

糖鎖遺伝子発現情報からの糖鎖構造推定ツール GlycoMaple の開発

細田 正恵¹, 木下 聖子¹, 藤田 盛久²

1. 糖鎖の役割と糖鎖合成

糖鎖は、単糖がグリコシド結合により連なった分子であり、糖タンパク質、糖脂質などの複合糖質として細胞の表面上に多く存在する。糖タンパク質のフォールディング制御に関与し、タンパク質の品質管理システムとしての糖鎖の役割が知られている¹⁾。また、生体内における糖鎖を認識するタンパク質であるレクチンとの相互作用により、感染やシグナル伝達に関与するなど、糖鎖の生物学的機能の重要性が注目されており、バイオマーカーや創薬など医療分野への発展が期待されている²⁾。

糖鎖合成は主に小胞体やゴルジ体で行われ、ドリコールリン酸などの脂質結合型糖や糖スクレオチドをドナー基質として糖をアクセプターへ転移させる糖転移酵素と、単糖間のグリコシド結合を加水分解する糖分解酵素により行われる。タンパク質の糖鎖修飾として、アスパラギン (Asp) 残基のアミノ基に共有結合する *N*-グリコシル化型、セリン (Ser) やトレオニン (Thr) 側鎖の水酸基に共有結合する *O*-グリコシル化型、 α -マンノースがトリプトファン (Trp) 残基の C2 原子に結合する *C*-マンノシル化型、エタノールアミンリン酸を介してタンパク質の C 末端に糖鎖付加する GPI アンカー型などがあげられる³⁾。

糖鎖関連遺伝子の変異によるグリコシル化の変化は、神経障害、骨格異常、筋ジストロフィーなどを引き起こすとともに、感染症、腫瘍の進行・転移への関与も報告されて

いる⁴⁻⁶⁾。このようにさまざまな生体内現象に関わるグリコシル化の生物学的重要性が認知されているにもかかわらず、これまでは限定された経路の糖鎖研究しか行われていない。糖鎖修飾が与える影響やその解明には、糖鎖代謝経路のより包括的な理解が望まれる。そのため、グリコシル化を体系的に理解するためのとりかかりとして、糖鎖関連遺伝子の発現情報から糖鎖代謝経路をマッピングするための Web ツールである GlycoMaple が開発された⁷⁾。

2. GlycoMaple ツール

1) ツールの説明


GlycoMaple は、ヒトの遺伝子発現レベルに基づいて糖鎖合成経路を視覚化し、構造を推定することができるツールである。棒グラフにより経路図の遺伝子発現レベルが閲覧できる。現在、19 種類の合成経路図を提供しており、約 950 の糖鎖関連遺伝子の発現プロファイルを可視化することができる。GlycoMaple では、遺伝子発現データの inputs は、Human Protein Atlas (HPA) の細胞や組織の遺伝子発現データが利用できる。また、ユーザーが用意した遺伝子発現データを指定のファイル形式にすることによりツールを利用することも可能である。

HPA は、ヒトの細胞や組織、臓器などのプロテオーム探索のためのデータベースである。タンパク質発現情報の RNA-seq データや画像データなどが公開されており、入手可能な RNA-seq データのうち細胞 64 種、組織 37 種を GlycoMaple に組み込んでいる。

2) ツールの使い方

GlycoMaple ツールは、<https://glycosmos.org/glycomaple/index> にて利用できる。ここでは、HPA データの選択による使用方法とファイル入力による使用方法や、ツールの機能について紹介する。

a. Human Protein Atlas (HPA) の細胞、または組織の選択からの利用

- i) HPA の各細胞または組織から糖鎖生合成経路を閲覧する場合、 1(1) の“HPA (cell)”や“HPA (tissue)”のプルダウンメニューから選択し「Add」ボタンをクリックする。

¹ 創価大学糖鎖生命システム融合研究所 (〒192-8577 東京都八王子市丹木町 1-236)

² 岐阜大学糖鎖生命コア研究所 (〒501-1193 岐阜県岐阜市柳戸 1-1)

Development of GlycoMaple Tool estimating glycan structures from glycogene expression data

Masae Hosoda¹, Kiyoko F. Aoki-Kinoshita¹ and Morihisa Fujita²

(¹Glycan & Life Systems Integration Center (GaLSIC), Soka University, 1-236 Tangi-machi, Hachioji City, Tokyo, 192-8577, Japan,

²Institute for Glyco-core Research (iGCORE), Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan)

本論文の図版はモノクロ (冊子版) およびカラー (電子版) で掲載。

DOI: 10.14952/SEIKAGAKU.2022.940623

© 2022 公益社団法人日本生化学会

テクニカルノート

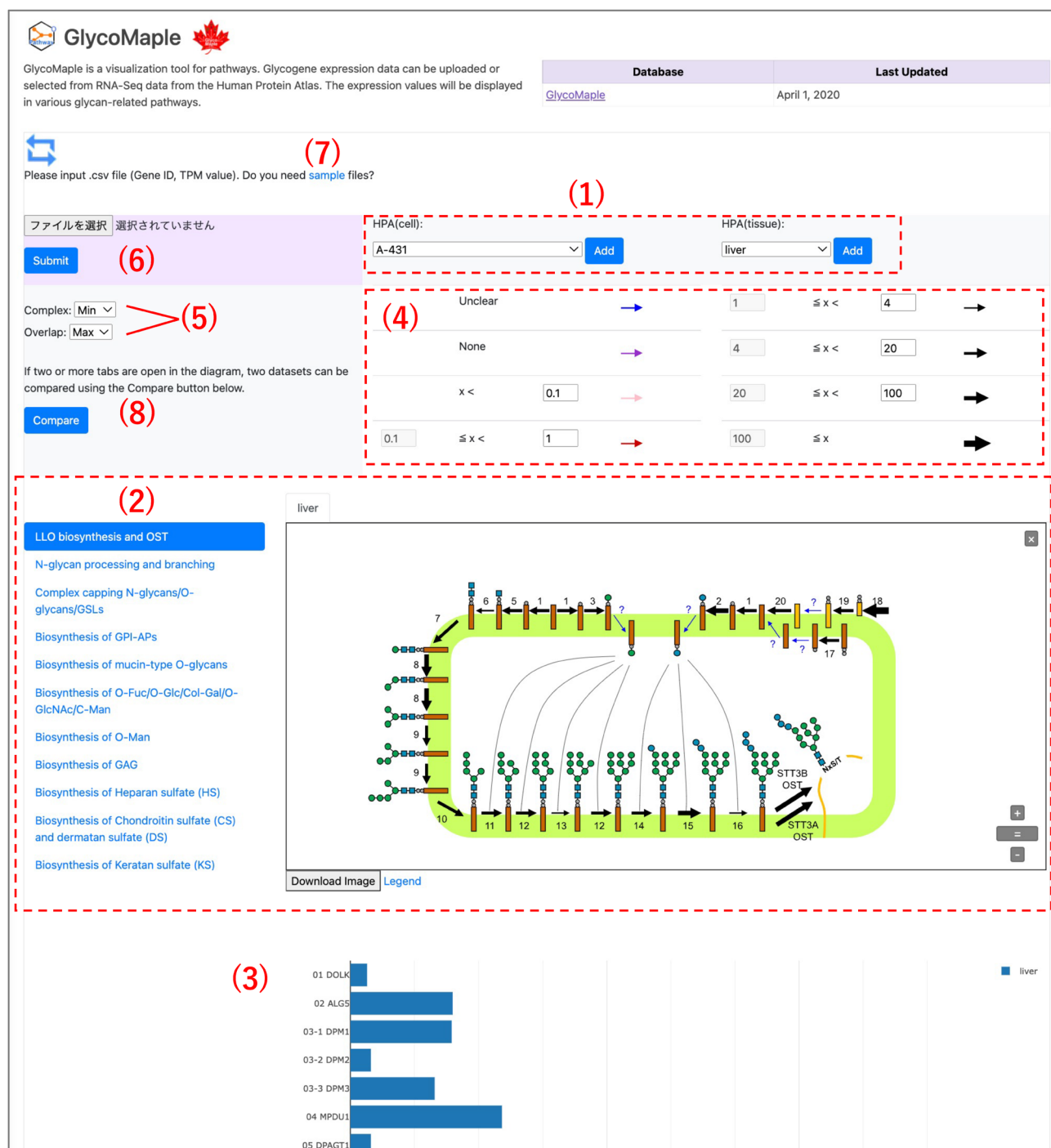


図1 GlycoMaple ツール画面

GlycoMaple ツールは、糖鎖関連遺伝子の発現量を糖鎖合成経路に可視化し糖鎖構造を推定できる。ツールへの入力 (1) の HPA の細胞や組織の選択、または (6) のユーザー独自のデータによる遺伝子の発現量のファイルから可能である。(2) の糖鎖合成経路に遺伝子の発現量が矢印の太さで反映される。(3) には、矢印の番号と対応する棒グラフが表示される。オプションとして、(4) では発現量に対する矢印の調整ができ、(5) では糖鎖合成に関与する遺伝子の複合体の発現レベル (Complex) や重複して同じ機能を持つ遺伝子の発現レベル (Overlap) について基準を設定できる。また、(8) では合成経路どうしの比較ができる機能が設置されている。(7) ではファイル入力のサンプルが入手できる。

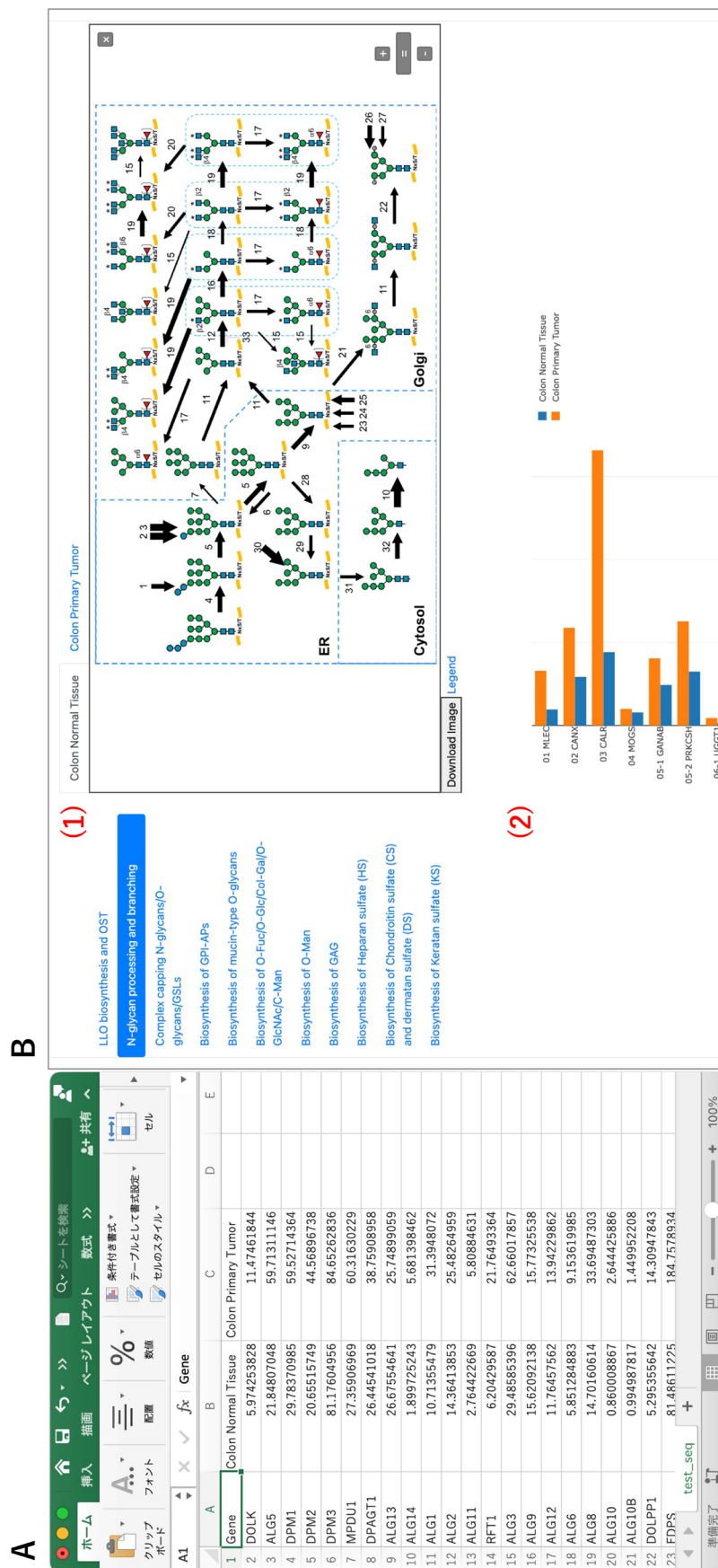


図2 GlycoMapleツールのファイル入力

(A) 図1の(7)で入手できるサンプルデータ。ファイルのA列に「Gene ID」、次のB列以降に「transcripts per kilobase million (TPM)」の値を入れる。サンプルでは、データセットが二つ入っており、B列に結腸の正常組織のTPM値、C列に結腸の腫瘍組織のTPM値が入力されている。1行目はラベルになっており、A1セルには「Gene」と入力し、B1, C1はツールの合成経路図に反映されるタブ名として表示される [図2B(1)]。 (B) (A)のファイル入力から「N-glycan processing and branching」経路をGlycoMapleで可視化した結果。(1)では入力したデータセット二つがタブで表示され、タブを切り替えるとそれぞれのデータのデータの合成経路を閲覧できる。(2)では、タブで表示される全データについて、各遺伝子の発現量が棒グラフとして同時に表示される。

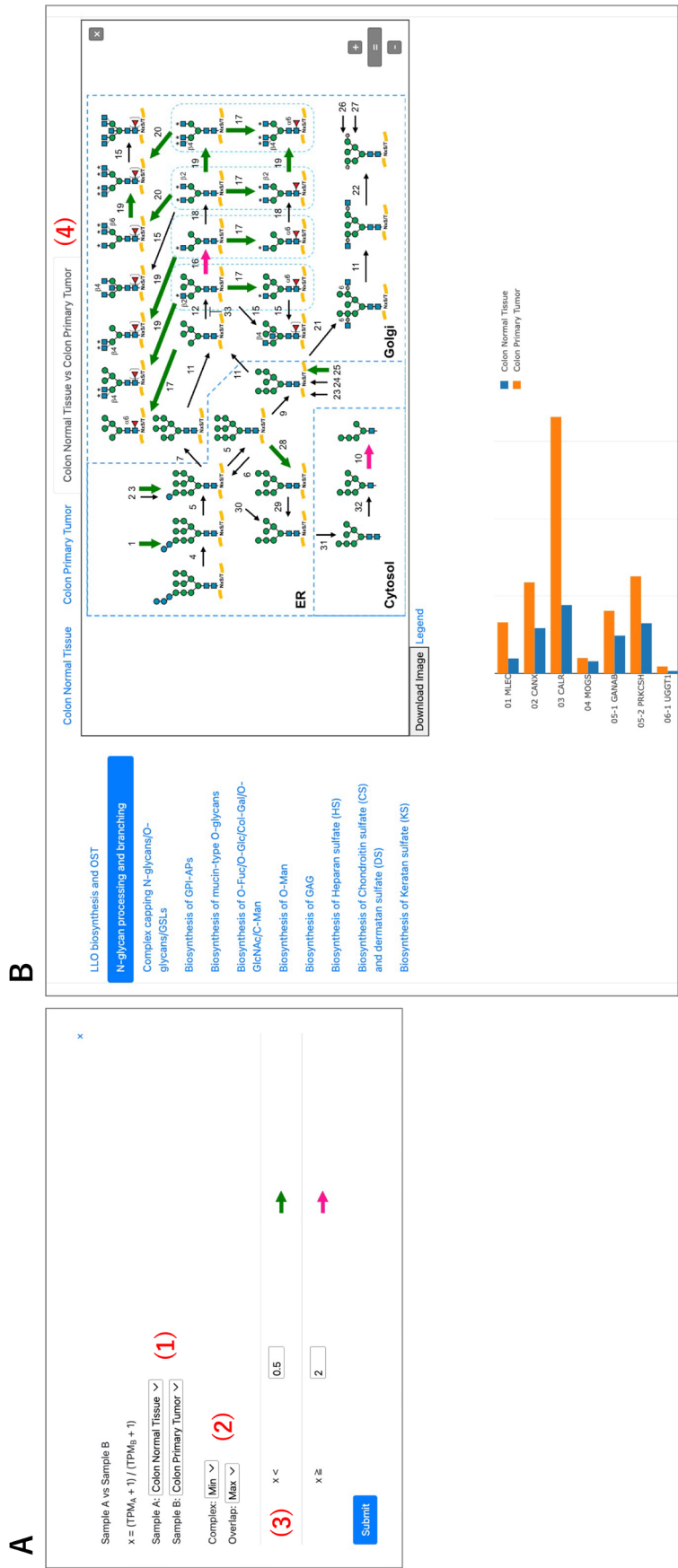


図3 GlycoMaple ツールの Compare 機能
(A) 図1(8)の「Compare」機能を設定する画面。(1)で比較するデータを選択し、(2)で「Complex」および「Overlap」の設定を行う。一つの合成経路図で二つの遺伝子の発現量を比較する際、表示する矢印の設定を(3)にて入力する。xは画面の2行目にある式を意味しており、初期設定では、緑の矢印はSample AがSample Bの半分より少ない発現量、ピンクの矢印はSample AがSample Bの2倍以上の発現量である場合、合成経路図にそれぞれの色が反映される(B)。(B)図2Aのファイル入力の二つのデータセットに対して「Compare」機能を用いて可視化した合成経路図。「Compare」機能を実行すると(4)のように「Colon Normal Tissue vs Colon Primary Tumor」のタブが新しく追加される。(A)で設定した結果として緑とピンクの矢印が表示され、発現量の比較が可視化される。

ii) 図1(2)の各糖鎖合成経路名を選択すると選択した細胞または組織の経路図が右に表示され、遺伝子発現レベルに基づいた経路を閲覧できる。図1は、“HPA (tissue)”の“liver”のデータを選び、“LLO biosynthesis and OST”の糖鎖生成経路を表示している。

①経路図中の各矢印の色や太さは発現量を表し、ヒト遺伝子命名法のサイト (Human gene nomenclature: <https://www.genenames.org/>) にリンクしている。

②経路図左下にダウンロード機能を設置しているため、表示している経路図をPNG形式で取得できる。

iii) 経路図を下にスクロールすると、図1(1)にて選択した細胞または組織における糖鎖合成経路に関わる矢印の各糖鎖関連遺伝子の発現量がグラフとして閲覧できる [図1(3)]。棒グラフは、RNA-seqのTPM (transcripts per kilobase million) 値を単位として表示している。プロット中の遺伝子名の前に記載した数字が経路図上の矢印の番号に対応している。

iv) ユーザーは、経路図の矢印の表示を図1(4)、(5)にて調整できるが、「Add」ボタンをクリックする前に設定する必要がある。

①図1(4)に表示されているxはTPM値の範囲を指し、ユーザーがxの値を設定することにより矢印の表示を調節することができる。

②青矢印“Unclear”は、合成経路において関わる遺伝子が明確になっていないことを示し、紫の矢印“None”はHPAにデータが存在していなかった場合を示す。

③図1(5)の“Complex”は、複数のタンパク質が複合体を形成して機能を獲得していることを意味し、“Overlap”はタンパク質が重複して機能を有していることを意味する。プルダウンメニューのパラメータ“Min”、“Max”、“Ave”、“Sum”は、それぞれ“Complex”または“Overlap”の遺伝子におけるTPM値の最小、最大、平均、合計を表す。初期設定として、“Complex”は“Min”、“Overlap”は“Max”にしている。

b. ファイル入力からの利用

GlycoMapleツールは、ユーザーが独自のRNA-seqデータを持っている場合、ファイル入力から利用できる。ここでは、ファイル入力および、二つの経路図を比較するCompare機能を紹介する。

i) 図1(6)のファイル入力では、CSV形式のファイルを用いる。そのため、ファイルの最初の列に「Gene ID」、次の列に「transcripts per kilobase million (TPM)」の値を入れたファイルを用意する。図2Aは、ツール上でCSVファイル入力のサンプルとして用意しているデータである。

①ツール上部の図1(7)の“sample”をクリックすると

test_seq.csvがダウンロードできる。CSVファイルはMacintosh, Windowsでのダウンロードおよびツール実行の入力に対応している。

②CSVファイルで用意するTPMの列は一つ以上のデータセットを記述できる。図2Aのtest_seq.csvでは、結腸の正常組織と結腸の腫瘍組織それぞれのデータを記述している。これら結腸組織のRNA-seqデータは、The Cancer Genome Atlas (TCGA) と Genotype-Tissue Expression (GTEx) の二つの公開データベースからの288個の結腸原発腫瘍サンプルと304個の正常結腸組織サンプルのRNA-seqデータから提供されている⁷⁾。

ii) 図1(6)の「Submit」ボタンをクリックすると結腸の正常組織および腫瘍組織の発現量が経路図に反映される。経路図の上方にCSVファイルのTPMの列の1行目に記述した“Colon Normal Tissue”、“Colon Primary Tumor”がタブで表示される [図2B(1)]。

①タブをクリックすることにより、それぞれの結腸組織を反映した糖鎖合成経路図を切り替えて表示できる。図2Bの経路図は、正常な結腸組織の“N-glycan processing and branching”経路を表示している。

②経路図の下には、タブで表示された二つのデータセットに対する棒グラフが同時に表示される [図2B(2)]。

③また、前述の細胞や組織の選択からツールを実行するのと同様に、矢印のパラメータ調節や“Complex”および“Overlap”のパラメータは「Submit」の前に調節可能である。図2Bでは、パラメータ値を初期設定の状態で行っている。

iii) 図1(8)の「Compare」をクリックすると図3Aが表示される。この機能は、経路図上のタブが二つ以上存在している状態で利用できる。

①Sample A, Sample Bのプルダウンメニューにて、比較したい組織を選択する。図3A(1)はSample Aを“Colon Normal Tissue”、Sample Bを“Colon Primary Tumor”に設定した。

②合成経路の糖鎖関連遺伝子の発現量について“Complex”や“Overlap”のパラメータを設定する [図3A(2)]。

③緑とピンク色の矢印は、経路図内の矢印の発現量の比較を表しており、ここでパラメータの設定ができる [図3A(3)]。初期設定として、緑の矢印は、Sample A : Colon Normal TissueがSample B : Colon Primary Tumorの1/2より少ない量であるとき表示され、ピンクの矢印は、Sample A : Colon Normal TissueがSample B : Colon Primary Tumorの2倍以上の量であるとき表示されるようになっている。

④図3Aの「Submit」ボタンを押すと新しく経路図の

タブに“Colon Normal Tissue vs Colon Primary Tumor”が表示される [図3B(4)]. タブをクリックすると, 「Compare」機能で設定した条件において図3Bの経路図のように緑やピンクの矢印が表示され, 発現量の違いを閲覧することができる.

3. GlycoMapleでできること

GlycoMapleを利用すると, 遺伝子発現プロファイルを紹介して糖鎖の代謝経路を視覚的に俯瞰することができる. 画像による代謝経路の大局的な観察は, 糖鎖構造の形成に関わる遺伝子が閲覧でき, 経路全体で糖鎖合成の触媒反応に関わる経路が複数みられるなど, 糖鎖形成の遺伝子の関わりが比較的わかりやすく, 簡単に情報が得られる. また, ツールであらかじめ用意されている細胞や組織の糖鎖合成の特性を理解することに役立ち, 糖鎖研究を始める糸口になると考えられる.

独自の実験データを用いたGlycoMapleの活用では, RNA-seqデータのTPM値に限らず, 遺伝子名とその遺伝子に付与する数値を入力して糖鎖プロファイルを予測的に推定することもできる. そのため, 糖鎖分析, 特定の疾患に対するバイオマーカーの探索, バイオ医薬品タンパク質

の研究などに貢献できると考えている⁷⁾.

文 献

- 1) Roth, J., Zuber, C., Park, S., Jang, I., Lee, Y., Kysela, K.G., Le Fourn, V., Santimaria, R., Guhl, B., & Cho, J.W. (2010) Protein N-glycosylation, protein folding, and protein quality control. *Mol. Cells*, **30**, 497–506.
- 2) Belický, Š., Katlík, J., & Tkáč, J. (2016) Glycan and lectin biosensors. *Essays Biochem.*, **60**, 37–47.
- 3) Schjoldager, K.T., Narimatsu, Y., Joshi, H.J., & Clausen, H. (2020) Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.*, **21**, 729–749.
- 4) Aoki, K., Perlman, M., Lim, J.M., Cantu, R., Wells, L., & Tiemeyer, M. (2007) Dynamic developmental elaboration of N-linked glycan complexity in the *Drosophila melanogaster* embryo. *J. Biol. Chem.*, **282**, 9127–9142.
- 5) Rodrigues, J.G., Balmaña, M., Macedo, J.A., Poças, J., Fernandes, A., de-Freitas-Junior, J., Pinho, S.S., Gomes, J., Magalhães, A., Gomes, C., et al. (2018) Glycosylation in cancer: Selected roles in tumour progression, immune modulation and metastasis. *Cell. Immunol.*, **333**, 46–57.
- 6) Magalhães, A., Duarte, H.O., & Reis, C.A. (2021) The role of O-glycosylation in human disease. *Mol. Aspects Med.*, **79**, 100964.
- 7) Huang, Y.F., Aoki, K., Akase, S., Ishihara, M., Liu, Y.S., Yang, G., Kizuka, Y., Mizumoto, S., Tiemeyer, M., Gao, X.D., et al. (2021) Global mapping of glycosylation pathways in human-derived cells. *Dev. Cell*, **56**, 1195–1209.e7.

著者寸描

●細田 正恵 (ほそだ まさえ)

創価大学糖鎖生命システム融合研究所助教. 博士 (工学).

■略歴 木下聖子教授のもとで創価大学工学部, 同大学院工学研究科博士前期課程および後期課程を経て2019年学位取得. 同年より創価大学理工学部助教. 21年より現職.

■研究テーマと抱負 レクチンが認識する糖鎖部位解明のための糖鎖構造マルチプルアラインメントアルゴリズムの開発および解析や糖鎖インフォマティクスを用いた研究を行っている.

●木下 聖子 (あおき-きのした きよこ)

創価大学糖鎖生命システム融合研究所副所長. Ph.D.

■略歴 1973年米国ミズーリ州カンザス・シティに生る. 96年米国ノースウエスタン大学コンピュータ科学の学士号と修士号を同時取得. 99年米国ノースウエスタン大学コンピュータ工学の博士号を取得. 台湾中央研究所のポストドクトラルフェロー, 米国バイオディスカバリーの上級ソフトウェア開発技術者, 京都大学化学研究所 (03年ポストドクトラルフェロー, 04年助手), 創価大学工学部 (06年講師, 08年准教授, 14年教授)を得て, 21年より現職.

■研究テーマと抱負 糖鎖とその他のオミクスデータのセマンティックウェブ化プロジェクトGlyCosmosの構築を始め, 糖鎖の機能解明を目指すためのバイオインフォマティクス技術の研究開発を行っている.

●藤田 盛久 (ふじた もりひさ)

岐阜大学糖鎖生命コア研究所教授. 博士 (理学).

■略歴 2006年筑波大学大学院生命環境科学研究科修了. 産業技術総合研究所糖鎖医工学研究センター, ジュネーブ大学生化学部門, 大阪大学微生物病研究所 (08年特任助教, 12年助教), 江南大学生物工程学院 (14年教授)を経て, 22年より現職.

■研究テーマと抱負 糖タンパク質 (特にGPIアンカー型タンパク質, N結合型糖タンパク質, 医薬糖タンパク質) の生合成, 細胞内輸送, 構造変化に着目し, 研究を行っている.