

GlyCosmos糖鎖科学ポータル の仕上げ開発 Putting the final touches on the GlyCosmos Glycoscience Portal

木下 聖子

Kiyoko F. Aoki-Kinoshita

糖鎖科学ポータルGlyCosmos(Yamada et al., 2020)は2019年度に公開され、糖鎖に関連するありとあらゆるオミクスデータを統合化し、本サイトから閲覧、検索、解析を可能にしている。ACGG-DBに格納されている糖タンパク質データベースGlycoProtDB(Kaji et al., 2017)や糖鎖遺伝子データベースGGDBをはじめ、これらのデータにGlyGen(York et al., 2020)の糖タンパク質情報やReactome(Good et al., 2021)のパスウェイ情報、UniProt(Bateman, 2019)のタンパク質に関する様々なメタデータを統合化している。また、レクチン情報に関しては本学独自で開発したMCAW-DB(Hosoda et al., 2018)もリンクしており、糖鎖アレイ実験で得られたレクチンが認識する糖鎖構造情報も可視化している。さらに、GlyCosmosにはリポジトリが設置しており、糖鎖構造を登録しアクセス番号が得られるGlyTouCan(Fujita et al., 2021)や質量分析関係のリポジトリUniCarb-DR(Rojas-Macias et al., 2019)およびGlycoPOST(Watanabe et al., 2021)にもアクセス可能になっている。図1にGlyCosmosポータルの概要を示しており、主に3つのセクションで構成されている: 標準、リポジトリおよびデータリソース。標準はGlyCosmosにおいて重要な要素であり、格納されたデータの標準形式やオントロジーを示しており、これらのデータをアクセスするために必要な情報である。リポジトリには、GlyTouCanやGlycoPOSTがアクセス可能になっている。データリソースがタンパク質や脂質、パスウェイ、疾患などの分類に整理されているが、相互リンクを通して関連性が明示的になっている。また、外部のリソースとの連携もしており、PubChemやProtein DataBank PDB、そしてGlySpace Alliance(Aoki-Kinoshita et al., 2020)へのリンクも用意している。

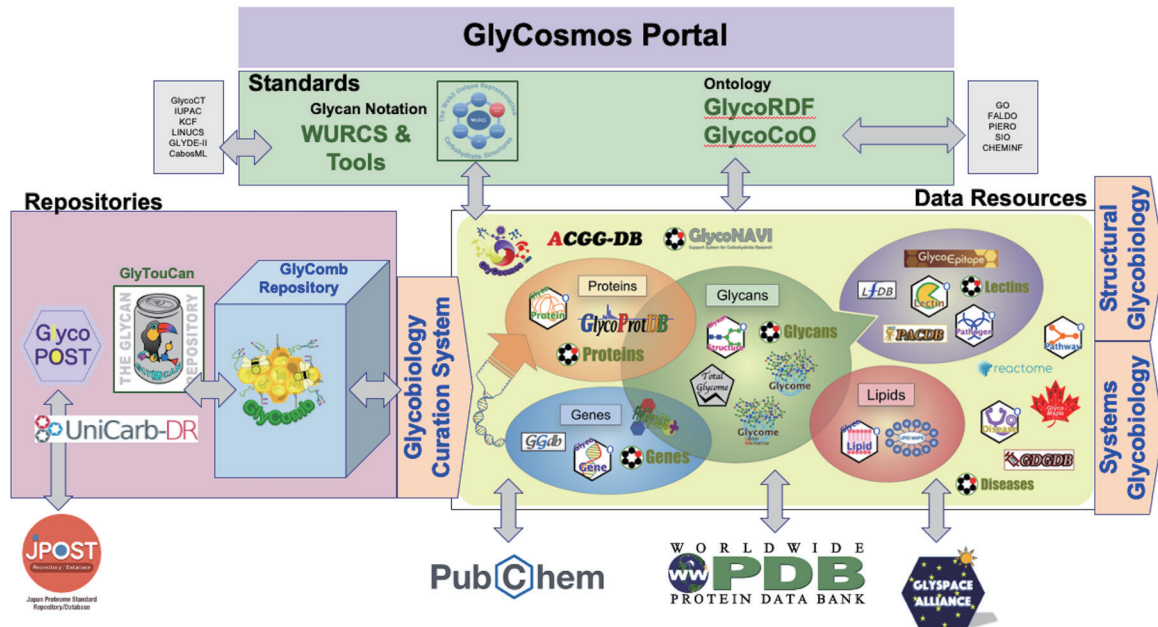


図1: GlyCosmosポータルの概要図。標準、リポジトリ、データリソースで構成されており、多様な糖鎖関連データを統合化している。

本年度は、GlyCosmosバージョン2にアップグレードし、新たな機能として糖鎖を検索するための様々なツールを提供している。単糖組成や新たな描画ツールSugarDrawer(Tsuchiya et al., 2021)を用いて、GlyTouCanから取得した糖鎖構造情報を目的別に検索が可能になっている。さらに、糖鎖遺伝子に関する疾患情報も統合化した(図2および図3)。なお、糖鎖を解析するために、糖鎖遺伝子の発現情報を入力し、合成される糖鎖構造を可視化するツールGlycoMaple(Huang et al., 2021)も公開した。本ツールでは、RNA-Seqなどの糖鎖関連酵素に関する数値情報を入力し、

様々な糖鎖合成経路(パスウェイ)上にその発現量を可視化する。これらのパスウェイに合成される糖鎖構造が表示されているため、遺伝子発現情報から合成される糖鎖構造を推測することができる。

本取り組みにより、GlyCosmos糖鎖科学ポータルプロジェクトの最終年度である本年度に、糖鎖科学者が知りたい情報の基盤が揃ったと言える。今後の発展として、複合糖質のリポジトリやレクチン情報の充実をはじめ、多くの情報を収集し、拡充させていく予定である。

GlyCosmos Diseases

List of diseases involving glycan related genes. The information of each database of Glyco-Disease Genes Database (GDGDB), DisGeNET, and Alliance of Genome Resources is integrated into one list.

Database	Last Updated
Glyco-Disease Genes Database (GDGDB)	January 25, 2017
Alliance of Genome Resources	February 1, 2021
DisGeNET	June 29, 2021

Show 10 entries

Cross Search

Disease ID	Disease Name	Gene Symbol	Gene ID
DOID:0040085	bacterial sepsis	DNM1L HDAC5	10014 10059
DOID:0050387	nonpapillary renal cell carcinoma	FHIT MTHFD2	10797 2272
DOID:0050424	familial adenomatous polyposis	KRAS	3845
DOID:0050441	mucosulfatidosis	SUMF1	285362
DOID:0050453	lissencephaly	POMGNT1 POMT1	100150181 10585

図2: GlyCosmosバージョン2で新たに導入した糖鎖関連疾患を統合化したリストGlyCosmos Diseasesである。Alliance of Genome ResourcesやDisGeNETより情報を収集し、リンクを提供している。このリストから、図3に示す疾患の詳細ページへ飛ぶことができ、詳細情報を閲覧することができる。

DOID:0050560

Summary	
DOID	DOID:0050560
Label	Walker-Warburg syndrome
Definition	A congenital muscular dystrophy that is characterized by hypotonia, seizures, severe intellectual and developmental disability, eye abnormalities and early death and has_material_basis_in mutations in multiple genes including POMT1, POMT2, ISPD, FKTN, FKRP, and LARGE1.
MeSH	D058494
Synonym	-

Contents
1. Summary
2. Related glycoproteins
3. Related glycoconjugates

Related glycoproteins

DisGeNET

Data not found.

Glyco-Disease Genes Database

UniProt	Protein Name
Q75072	Fukutin

Related glycoconjugates

DisGeNET

Data not found.

Glyco-Disease Genes Database

Gene ID	Gene Symbol	Gene Name
55624	POMGNT1	protein O-linked mannose N-acetylglucosaminyltransferase 1 (beta 1,2-)
2218	FKTN	fukutin

Alliance of Genome Resources

Gene ID	Gene Symbol	Gene Name	Organism
3565692	bgnt-1.5	-	Caenorhabditis elegans
182795	bgnt-1.4	-	Caenorhabditis elegans
557187	rxylt1	-	Danio rerio
569769	pomt1	-	Danio rerio
563878	pomt2	-	Danio rerio
571876	pomgnt1	-	Danio rerio
101669768	b4gat1	-	Danio rerio
33800	CG11149	-	Drosophila melanogaster

図3: 疾患の詳細ページ。Disease Ontology ID (DOID)が上部に表示されており、そのIDに基づいて抽出した疾患の説明文などがSummaryに表示している。その下に、関連糖タンパク質や遺伝子の情報が表示されている。



GlycoMaple is a visualization tool for pathways. Glycogene expression data can be uploaded or selected from RNA-Seq data from the Human Protein Atlas. The expression values will be displayed in various glycan-related pathways.

Database	Last Updated
GlycoMaple	April 1, 2020

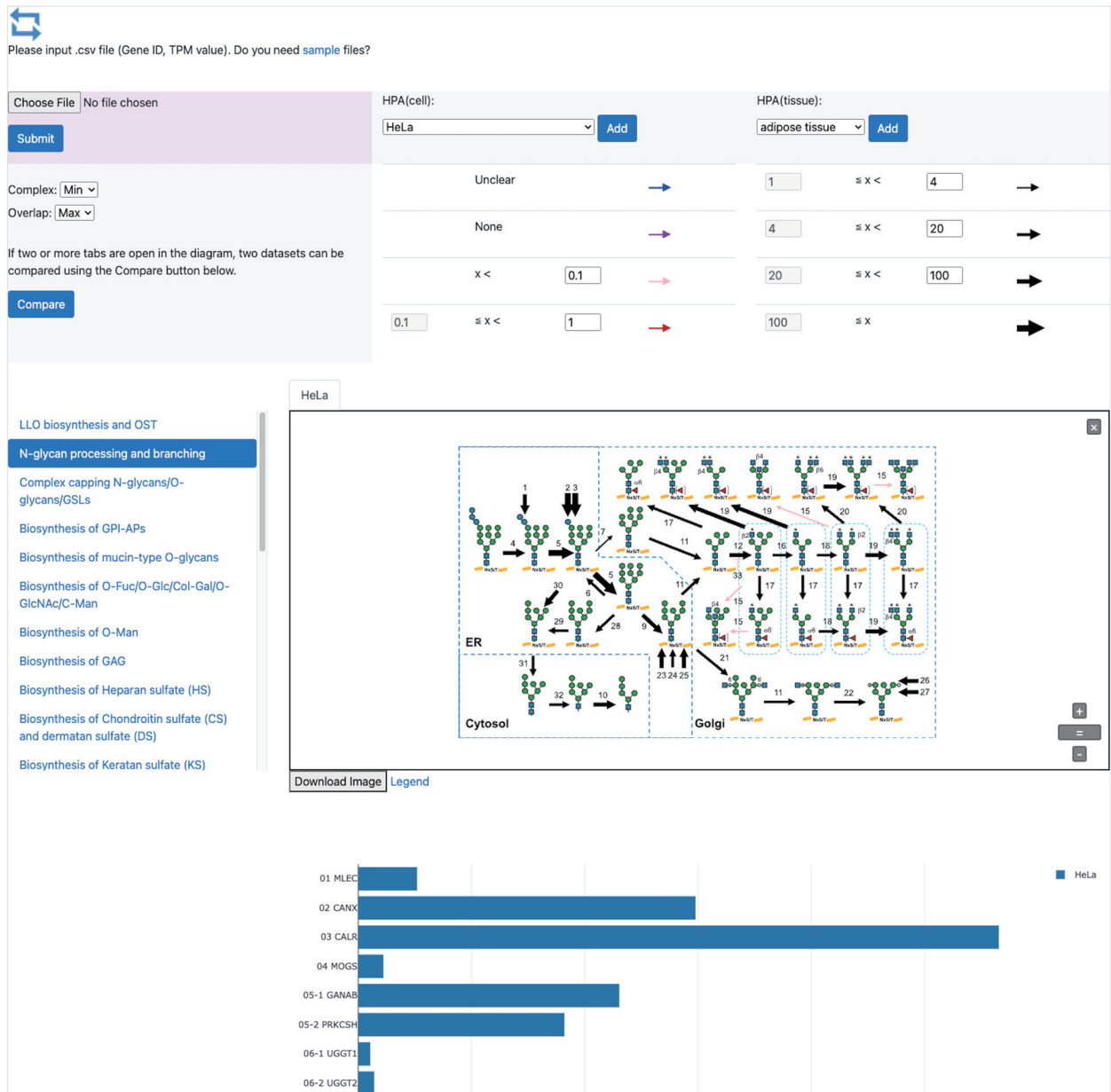


図4：GlycoMapleの画面。HeLa細胞における遺伝子発現情報を反映したN型糖鎖合成経路図を示す。

【謝辞】

本研究は、科学技術振興機構 (JST)、バイオサイエンスデータベースセンター (NBDC) ライフサイエンスデータベース統合推進事業「統合化推進プログラム」の助成を受けています。

【引用文献】

Aoki-Kinoshita, K. F., Lisacek, F., Mazumder, R., York, W. S., & Packer, N. H. (2020). The GlySpace Alliance: Toward a collaborative global glycoinformatics community. *Glycobiology*, 30(2).
<https://doi.org/10.1093/glycob/cwz078>

- Bateman, A. (2019). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research*, 47(D1), D506–D515. <https://doi.org/10.1093/nar/gky1049>
- Fujita, A., Aoki, N. P., Shinmachi, D., Matsubara, M., Tsuchiya, S., Shiota, M., Ono, T., Yamada, I., & Aoki-Kinoshita, K. F. (2021). The international glycan repository GlyTouCan version 3.0. *Nucleic Acids Research*, 49(D1), D1529–D1533. <https://doi.org/10.1093/nar/gkaa947>
- Good, B. M., Van Auken, K., Hill, D. P., Mi, H., Carbon, S., Balhoff, J. P., Albou, L.-P., Thomas, P. D., Mungall, C. J., Blake, J. A., & D'Eustachio, P. (2021). Reactome and the Gene Ontology: Digital convergence of data resources. *Bioinformatics (Oxford, England)*, 37(19), 3343–3348. <https://doi.org/10.1093/BIOINFORMATICS/BTAB325>
- Hosoda, M., Takahashi, Y., Shiota, M., Shinmachi, D., Inomoto, R., Higashimoto, S., & Aoki-Kinoshita, K. F. (2018). MCAW-DB: A glycan profile database capturing the ambiguity of glycan recognition patterns. *Carbohydrate Research*, 464, 44–56. <https://doi.org/10.1016/J.CARRES.2018.05.003>
- Huang, Y. F., Aoki, K., Akase, S., Ishihara, M., Liu, Y. S., Yang, G., Kizuka, Y., Mizumoto, S., Tiemeyer, M., Gao, X. D., Aoki-Kinoshita, K. F., & Fujita, M. (2021). Global mapping of glycosylation pathways in human-derived cells. *Developmental Cell*, 56(8), 1195–1209.e7. <https://doi.org/10.1016/J.DEVCEL.2021.02.023>
- Kaji, H., Shikanai, T., Suzuki, Y., & Narimatsu, H. (2017). GlycoProtDB: A Database of Glycoproteins Mapped with Actual Glycosylation Sites Identified by Mass Spectrometry. In *A Practical Guide to Using Glycomics Databases* (pp. 215–224). Springer Japan. https://doi.org/10.1007/978-4-431-56454-6_11
- Rojas-Macias, M. A., Mariethoz, J., Andersson, P., Jin, C., Venkatakrishnan, V., Aoki, N. P., Shinmachi, D., Ashwood, C., Madunic, K., Zhang, T., Miller, R. L., Horlacher, O., Struwe, W. B., Watanabe, Y., Okuda, S., Levander, F., Kolarich, D., Rudd, P. M., Wuhrer, M., ... Karlsson, N. G. (2019). Towards a standardized bioinformatics infrastructure for N- and O-glycomics. In *Nature Communications* (Vol. 10, Issue 1, pp. 1–10). Nature Publishing Group. <https://doi.org/10.1038/s41467-019-11131-x>
- Tsuchiya, S., Matsubara, M., Aoki-Kinoshita, K. F., & Yamada, I. (2021). SugarDrawer: A Web-Based Database Search Tool with Editing Glycan Structures. *Molecules* 2021, Vol. 26, Page 7149, 26(23), 7149. <https://doi.org/10.3390/MOLECULES26237149>
- Watanabe, Y., Aoki-Kinoshita, K. F., Ishihama, Y., & Okuda, S. (2021). GlycoPOST realizes FAIR principles for glycomics mass spectrometry data. *Nucleic Acids Research*, 49(D1), D1523–D1528. <https://doi.org/10.1093/nar/gkaa1012>
- Yamada, I., Shiota, M., Shinmachi, D., Ono, T., Tsuchiya, S., Hosoda, M., Fujita, A., Aoki, N. P., Watanabe, Y., Fujita, N., Angata, K., Kaji, H., Narimatsu, H., Okuda, S., & Aoki-Kinoshita, K. F. (2020). The GlyCosmos Portal: A unified and comprehensive web resource for the glycosciences. *Nature Methods*, 17(July), 649–650. <https://doi.org/10.1038/s41592-020-0879-8>
- York, W. S., Mazumder, R., Ranzinger, R., Edwards, N., Kahsay, R., Aoki-Kinoshita, K. F., Campbell, M. P., Cummings, R. D., Feizi, T., Martin, M., Natale, D. A., Packer, N. H., Woods, R. J., Agarwal, G., Arpinar, S., Bhat, S., Blake, J., Castro, L. J. G., Fochtman, B., ... Zhang, W. (2020). GlyGen: Computational and Informatics Resources for Glycoscience. *Glycobiology*, 30(2). <https://doi.org/10.1093/glycob/cwz080>