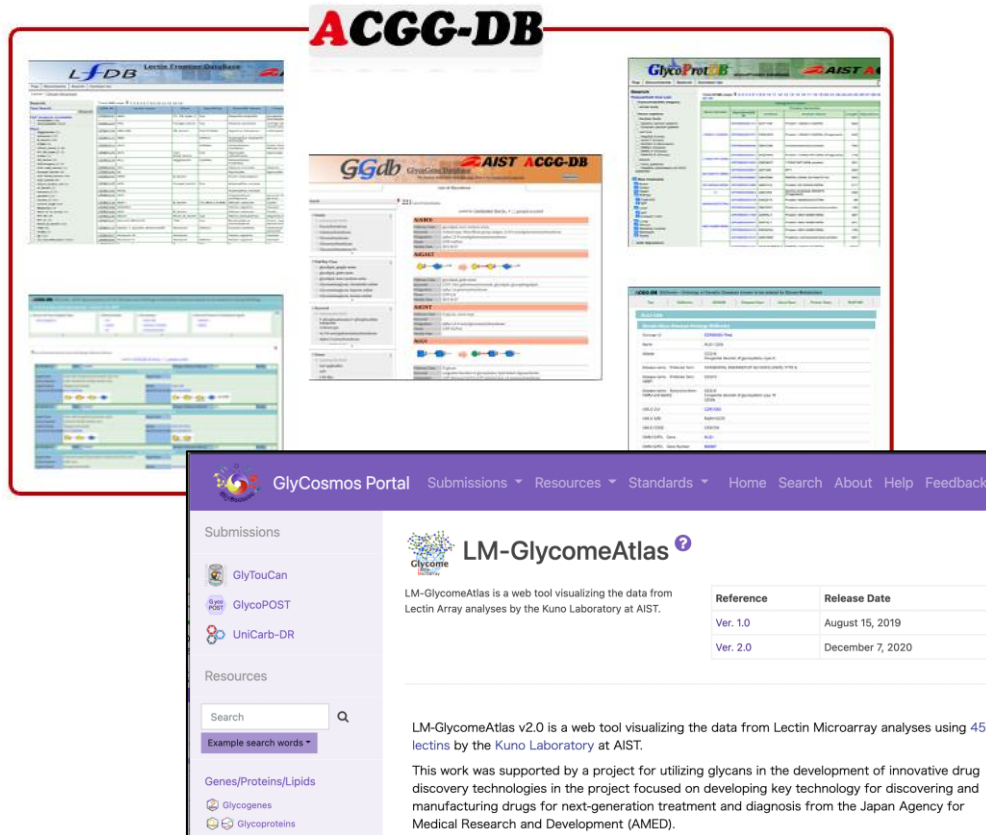


糖鎖関連データベースの連携強化 (ACGG-DB)

○藤田典昭^{1, 2)}、新町大輔^{1, 2)}、富永大介¹⁾、安形清彦^{1, 2)}、岡谷千晶¹⁾、成松久¹⁾、木下聖子²⁾、久野敦¹⁾

1. 産業技術総合研究所、2. 創価大学糖鎖生命システム融合研究所



糖鎖関連遺伝子(GGDB)、レクチン(LfDB)、糖タンパク質(GlycoProtDB)、糖鎖と感染症(PACDB)や糖鎖疾患遺伝子(GDGDB)といった糖鎖関連データベース(DB)や、レクチンマイクロアレイデータの可視化ツール(LM-GlycomeAtlas)を開発・公開し、各DBの共通言語である糖鎖構造に、共通のID(GlyCosmos GlycansやGlyTouCan ID)を紐づけることにより、DB間の相互連携を促進しています。

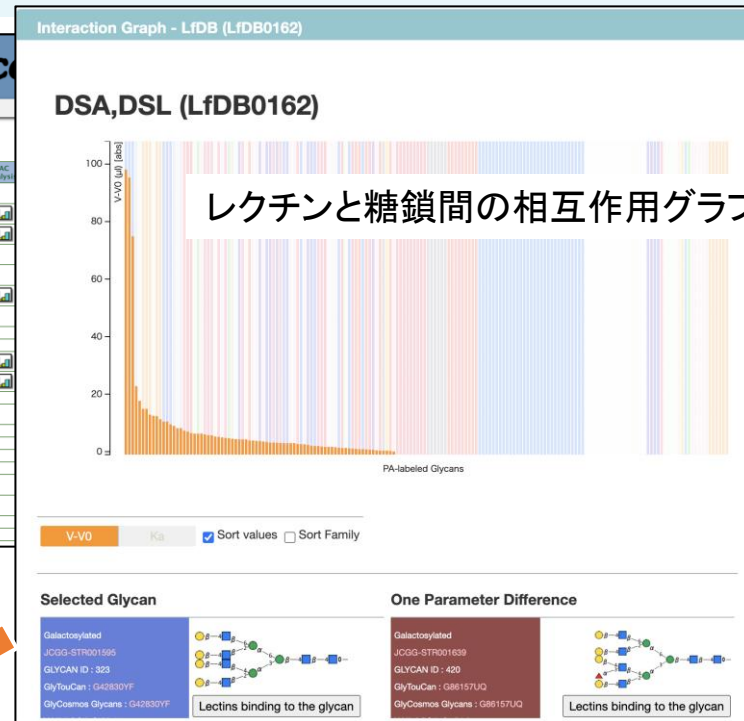
今回は、特に以下の改善を実施しましたので紹介いたします。

(1) LfDBとGlycoProtDB間で共通するレクチンによる横断を可能にしました。

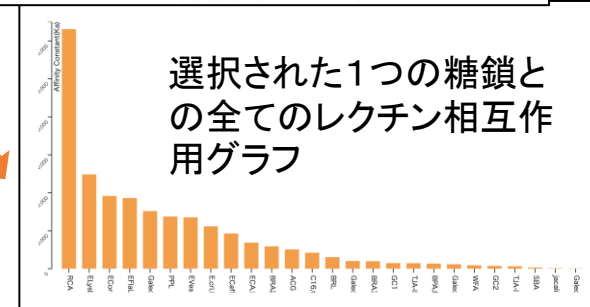
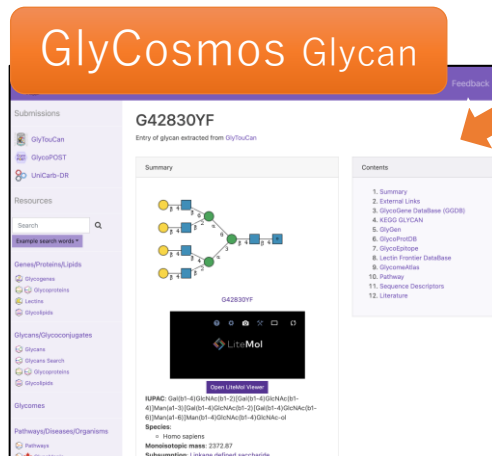
(2) GlycoProtDBでは糖鎖構造の多様性を可視化するビューワー(Glycofrom)を新たに搭載し、糖鎖付加部位特異的グライコム解析法(Glyco-RIDGE法)により取得したデータを表示可能としました。

(3) LM-GlycomeAtlas(ポスターNo.33)ではインターフェースの改良、新規データの追加を実施しました。

Lectin Frontier DataBase (LfDB)



- Lectin Frontier DataBase (LfDB) には、蛍光検出を用いた自動化フロンタルアフィニティークロマトグラフィーシステム(FAC-FD)を用いて取得した各種レクチンと糖鎖間の相互作用データを格納しています。
- 共通する糖鎖構造のIDを表記し、GlyCosmos Glycan へのリンクも提供しています。



37 GlycoProtein DataBase (GlycoProtDB)

The screenshot displays the GlycoProtDB website. The top navigation bar includes 'Top', 'Documents', 'Search', and 'Contact Us'. A search bar contains the text 'TFRC'. The left sidebar shows a 'Search' section with 'Tissue/Cell line List' and various filters for 'Homo sapiens' (e.g., Ascites fluids, Gastric cancer patient, Ovarian cancer patient, cell line, serum, HCC patients, Healthy volunteers (cf HCC patients)) and 'Mus musculus' (e.g., Brain, Colon). The main content area shows search results for 'TFRC' with a total of 3 pages. The first page displays a table of assigned proteins and sample(s) detected. The table has columns for Gene Symbol, GlycoProtDB ID, UniProt, Protein Name, Length, Glycoform, Organism, and Tissue. The results show two entries for TFRC: GPDB0010560 (Transferrin receptor (P90, CD71), isoform CRA_c) and GPDB0011338 (Transferrin receptor protein 1). Below the table, a detailed view for GPDB0011338 is shown, including a 'Glycosylation Sites (GPDB0011338)' section with a 'Reset' button and 'Zoom In'/'Zoom Out' controls. The section displays a sequence of the protein with N-glycosylation sites marked. The sites are labeled as 'Potential', 'Identified', and 'Identified with Glycoform'. The sequence is shown in a box with a 'mouse over to show the site and the motif sequence' instruction. Below the sequence, a 'Sequence / Glycosylation site(s) - GlycoProtDB (GPDB0011338)' section is visible, showing the full protein sequence and the glycosylation sites.

Assigned Protein					Sample(s) detected		
Gene Symbol	GlycoProtDB ID	UniProt	Protein Name	Length	Glycoform	Organism	Tissue
TFRC	GPDB0010560	G3V0E5	Transferrin receptor (P90, CD71), isoform CRA_c	679		Homo sapiens	cell line/HuH-7 (Liver) cell line/NUGC-4 (Stomach)
	GPDB0011338	P02786	Transferrin receptor protein 1	760		Homo sapiens	cell line/HuH-7 (Liver) cell line/NUGC-4 (Stomach)

Glycosylation Sites (GPDB0011338)

Reset Zoom In Zoom Out ToLeft ToRight

Tissue Lectin

cell line / HuH-7 medium AAL

cell line / HuH-7 medium DSA

cell line / NUGC-4 AAL(-)/RCA120(+)

N-glycosylation site(s) (: Potential, ↑ Identified, ↓ Identified with Glycoform)

1 200 400 600 760

mouse over to show the site and the motif sequence

identified peptide (mouse over to show peptide position)

Sequence / Glycosylation site(s) - GlycoProtDB (GPDB0011338)

Sequence

MMDQARSASF NLFGEPLSY TRFSLARQVD GDNHSEMKL AVDEEENADN NTKANVTKEP RCGSICYGT IAVIVFFLIG FMIGYLYGCK GVEPKTECER LAGTESPVRE EPGEDFPAAR RLYWDDLKRR LSEKLDSTDF TGTIKLLNEN SYVPREAGSQ KDENLALYVE NQREFKLSK VWRDQHFVKI QVKDSQNSV IIVDKNGRLV YLVENPGGYV AYSKAATVTG KLVHANFGTK KDFEDLYTPV GSGIVIVRAG KITFAEKVAN AESLNAIGVL IYMDQTKFPI VNAELSPFGH AHLGTGDPYT PGFSPFNHTQ FPPSRSSGLP NIPVQTSIRA AAEKLFQNM GDCPSDWKTD STCRMVTSSE KNVKLTVSNV LKEIKILNIF GVIKGFVEPD HYVVVGARQD AMGPAAKSG VCTALLKLA QMFSDMVLKD GFQPSRSIIF ASWSAGDFGS VGATEWLEGY LSSLHLKAPT YINLDKAVLG TSNFKVSASP LLYTLIEKTH QNVKHPVTGQ FLYQDSNWS KVEKLTLDNA AFFFLAYSGI PAVSFCEFD TDYPYLGTM DTYKELIERI PELNKVARAA AEVAGQFVHK LTHDVELNLD YERYNSQLLS FVRDLNQYRA DIKEMGLSLQ WLYSARGDFP RATSRLTDDF GNAEKTDREV MKKLNDVRMR VEYHFLSPYV SPKESPFPHV FWGSGSHTLP ALLENLKLKQ QNNGAFNETL FRNQLALATW TIQGAANALS GDVWDIDNEF

- GlycoProtein DataBase (GlycoProtDB)に格納されているデータは、糖鎖付加部位を同定できる方法 (IGOT法) から得られた実験データです。
- 現在、マウスの各種臓器、ヒトのがん細胞や肝がん患者と健常人のそれぞれの血清の糖タンパク質のN型糖鎖の付加位置情報を可視化しています。
- 糖ペプチドの捕集に使ったレクチンにより、その部位の持つ糖鎖モチーフが推定できます。
- 人の卵巣や胃などの組織のデータを追加しました。

- ・LfDB上のGlycoProtDBで公開しているデータに関するレクチンに、GlycoProtDBへのリンク付けを行いました。
- ・GlycoProtDBではレクチンを指定した検索が出来ます。

LfDB Lectin Frontier DataBase

Top Documents Search Contact Us

Lectin | Glycan Structure

Search

Text Search

Captured Glycoprotein

FAC Analysis Available

Pfam

Total: 398 page: 1 2 3 4 5 6 7 8 9 10 11 12 13 14

LfDB ID	Lectin name	Pfam	Specificity	Scientific Name	Common Name	Kingdom	FAC Analysis	GlycoProtDB Link
LfDB0045	AAA	F5_F8_type_C	Fuc	Anguilla anguilla	European eel, European freshwater eel	Animal		
LfDB0124	AAL	Fungal_lectin	Fuc	Aleuria aurantia	orange peel mushroom	Fungi		GPDB link
LfDB0126	ABA, ABL	FB_lectin	Gal, GlcNAc	Agaricus bisporus	cultivated mushroom	Fungi		
LfDB0115	ABD		Others	Aspergillus kawachii	IFO4308			
LfDB0221	ACA		GalNAc	Amaranthus caudatus	Grain Amaranthus, Inca Wheat, Tassel flower	Plant		
LfDB0125	ACG	Gal-bind_lectin	Gal	Agrocye cylindracea	Agrocye	Fungi		
LfDB0239	ACL	Agglutinin	GalNAc	Amaranthus cruentus				
LfDB0343	AKL			Aplysia kurodai	Aplysia	Animal		
LfDB0344	AL			Agrocye	Agrocye	Fungi		
LfDB0243	AMA	B_lectin		Arum maculatum				

GlycoProtDB GlycoProtein DataBase

Top Documents Search Contact Us

Search

Tissue/Cell line List

Caenorhabditis elegans

Homo sapiens

Mus musculus

Brain

Kidney

Fut9 KO

WT

Liver

WT

b4GalT-1 KO

Lung

Serum

Skeletal muscle

Stomach

Testis

with Glycoform

captured by Lectin: **AAL**

Total: 2263 page: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46

Gene Symbol	Assigned Protein		Length	Glycoform	Sample(s) detected	
	GlycoProtDB ID	Protein Variant(s)			Organism	Tissue
0105 (Fragment)	GPDB0015575	O19746	MHC class I antigen GN00105 (Fragment)	181	Homo sapiens	cell line/COLO205 (Colon)
1300017J02Rik	GPDB002147	D3Y36	Protein 1300017J02Rik	622	Mus musculus	Heart Kidney/Fut9 KO Liver/WT Serum Skeletal muscle
1300017J02Rik	GPDB002727	F6W4D3	Protein 1300017J02Rik (Fragment)	242	Mus musculus	Heart Kidney/Fut9 KO Liver/WT Serum Skeletal muscle
GPDB0006936	Q8VC96		Uncharacterized protein	700	Mus musculus	Heart Kidney/Fut9 KO Liver/WT

Glycosylation Sites (GPDB0011338)

Reset

Tissue

Lectin

cell line / HuH-7 medium

cell line / HuH-7 medium

cell line / NUGC-4

AAL

DSA

AAL(-)/RCA120(+)

N-glycosylation site(s) (: Potential, • Identified, • Identified with Glycoform)

Sequence / Glycosylation site(s) - GlycoProtDB (GPDB0011338)

Sequence

MMDQARSFAS NLFGEPLSY TRFSLARQVD GDNHVMKL AVDEENADN NTKANVTXPK RCGSGICYGT IAVIVFFLLG FMIGLYGCK GVEPKTECR LACTESPVRE EPEDFPAAR RLYWDLKRR LSEKLDSTDF TGTIKLLNEN SYVPREAGSQ KENLALYVE NQPREFKLSK VWRQHFVKI QVKDSQNSV IIVDKNGRLV YLVENPGGVV AYSKAATVTG KLVHANFGTK KDFEDLYTPV IGSIVIVRAG KITFAEKVAN AESNAIGVL IYMDQKFPPI VNAELSPFGH

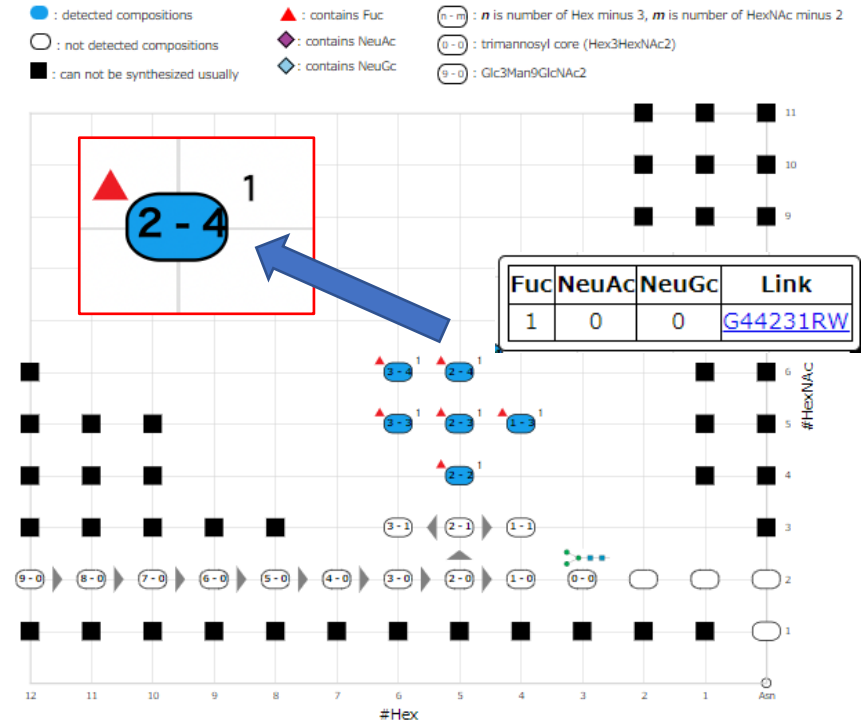
- Glycoform viewer は、糖タンパク質に対して、1つの糖鎖修飾位置の単糖組成の多様性を可視化したツールです。
- 横軸がヘキソース(Hex)、縦軸がヘキソサミン(HexNAc)の個数を示し、青丸は、実験で得られた単糖組成を示しています。※(Hex - HexNAc)
- 例で示す(2 - 4)は、(5 - 6)の単糖組成から N型糖鎖のコア構造の単糖組成(3 - 2)を引いた数を表しています。
- 付加情報として、フコース(Fuc)、シアル酸(NeuAc、NeuGc)の数や単糖組成に関連したGlyCosmosのリンクを表示します。
- Related Glycoproteinsのテーブルでは、GlycoProtDB内で関連のある糖タンパク質を表示しています。

N-glycosylation sites

Potential sites	Identified Peptide	Unique or Shared	Brain	Heart	Kidney	Kidney(Fut9 KO)	Kidney(Fut9 WT)	Liver: b4GalT-1(+/-)	Liver: b4
Position	Sequon	Position	Sequence						
140	NTT	139 - 152	CTTQNEVTSILR	Share (6)					
		137 - 152	TRCTTQNEVTSILR	Share (6)					
230	NRT	230 - 241	RTDVEYELDEK	Share (5)					
271	NRT								
303	NLT								
		423 - 450	VVDGEREVSVDYAHNNYQAQSAVPLR	Share (5)					
		423 - 464	VVDGEREVSVDYAHNNYQAQSAVPLRHETHGGEDVAFAK	Share (4)					
		423 - 439	VVDGEREVSMDVYAHN	Share (5)					
		423 - 440	VVDGEREVSMDVYAHNN	Share (5)					
		423 - 443	VVDGEREVSMDVYAHNNYQA	Share (5)					
		423 - 444	VVDGEREVSMDVYAHNNYQAQ	Share (5)					
		423 - 450	VVDGEREVSMDVYAHNNYQAQSAVPLR	Share (5)					
		423 - 464	VVDGEREVSMDVYAHNNYQAQSAVPLRHETHGGEDVAFAK	Share (4)					

*1: _ : Potential Sequon A: Asn (glycosylated) G: Gln (deaminated; pyroGlu) M: Met (oxidized) C: Cys (carbamidomethylated and deaminated)

Glycoform



Related Glycoproteins from GlycoProtDB

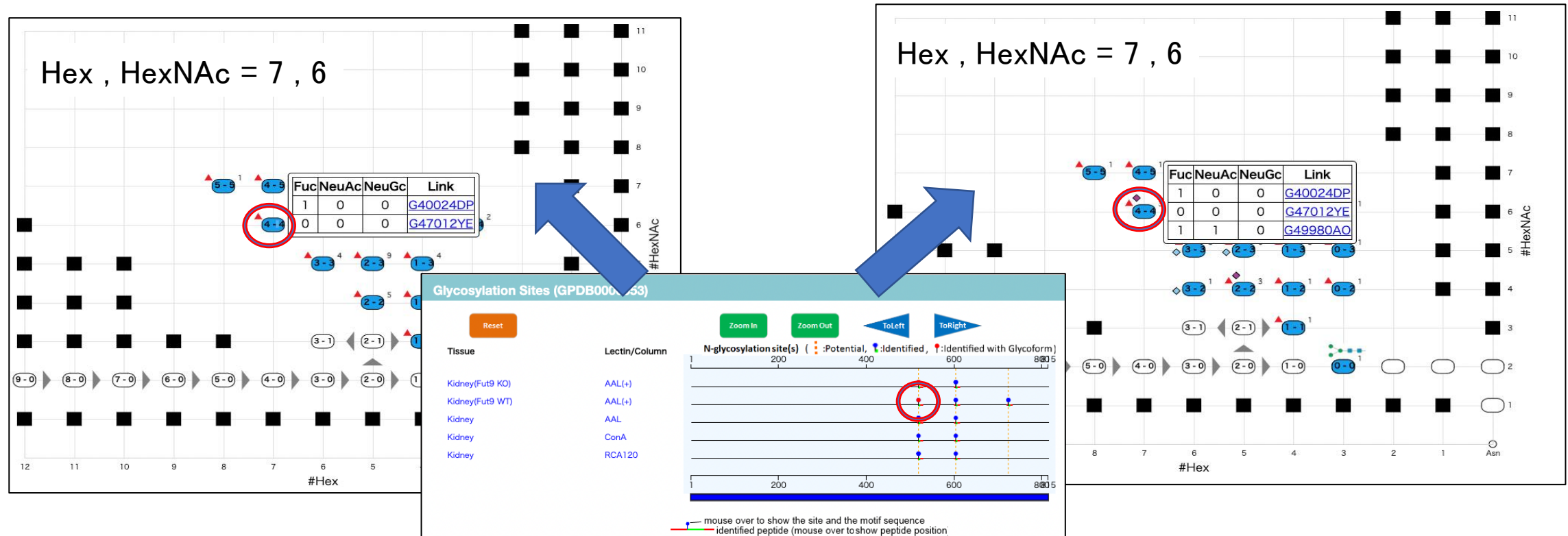
Protein Name	Protein Accession	Position	Lectin	Tissue	Taxon ID
Alkaline phosphatase	B7XGA6	430	AAL(+)	Kidney(Fut9 KO)	10090
Alkaline phosphatase	Q3TQ02	430	AAL(+)	Kidney(Fut9 KO)	10090
Alkaline phosphatase	Q8JZS4	430	AAL(+)	Kidney(Fut9 KO)	10090
Alkaline phosphatase	Q9EQ79	430	AAL(+)	Kidney(Fut9 KO)	10090
Alkaline phosphatase, tissue-nonspecific isozyme	P09242	430	AAL(+)	Kidney(Fut9 KO)	10090

Cadherin 16 の N519のサンプル間のGlycoform 比較

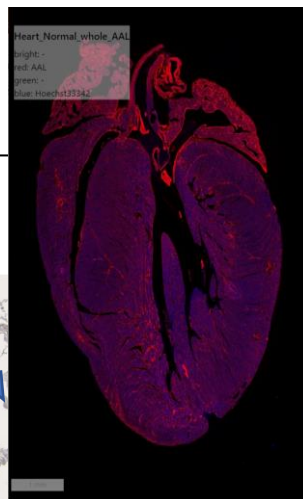
野生型(WT)とノックアウト(KO)のサンプルを比較することで、単糖組成の違いを確認することができます。

WT / AAL

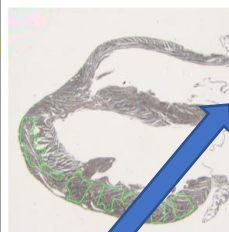
Fut9 KO / AAL



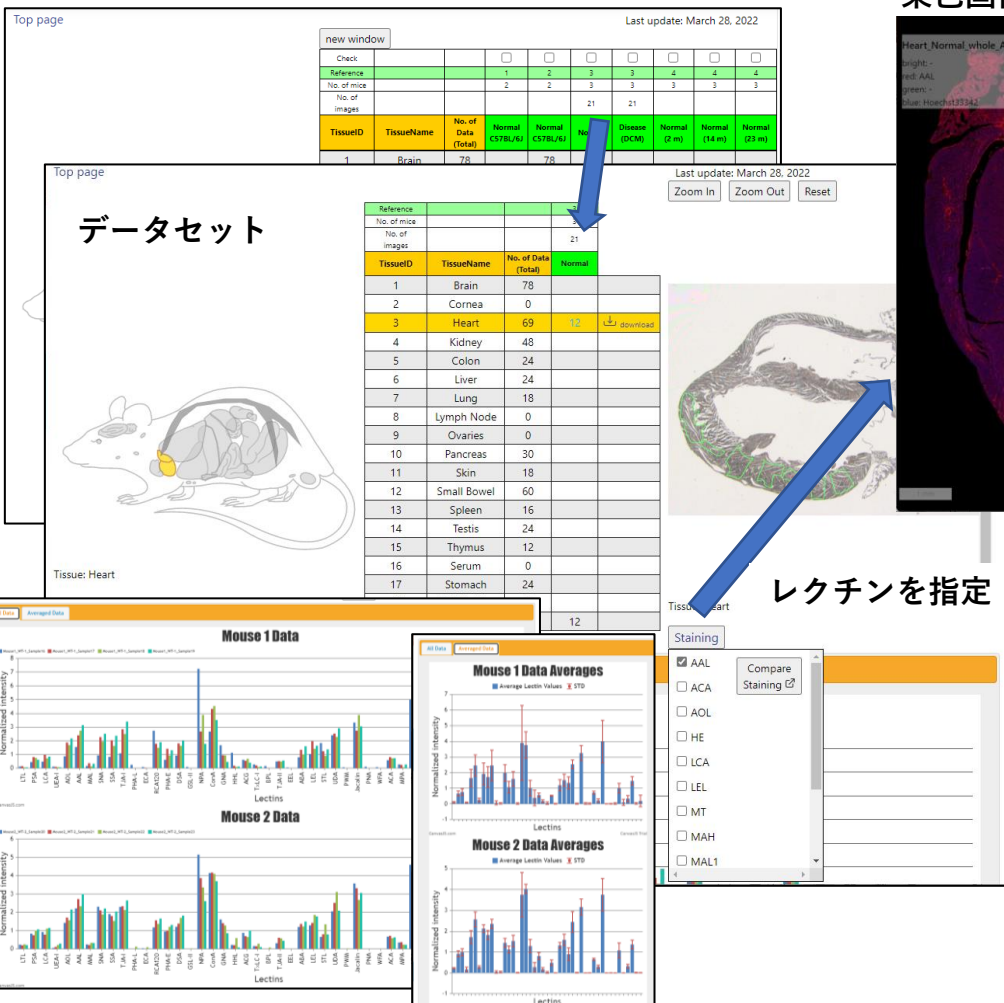
染色画像



レクチンを指定



- LM-GlycomeAtlasは、レクチンマイクロアレイ(45種類のレクチン)データの可視化ツールで、GlyCosmosにて公開しています。
- データセット単位でマウス組織ごとにデータが区別され、組織画像やレクチンマイクロアレイデータを表示しダウンロードして利用できます。
- 正常・疾患モデルマウスのレクチンマイクロアレイデータや組織のレクチン染色画像を格納しています。
- 東京都健康長寿医療センターのマウス組織データを公開しました。



37 Glyco Gene DataBase (GGDB)

The screenshot displays the GGDB interface for the ALG6 gene. The top section shows the gene name, its function (alpha-1,3-glucosyltransferase), and its involvement in the synthesis of lipid-linked oligosaccharides. Below this, a diagram illustrates the glycan structure of the enzyme. The bottom section lists the associated disease, Congenital Disorder of Glycosylation type 1c (CDG1c), and provides a link to the OMIM database.

ALG6
ALG6 encodes alpha-1,3-glucosyltransferase that is involved in the synthesis of lipid-linked oligosaccharides in the endoplasmic reticulum. Dolichol-phosphate-glucose is used for the donor substrate. ALG6 complements the defect of the yeast alg6 mutant. Mutation of ALG6 is associated with congenital disorder of glycosylation type 1c.

Keyword: congenital disorders of glycosylation, lipid-linked oligosaccharide

Diagram: A diagram showing the glycan structure of the enzyme, with a blue arrow indicating the transfer of a glucose unit from the donor substrate to the acceptor.

Table:

CAZy	GT57
GlyCosmos Glycogenes	ALG6
OMIM	604566
Disease name	CDG 1c, Congenital Disorder of Glycosylation type 1c
GDGDB	GN00345

Disease

GlyCosmos Glycan

G60230HH
Entry of glycan extracted from GlyTouCan

Summary

Diagram of the glycan structure G60230HH.

Contents

- Summary
- External Links
- GlycoGene DataBase (GGDB)
- KEGG GLYCAN
- Glycan
- GlycoProDB
- GlycoPathway
- Lectin Frontier Database
- GlycoMol
- Pathway
- Sequence Descriptors
- Literature

- Glyco Gene DataBase (GGDB)は、糖鎖関連遺伝子(合成)に関するデータベースです。
- 糖転移酵素の基質特異性、ヒトの遺伝子をベースにした糖鎖関連のオーソログ遺伝子などの情報を表示します。
- 共通した糖鎖構造を元に、GlyCosmos Glycanへのリンクも提供しています。
- 糖鎖関連遺伝子の変異による遺伝性疾患のデータベース (GDGDB) へのリンクも提供しています。
- 糖鎖遺伝子222件、情報追加: 65件、新規: 1件

Pathogen Adherence to Carbohydrate DataBase (PACDB) Glyco-Disease Genes DataBase (GDGDB)

PACDB

ACGG-DB Pathogen Adherence to Carbohydrate DataBase (PACDB)

Diseases known to be related to Glycan Binding

Helicobacter pylori

Diseases: Digestive System Disease, Tumor

Glycans and Glycoconjugates Types

- 42 Glycoconjugates
- 2 Non-glycosylated Lipids
- 16 Saccharides

Monosaccharides

- 21 D-Neup5Ac
- 3 D-Neup5Gc
- 15 Fuc

Glycoepitopes

- 1 2-3 Sialylparagloboside
- 1 2-6 Sialylparagloboside
- 1 Alpha-Galactosyl Epitope

Structural Features of Carbohydrate Ligands

- 16 Fucosylated
- 26 Sialylated
- 5 Sulfated

60 List of Interactions between Glycan and Pathogen Adherence Molecule.

sorted by: PACDB_REF_ID then by... grouped as sorted

PACDB REF ID	PMID	Pathogen Adherence Molecule	Glycan/Glycoconjugate Ligand	Target Source	Binding
4	10341176	bacteria adhesins/Leb-binding adhesin	Fucosylated Lewis b (Leb) histoblood group antigens	Lewis b	✓
25	12628483	bacterial glycan-binding proteins/Unknown	GM3(NeuAc)	N-Acetyl GM3	✓

GDGDB

ACGG-DB Glyco-Disease Genes DataBase (GDGDB)

Glycan Metabolism

Desc. Protein Desc. ENZYME

GDGDB ID	37
Disease names	Congenital Disorder of Glycosylation type Ic CDG Ic
Pathogenesis	Hypotonia, psychomotor retardation, seizures, strabismus, feeding problem, coagulopathy Glycoforum -Beyond Glycogenes, Dr. Hudson H. Freeze http://www.glycoforum.gr.jp/science/glycogenes/08/08E.html
OMIM phenotype	CONGENITAL DISORDER OF GLYCOSYLATION, TYPE Ic; CDG1C
OMIM phenotype mim number	603147
Gene names	asparagine-linked glycosylation 6, alpha-1,3-glucosyltransferase homolog (S. cerevisiae)
Official gene symbol	ALG6
Alias gene names	Man9GlcNAc(2)-PP-Dol alpha-1,3-glucosyltransferase dolichyl pyrophosphate Man9GlcNAc2 alpha-1,3-glucosyltransferase dolichyl-P-Glc:Man9GlcNAc2-PP-dolichylglucosyltransferase asparagine-linked glycosylation 6 homolog (S. cerevisiae, alpha-1,3-glucosyltransferase) asparagine-linked glycosylation 6 homolog (yeast, alpha-1,3-glucosyltransferase)
Locus	1p31.3
Link	GDGDB : http://cgdb.jp/cmg/ggdb/Homolog?cat=symbol&symbol=ALG6 NCBI Gene : http://www.ncbi.nlm.nih.gov/gene/29929

- PACDB: Pathogen Adherence to Carbohydrate DataBaseには、97の疾患名と127の病原体(370の株)についての情報が214の学術論文から抽出され、約1700件の糖鎖と病原体の相互作用のデータが収録されています。
- 感染症を引き起こす病原体のタンパク質(レクチン)とそのレクチンが結合する宿主(主にヒト)側の糖タンパク質や糖脂質(糖鎖リガンド)の情報が提供されています。

- Glyco-Disease Genes DataBase (GDGDB)は、糖鎖関連遺伝子の変異によって引き起こされる疾患の情報と遺伝子の情報を提供します。
- UniProtKB、ENZYME database等から糖転移酵素や糖鎖代謝に関与する酵素などの機能情報を取得できます。

本研究は、科学技術振興機構(JST)、NBDC事業推進部(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。