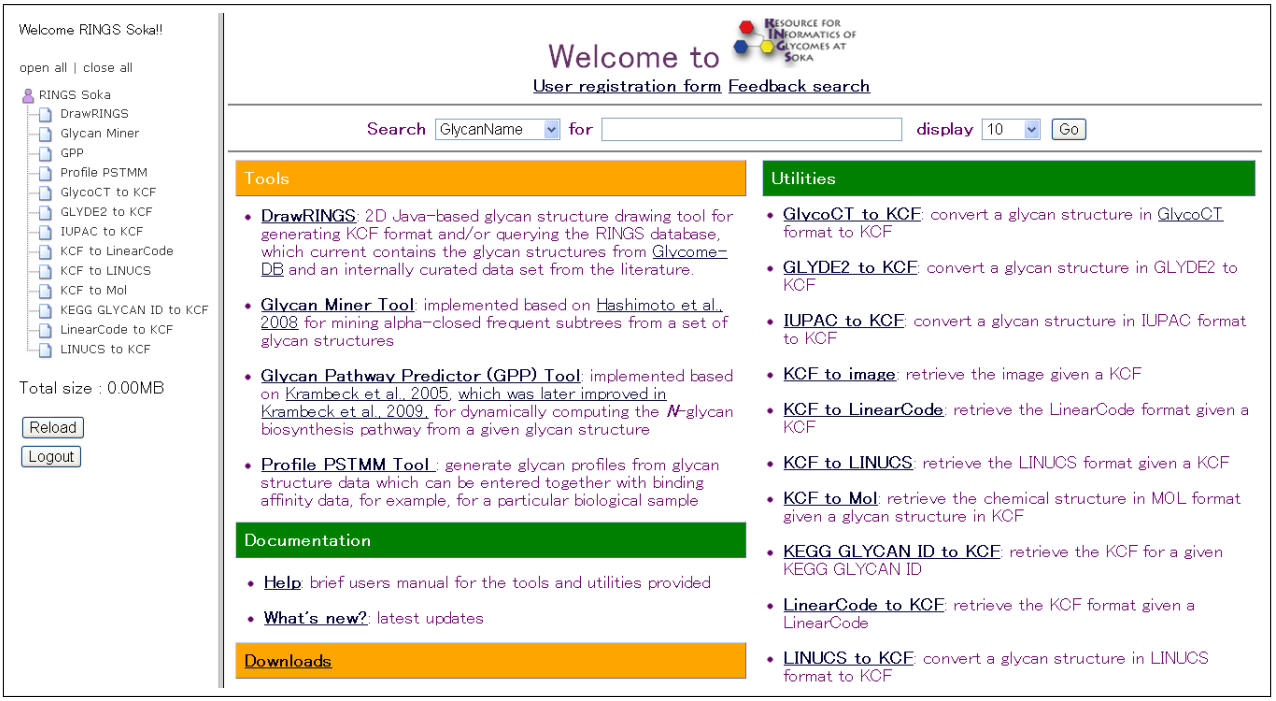

図 17: ログイン後のトップ画面。ログインすると画面左側にデータツリーが表示される。この状態でツールを実行すると入力データと解析結果が保存される。

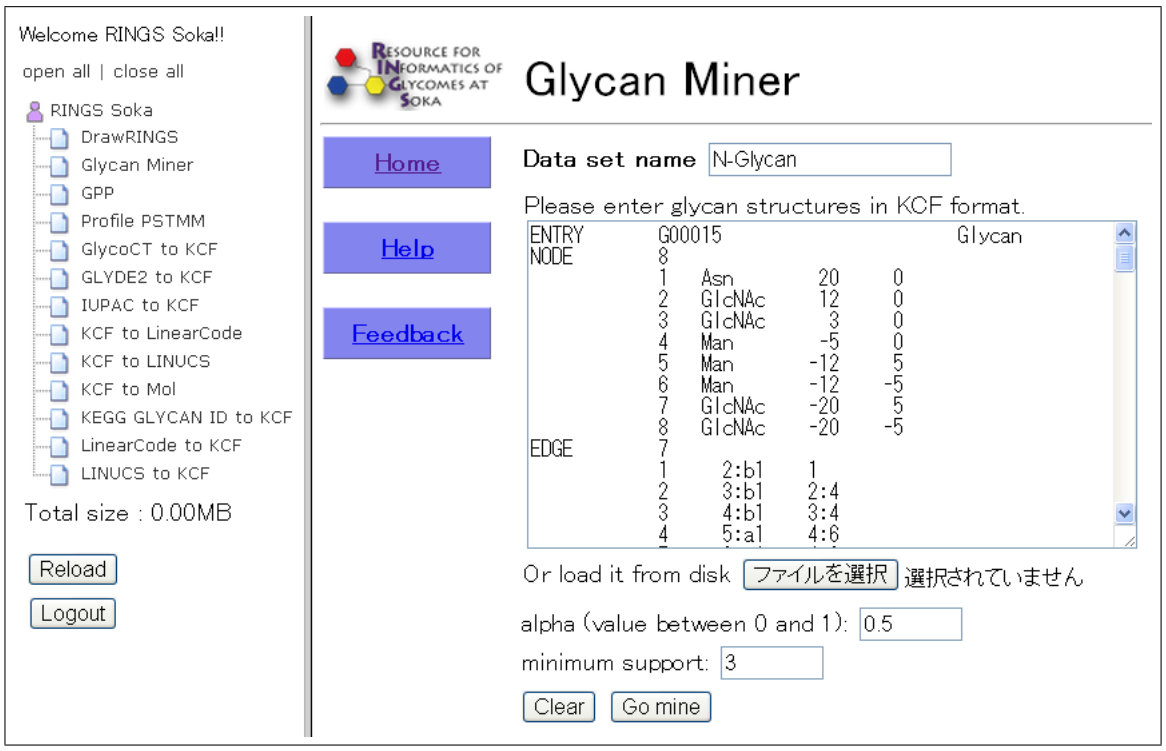

図 18: ログイン状態の GlycanMiner の入力画面。まず、任意のデータセット名を入力する（この例では N-Glycan）。次に、KCF 形式の糖鎖情報、必要なパラメータを入力する。ツールを実行させると解析結果が表示され、入力データと解析結果が保存される。

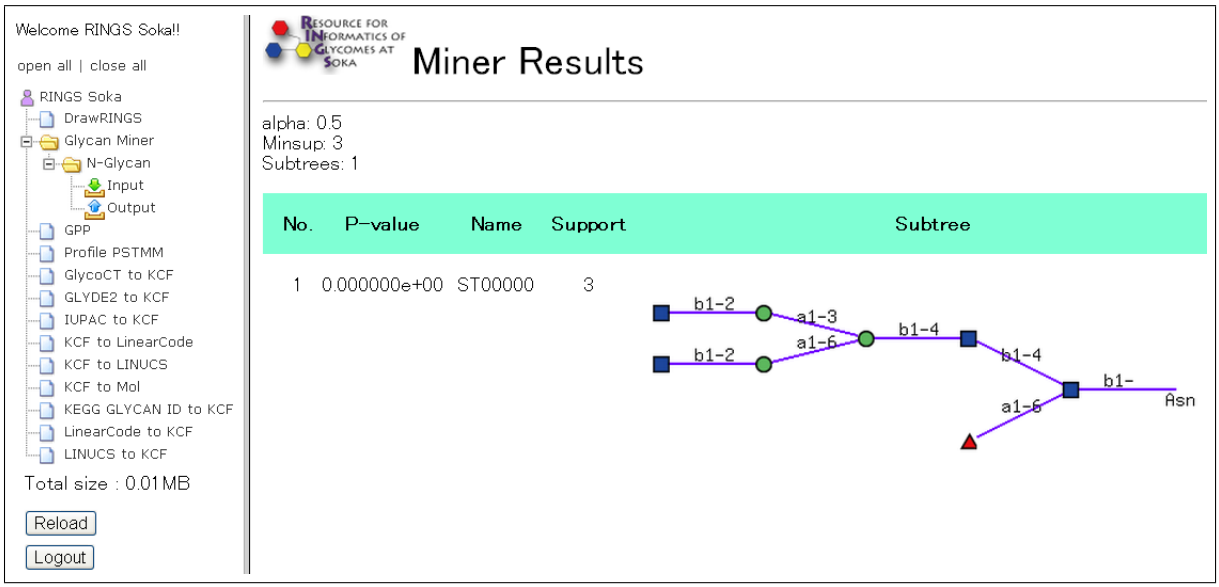

図 19: ログイン状態の GlycanMiner の解析結果画面。左側のデータツリーに入力したデータセット名が表示され、保存されたことを示す。保存された解析結果は、データツリーの Output をクリックすると表示することができる。

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。

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。

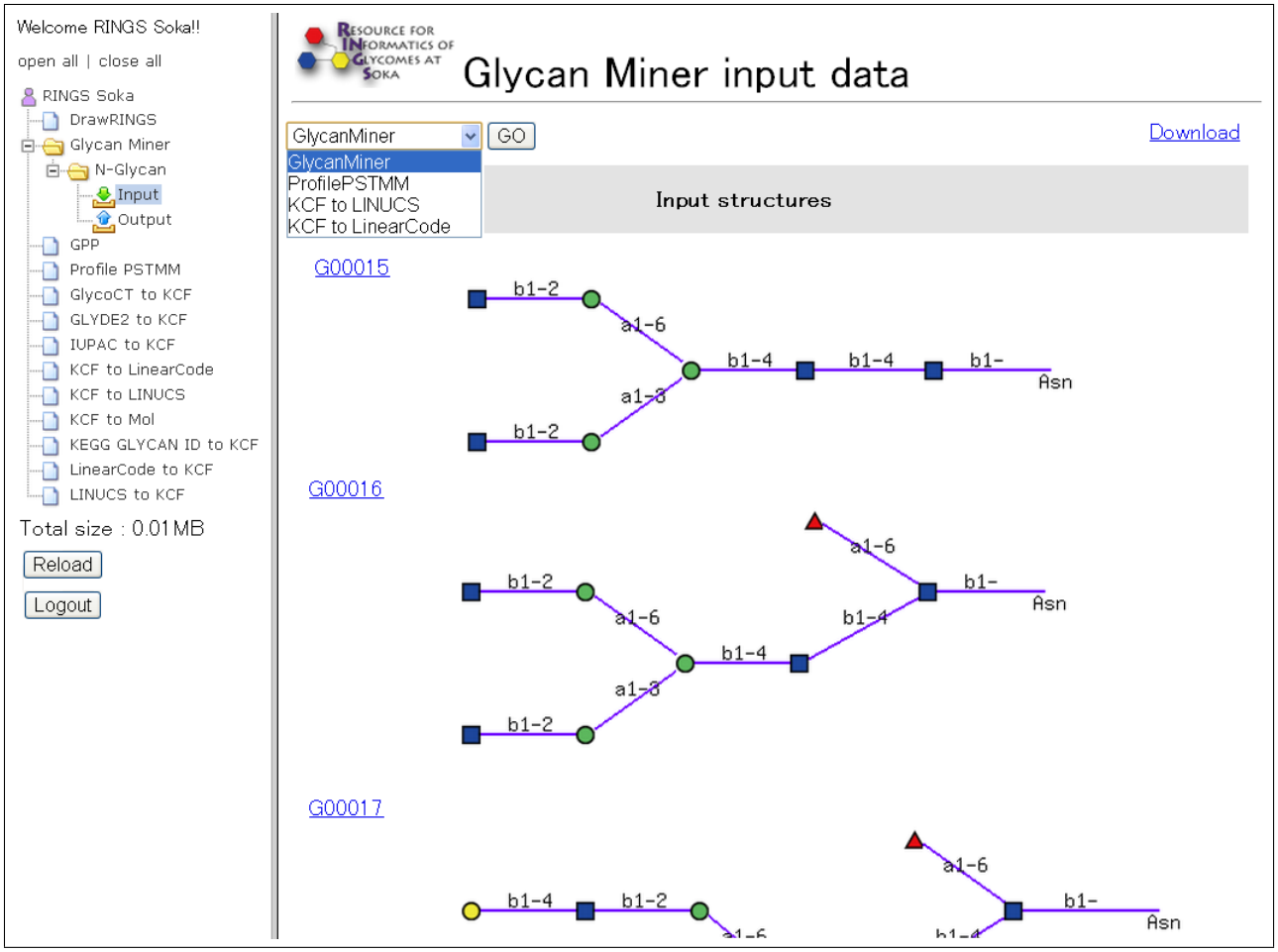

図 20: 保存された入力データ。データツリーの Input をクリックすると保存された入力データが表示される。右上の Download をクリックするとテキスト形式で表示される。また、プルダウンメニューからツールを選んで「Go」ボタンを押すと選んだツールの入力画面に保存された入力データが表示される。

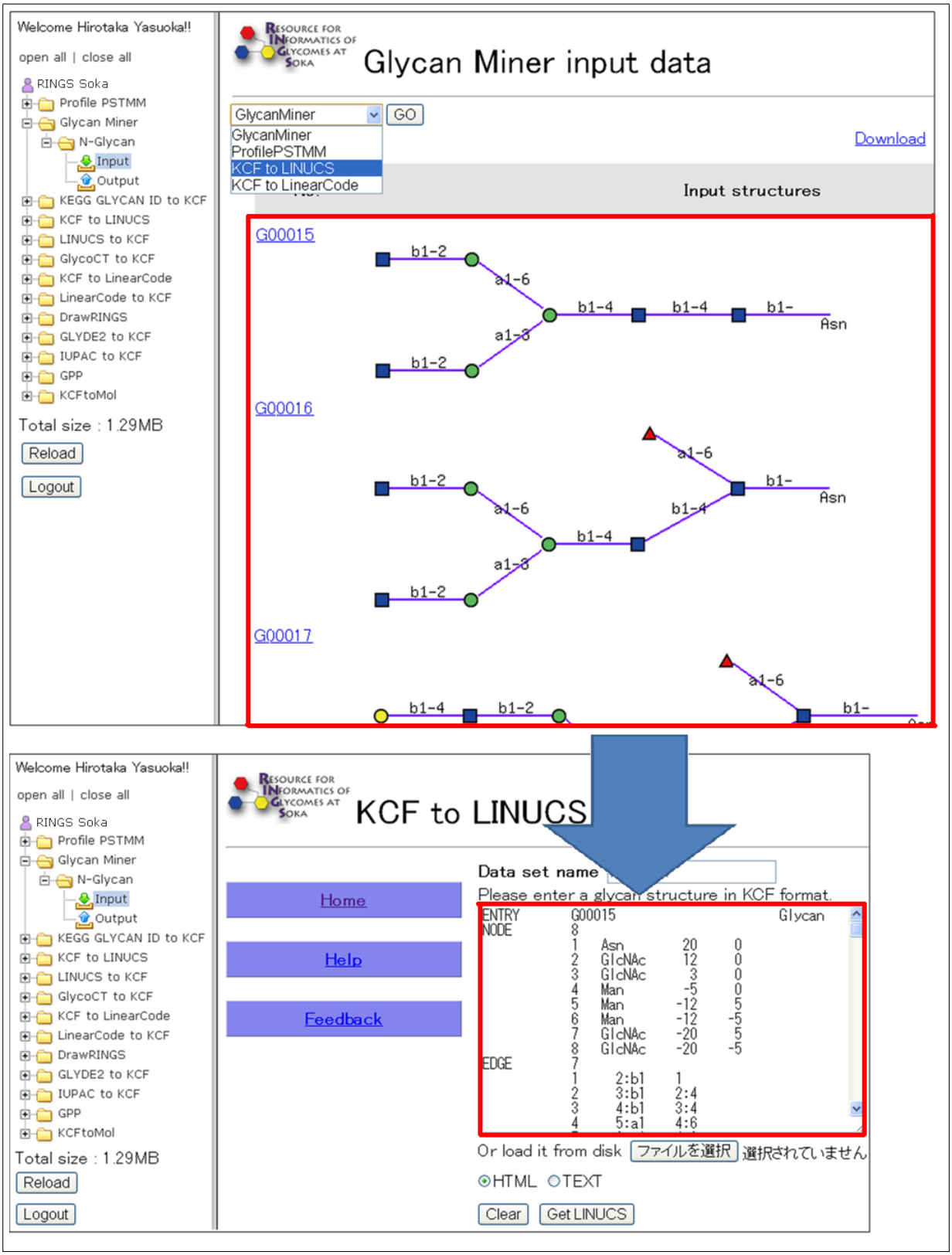

図 21: 上図が GlycanMiner の保存した入力データ、下図が KCFtoLINUCS の入力画面。上図のプルダウンメニューからツールを選び「Go」ボタンをクリックすると下図のように上図の入力データが入力された状態の入力画面が表示される。

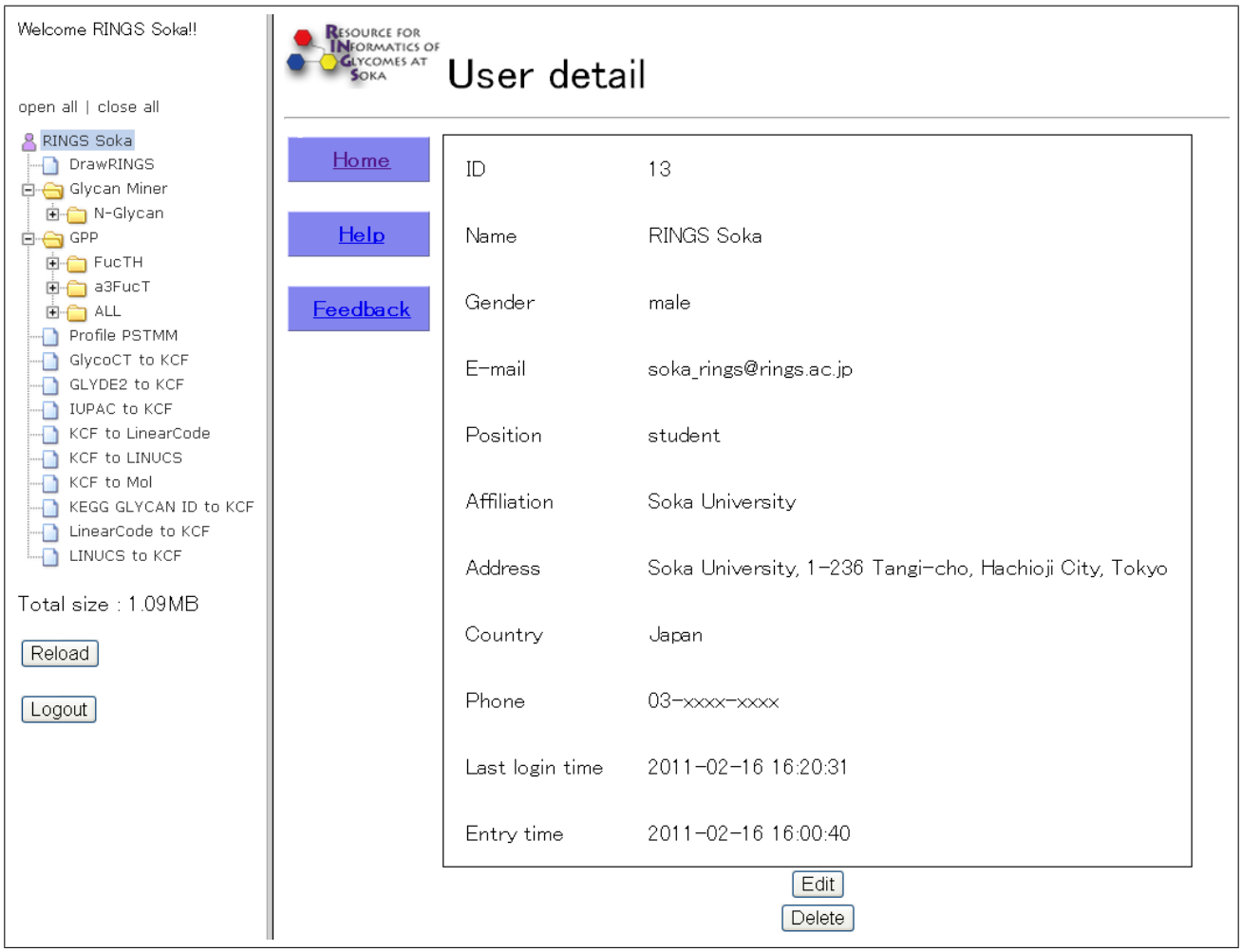

図 22: ユーザーの詳細画面。ユーザーの名前をクリックすると表示される。また、Edit ボタンを押すことで詳細を変更することができる。

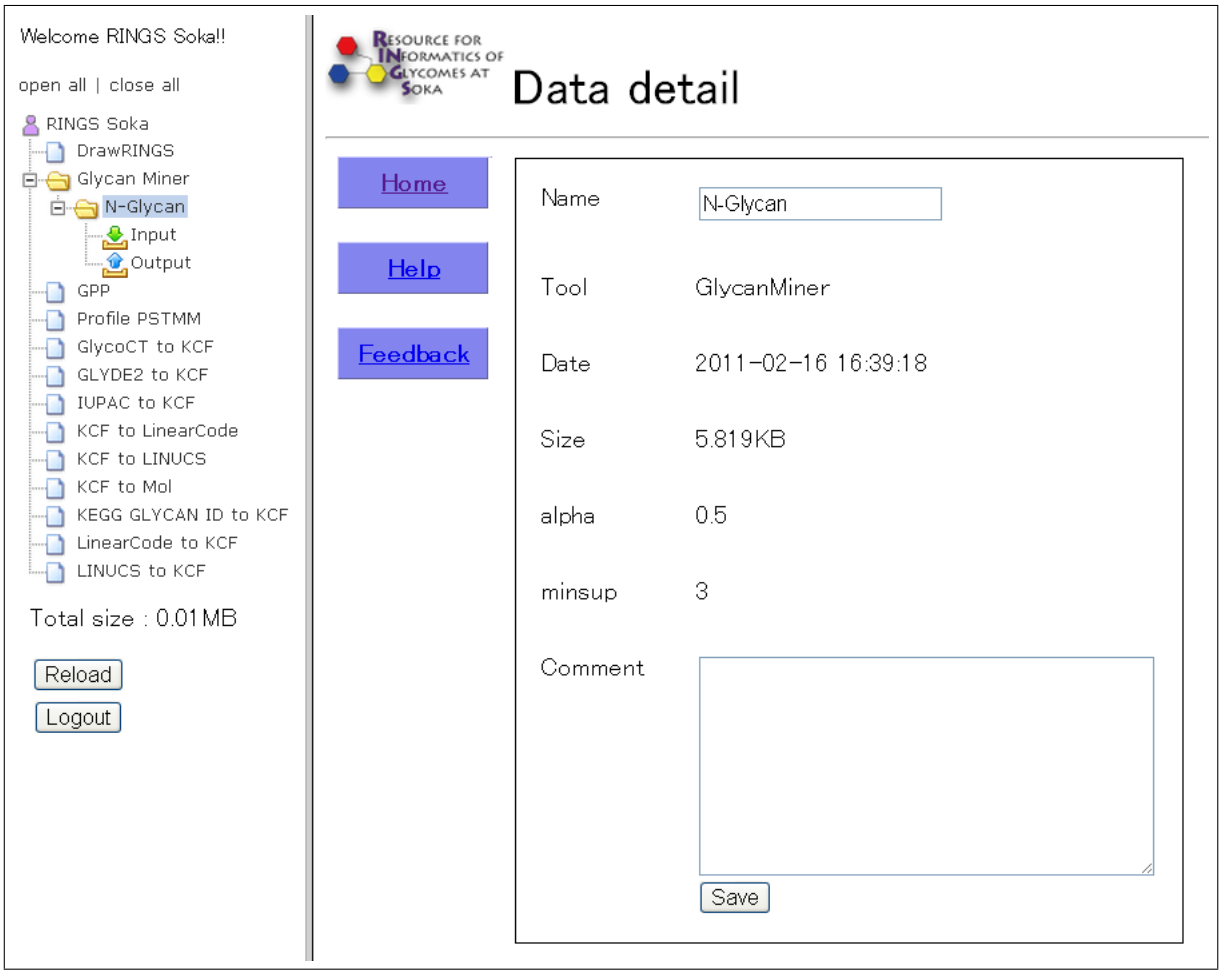

図 23: データの詳細画面。データツリーのデータセット名をクリックするとデータの詳細が表示される。ここでは、データセット名の変更やコメントを保存することができる。ただし、データセット名は英数字と「_」のみ使用可能である。

参考URL

1. RINGS - http://www.rings.t.soka.ac.jp/
2. KEGG GLYCAN - http://www.genome.jp/kegg/glycan/
3. GlycomeDB - http://www.glycome-db.org/
4. Consortium for Functional Glycomics (CFG) - http://www.functionalglycomics.org/

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.