SOKA University

Discover your potential

Guide to using Glycomics Databases

Kiyoko F. Aoki-Kinoshita

15 July 2018 ICS Glycoinformatics Masterclass

Overview of lecture

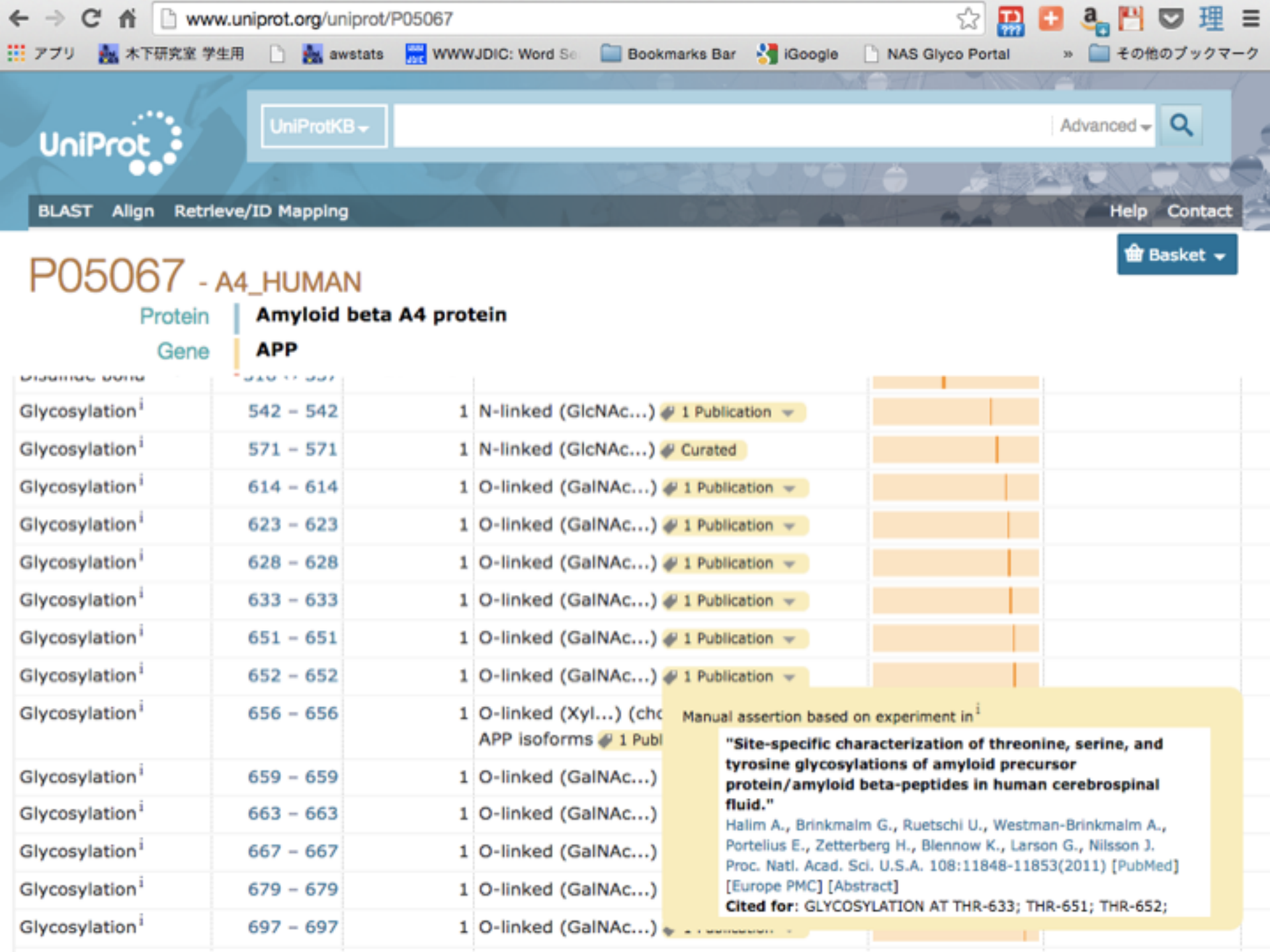
SOKA University

- The Semantic Web and Databases
- WURCS representation of carbohydrate structures
- Glyco-related database examples
- Applicability of the Semantic Web

Overview of lecture

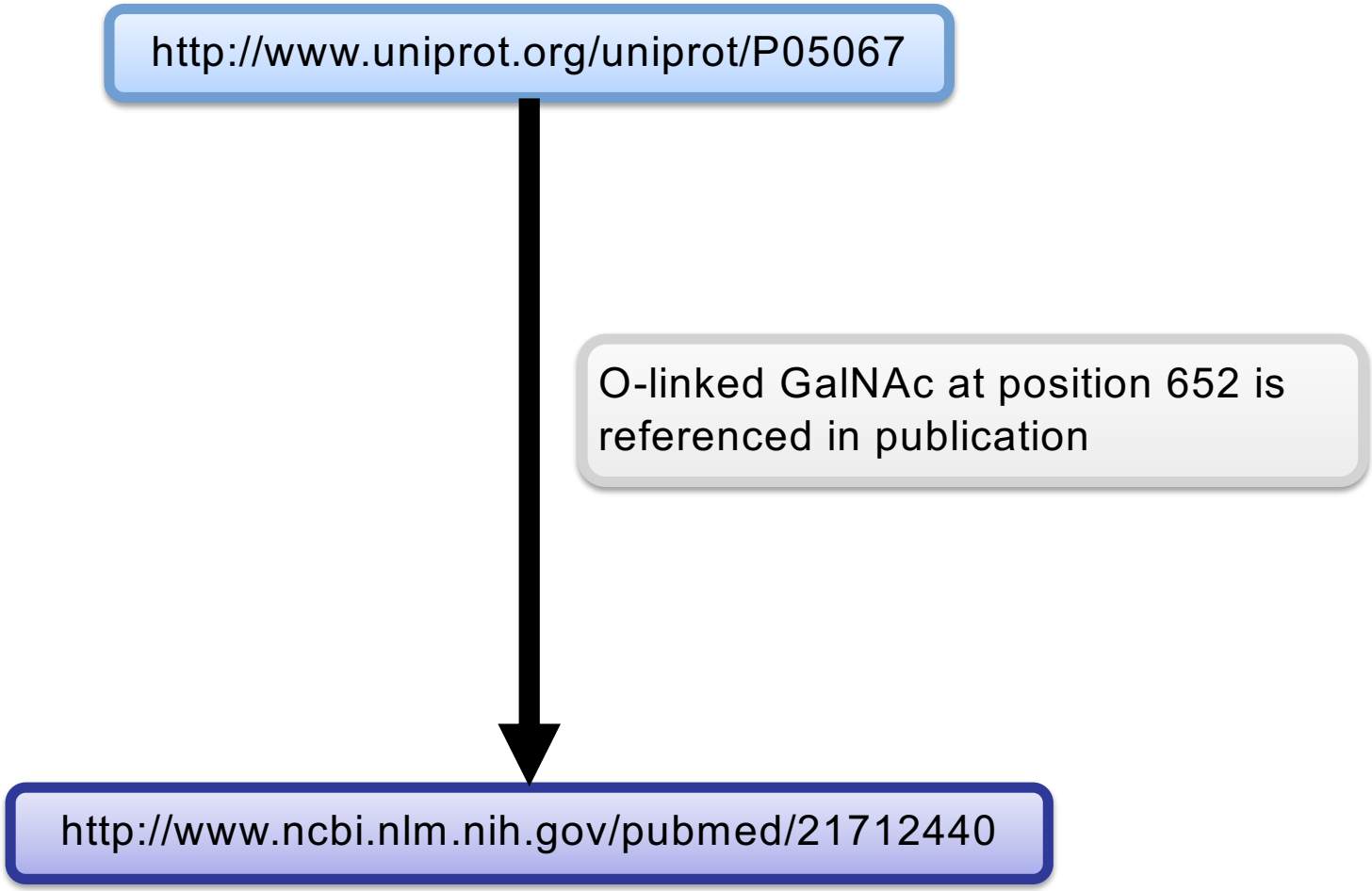
SOKA University

- The Semantic Web and Databases
- WURCS representation of carbohydrate structures
- Glyco-related database examples
- Applicability of the Semantic Web


www.uniprot.org/uniprot/P05067
UniProtKB
Advanced
BLAST Align Retrieve/ID Mapping
Help Contact
Basket
P05067 - A4_HUMAN
Protein Amyloid beta A4 protein
Gene APP
Glycosylation 542 – 542 1 N-linked (GlcNAc...) 1 Publication
Glycosylation 571 – 571 1 N-linked (GlcNAc...) Curated
Glycosylation 614 – 614 1 O-linked (GalNAc...) 1 Publication
Glycosylation 623 – 623 1 O-linked (GalNAc...) 1 Publication
Glycosylation 628 – 628 1 O-linked (GalNAc...) 1 Publication
Glycosylation 633 – 633 1 O-linked (GalNAc...) 1 Publication
Glycosylation 651 – 651 1 O-linked (GalNAc...) 1 Publication
Glycosylation 652 – 652 1 O-linked (GalNAc...) 1 Publication
Glycosylation 656 – 656 1 O-linked (Xyl...) (cho
APP isoforms 1 Publ
Glycosylation 659 – 659 1 O-linked (GalNAc...)
Glycosylation 663 – 663 1 O-linked (GalNAc...)
Glycosylation 667 – 667 1 O-linked (GalNAc...)
Glycosylation 679 – 679 1 O-linked (GalNAc...)
Glycosylation 697 – 697 1 O-linked (GalNAc...)
Manual assertion based on experiment in
"Site-specific characterization of threonine, serine, and tyrosine glycosylations of amyloid precursor protein/amyloid beta-peptides in human cerebrospinal fluid."
Halim A., Brinkmalm G., Ruetschi U., Westman-Brinkmalm A., Portelius E., Zetterberg H., Blennow K., Larson G., Nilsson J.
Proc. Natl. Acad. Sci. U.S.A. 108:11848-11853(2011) [PubMed] [Europe PMC] [Abstract]
Cited for: GLYCOSYLATION AT THR-633; THR-651; THR-652;

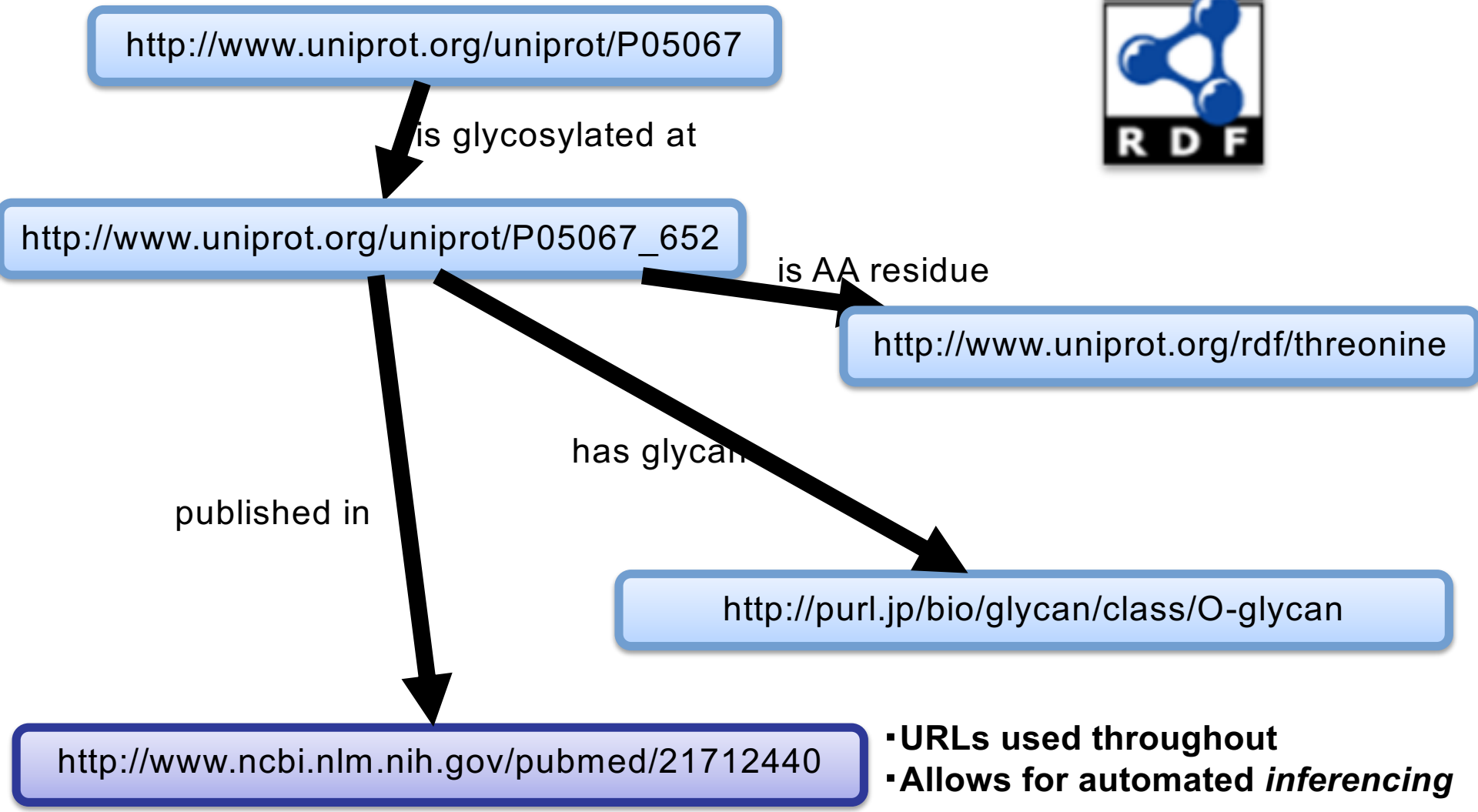
Semantic Web: human view

SOKA University



Semantic Web: computer view

SOKA University



- **URLs used throughout**
- **Allows for automated *inferencing***

Web services vs. SemWeb

SOKA University

What families are my CG numbers in?

Web Services
- Get FBgn number for CG no.

Server A (FlyGlycoDB)

CG No.	Name	FBgn No.
CG4084	dARG3	FBgn0011297
CG6850	dUGT	FBgn0014075
CG18799	dGM1	FBgn0259170
CG13431	dMGAT1	FBgn0034521
CG18802	dGM2	FBgn0011740

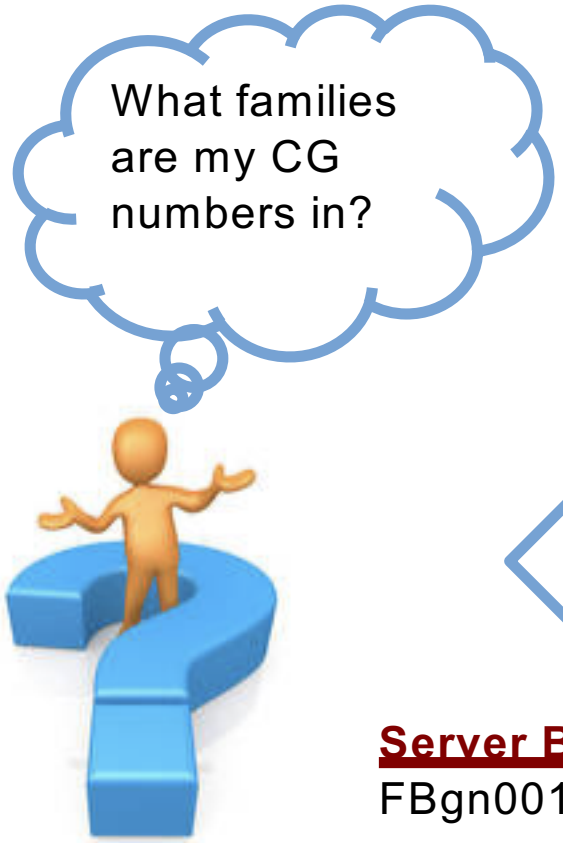
Web Services
- Get UniProt ID for FBgn number

Server B (FlyBase)

FBgn No.	Name	Family	UniProt ID
FBgn0011297	l(2)not	GT 58 Family	P82149
FBgn0014075	Ugt	GT 8 family	Q09332
FBgn0259170	α-Man-Ia	GH 47 family	P53624
FBgn0034521	Mgat1	-	-
FBgn0011740	α-Man-IIa	GH 38 family	Q24451

Web services vs. SemWeb

SOKA University



Server A (FlyGlycoDB)

CG4084	has_name	dARG3 .
CG4084	has_fbgn_id	FBgn0011297 .
CG6850	has_name	dUGT .
CG6850	has_fbgn_id	FBgn0014075 .
...		

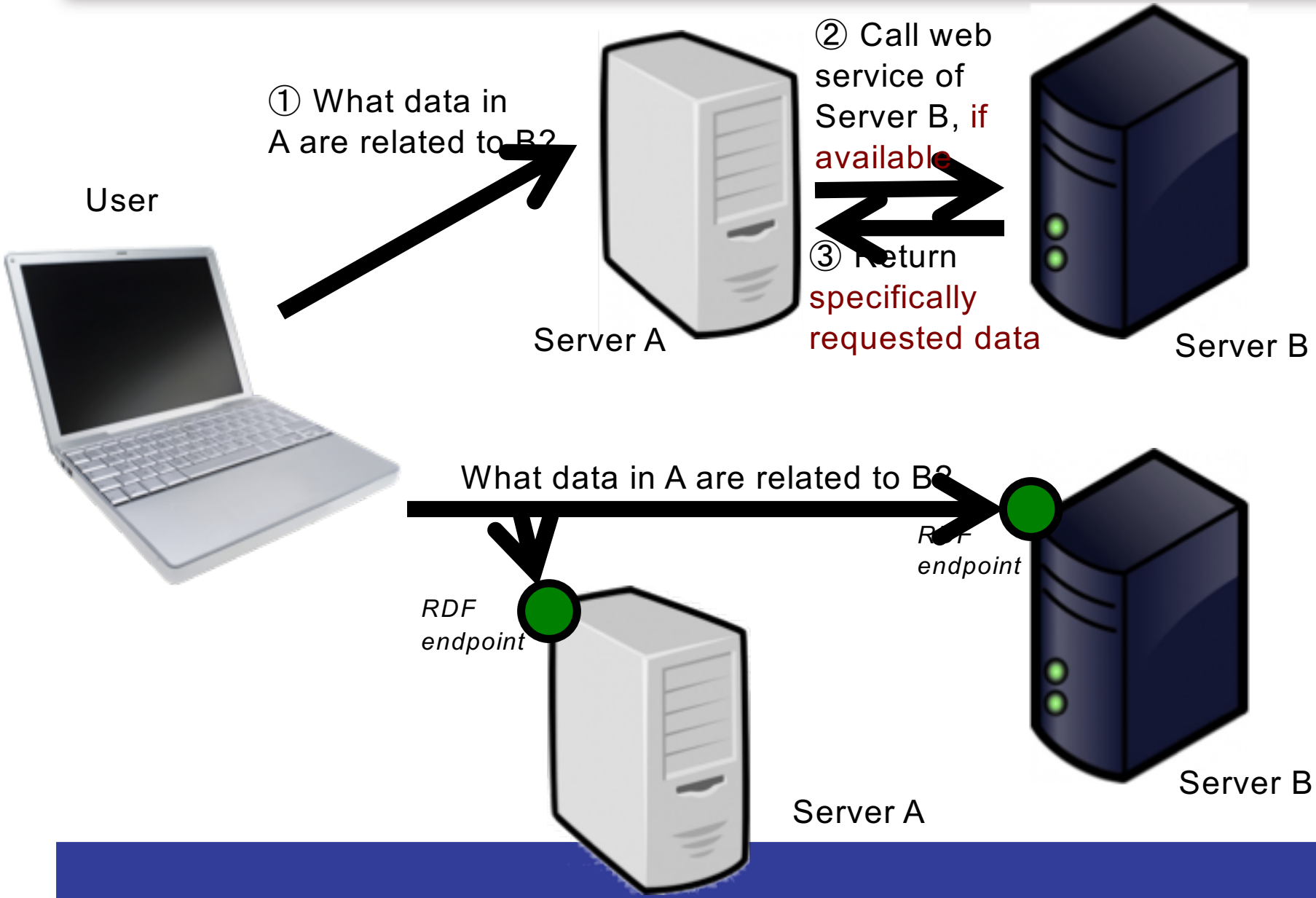
SPARQL!

Server B (FlyBase)

FBgn0011297	has_name	l(2)not ;
	has_family	GT 58 Family ;
	has_uniprot_id	P82149 .
FBgn0014075	has_name	Ugt ;
	has_family	GT 8 family ;
	has_uniprot_id	Q09332 .

Web services vs. SemWeb

SOKA University



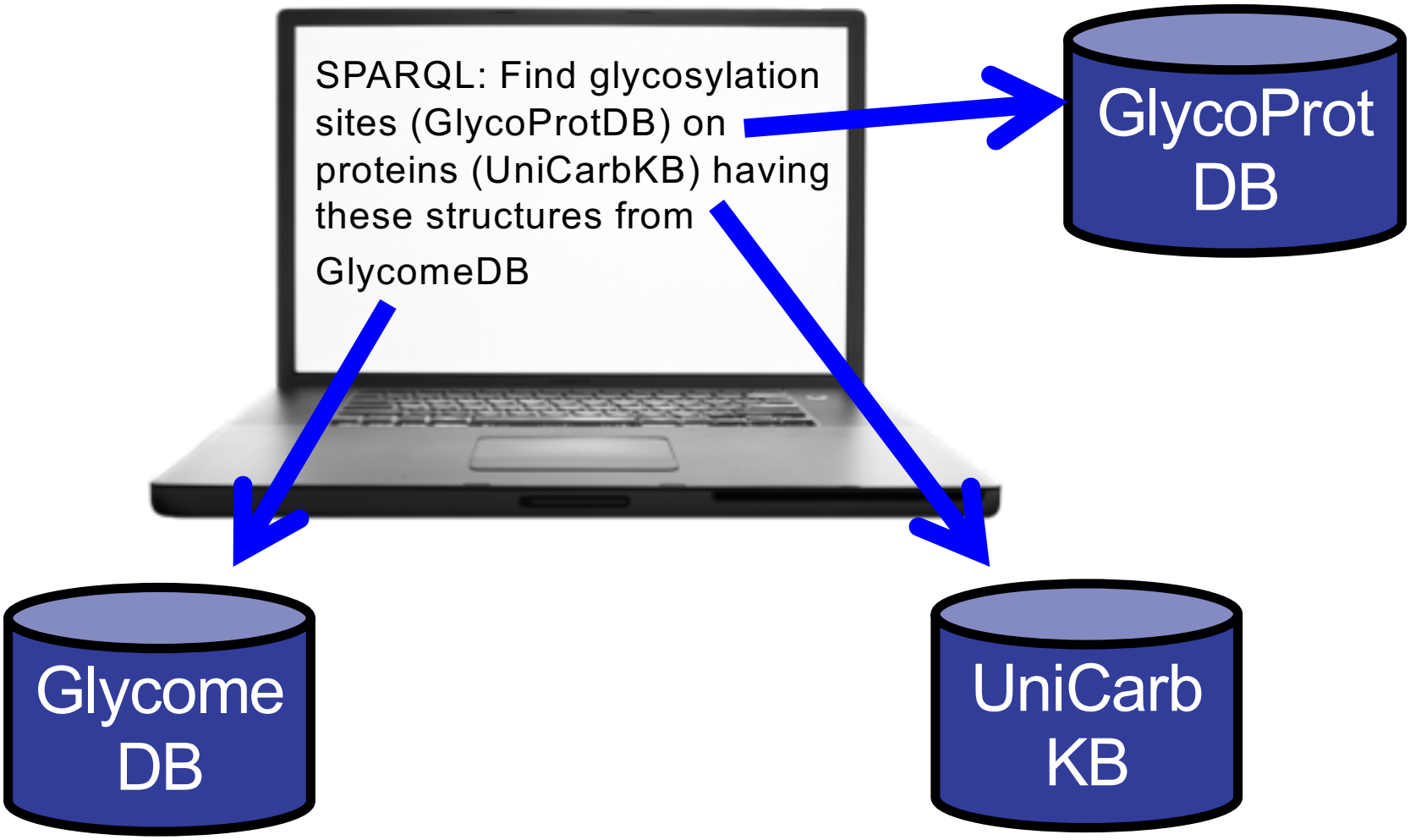
Web services vs. SemWeb

SOKA University

- Web services
  - Pros: allows other programs and web sites to obtain selected data available
  - Cons: many services need to be developed specifically to provide all data; lots of data may potentially need to be sent over the Internet
- SemWeb
  - Pros: allows flexible queries for all data in endpoint
  - Cons: requires an endpoint; the data model needs to be made clear and available; SPARQL language is not very user-friendly

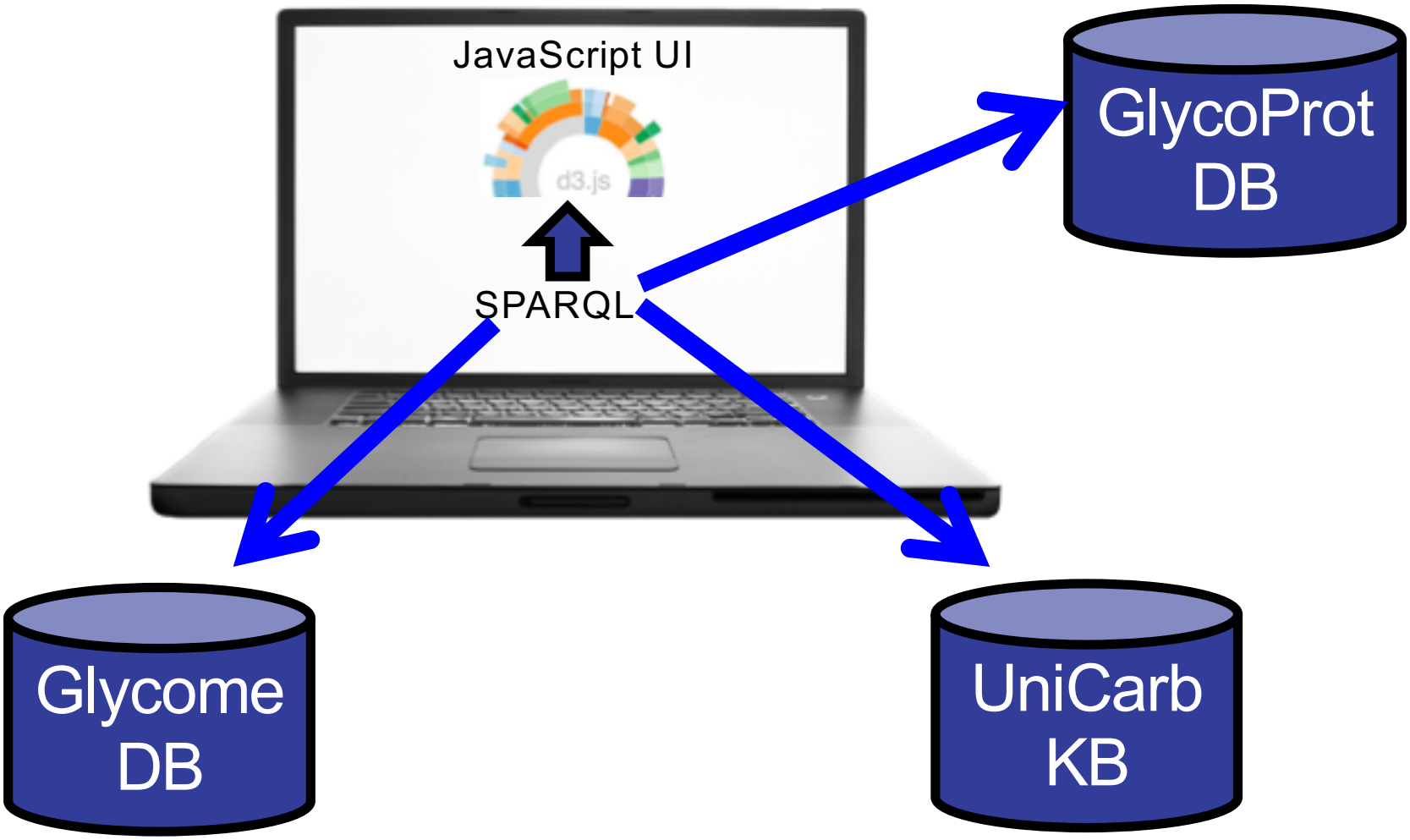
SPARQL queries

SOKA University



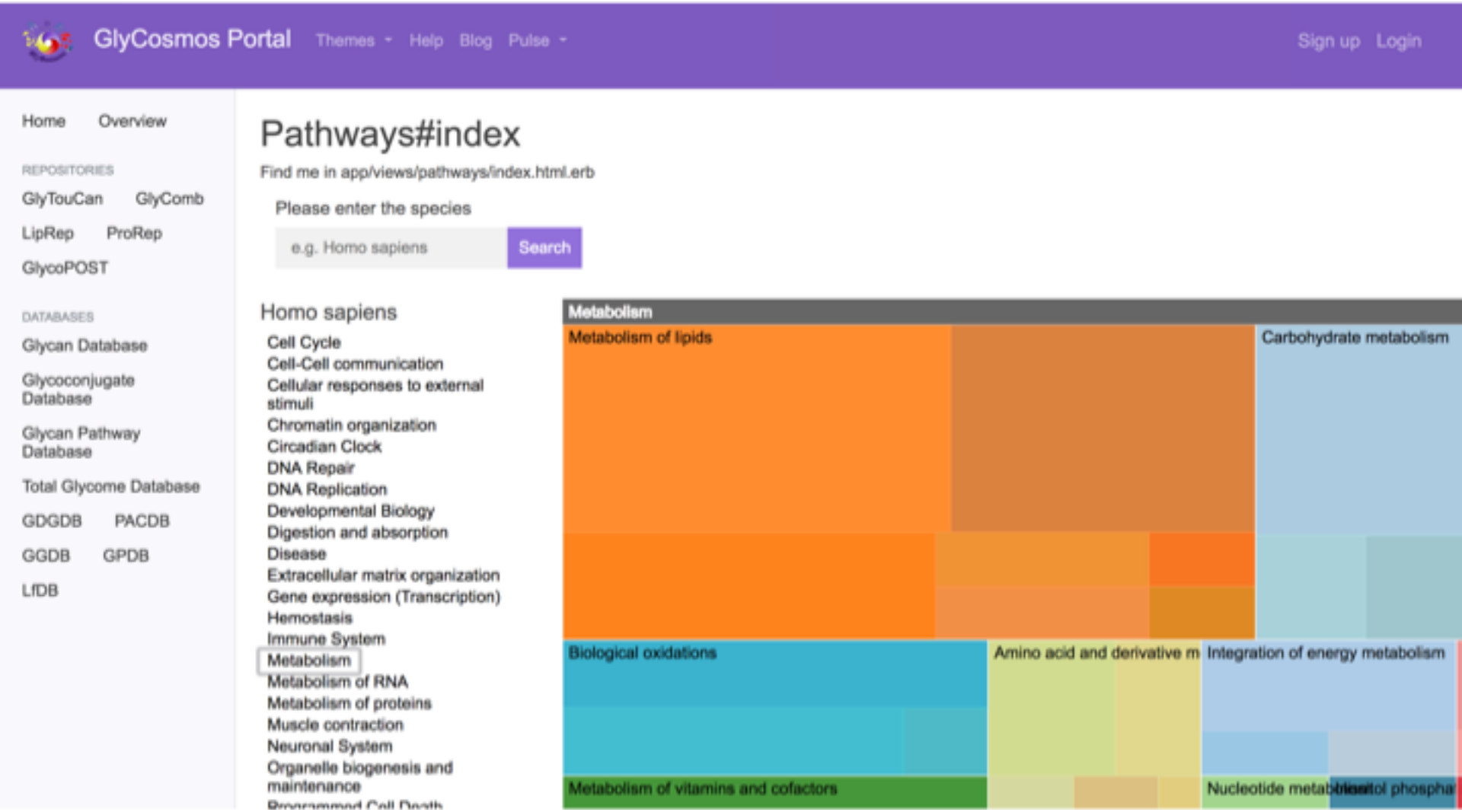
UIs for SemWeb: SPARQLists

SOKA University



UIs for SemWeb: SPARQLists

SOKA University



Overview of lecture

SOKA University

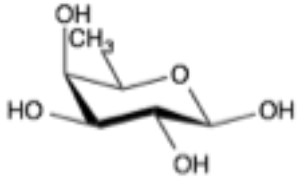
- The Semantic Web and Databases
- WURCS representation of carbohydrate structures
- Glyco-related database examples
- Applicability of the Semantic Web

Necessity for glycan data standardization

SOKA University

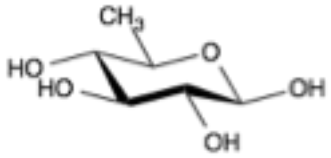
- Ambiguity of monosaccharides to start

D-Fucose
6-Deoxy-D-galactose



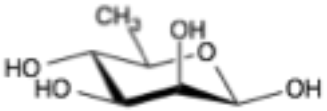
β-D-Fuc*p*

D-Quinovose
6-Deoxy-D-glucose



β-D-Qui*p*

D-Rhamnose
6-Deoxy-D-mannose



β-D-Rha*p*

- Bacterial and plant oligosaccharides?

Glycan structure formats

SOKA University

LINUCS

[][b-D-GlcpNAc]{[(4+1)][b-D-GlcpNAc]{[(3+1)][b-D-Manp]{}[(6+1)][b-D-Manp]{}}}

LinearCode

Mb3(Mb6)GNb4GNb

linear

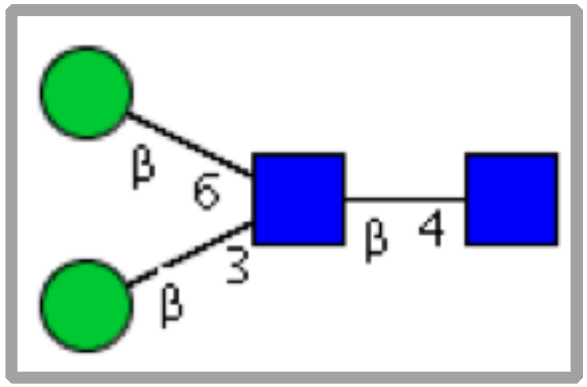
GlycoCT

```
RES
1b:b-dglc-HEX-1:5
2s:n-acetyl
3b:b-dglc-HEX-1:5
4s:n-acetyl
5b:b-dman-HEX-1:5
6b:b-dman-HEX-1:5
LIN
1:1d(2+1)2n
2:1o(4+1)3d
3:3d(2+1)4n
4:3o(3+1)5d
5:3o(6+1)6d
```

KCF format

```
ENTRY CT-1 Glycan
NODE 4
1 GlcNAc 0 0
2 GlcNAc -8 0
3 Man -16 2
4 Man -16 -2
EDGE 3
1 2:b1 1:4
2 4:b1 2:3
3 3:b1 2:6
///
```

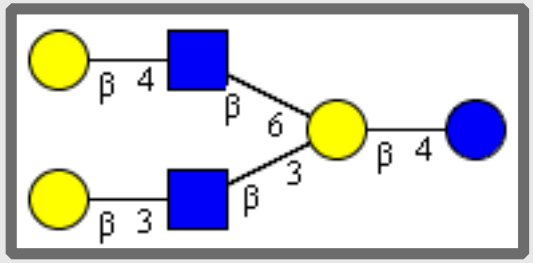
graph-based



Text formats for glycans

SOKA University

```
ENTRY   Structure  Glycan
NODE  6
    1  glcnac   -16   2
    2  glcnac   -16   -2
    3  gal   -24   2
    4  gal   -8   0
    5  gal   -24   -2
    6  D-glc   0   0
EDGE  5
    1  4:b1     6:4
    2  2:b1     4:3
    3  5:b1     2:3
    4  1:b1     4:6
    5  3:b1     1:4
///
```




```
ENTRY   Structure  Glycan
NODE  6
    1  GlcNAc   -16   -2
    2  Gal   -8   0
    3  GlcNAc   -16   2
    4  Gal   -24   -2
    5  D-Glc   0   0
    6  Gal   -24   2
EDGE  5
    1  2:b1     5:4
    2  1:b1     2:3
    3  4:b1     1:3
    4  3:b1     6:6
    5  3:b1     2:4
///
```

Uniquely represent all glycan structures as a text string

Web3 Unique Representation of Carbohydrate Structures*

- Tanaka et al., WURCS: Web3 Unique Representation of Carbohydrate Structures, J. Chem. Inf. Model., 2014, 54 (6), pp 1558–1566.
- Matsubara, M. et al., WURCS 2.0 update to encapsulate ambiguous carbohydrate structures. J. Chem. Inf. Model, 57(4):632-637. 2017. (doi: 10.1021/acs.jcim/6b00650) (PMID: 28263066)

WURCS

SOKA University

- Original purpose:
  - Develop a glycan representation that could be used as a URI on the Semantic Web
    - http://www.glytoucan.org/WURCS=2.0/3,5,6/xxx…
  - Single text string (no newlines)
  - Unique for ANY glycan sequence or composition, including ambiguous structures
- Why not existing formats?
  - GlycoCT, KCF are multi-line and could not describe substituents universally and consistently
  - IUPAC, LinearCode cannot cover the full spectrum of glycans/monosaccharides without extending them

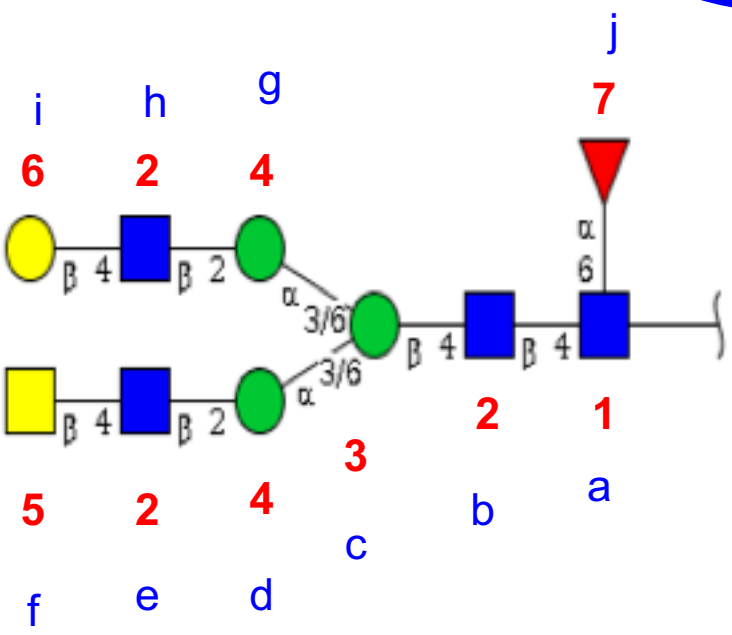
- Tanaka et al., WURCS: Web3 Unique Representation of Carbohydrate Structures, J. Chem. Inf. Model., 2014, 54 (6), pp 1558–1566.
- Matsubara, M. et al., WURCS 2.0 update to encapsulate ambiguous carbohydrate structures. J. Chem. Inf. Model, 57(4):632-637, 2017. (doi: 10.1021/acs.jcim/6b00650) (PMID: 28263066)

WURCS example

SOKA University

Glytoucan ID:
G00219DO

WURCS=2.0/7,10,9/[c2122h-1x_1-5_2*NCC/3=O][c2122h-1b_1-5_2*NCC/3=O][c1122h-1b_1-5][c1122h-1a_1-5][c2112h-1b_1-5_2*NCC/3=O][c2112h-1b_1-5][c1221m-1a_1-5]/1-2-3-4-2-5-4-2-6-7/a4-b1_a6-j1_b4-c1_d2-e1_e4-f1_g2-h1_h4-i1_d1-c3|c6_g1-c3|c6



WURCS=
2.0
7
10
9
[c2122h-1x_1-5_2*NCC/3=O]
[c2122h-1b_1-5_2*NCC/3=O]
[c1122h-1b_1-5]
[c1122h-1a_1-5]
[c2112h-1b_1-5_2*NCC/3=O]
[c2112h-1b_1-5]
[c1221m-1a_1-5]
1-2-3-4-2-5-4-2-6-7
a4-b1
a6-j1
b4-c1
d2-e1
e4-f1
g2-h1
h4-i1
d1-c3|c6
g1-c3|c6

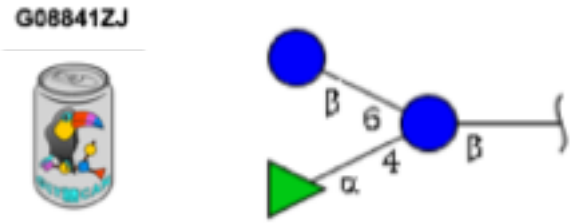
J. Chem. Inf. Model., 54(6):1558–1566, 2014, .

J. Chem. Inf. Model, 57(4):632-637, 2017.

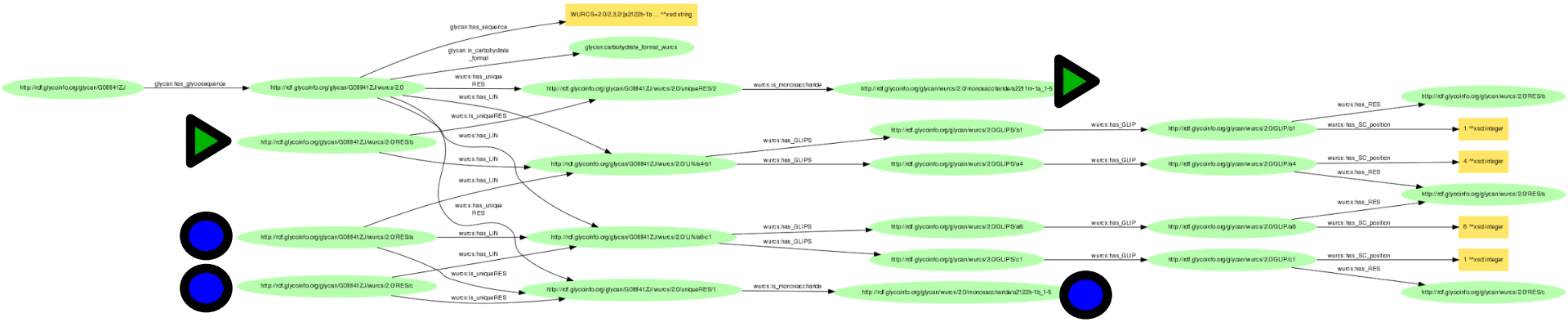
WURCS-RDF

SOKA University

- Can generate RDF data that describes a glycan via WURCS
- Can be used for substructure search of glycans using Semantic Web queries



WURCS=2.0/2,3,2/[a2122h-1b_1-5][a2211m-1a_1-5]/1-2-1/a4-b1_a6-c1

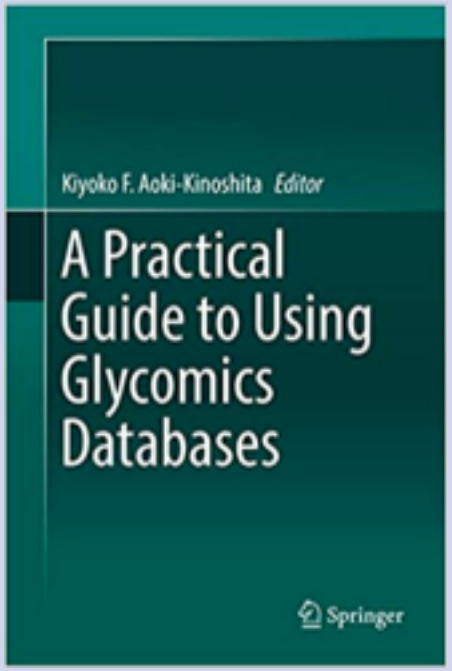


Overview of lecture

SOKA University

- The Semantic Web and Databases
- WURCS representation of carbohydrate structures
- Glyco-related database examples
- Applicability of the Semantic Web

A Practical Guide to Using Glycomics Databases



Open-access Databases for glycoscience just 10 years ago

SOKA University

Until 1996: CarbBank

- Glycosciences.de: 3D structures of glycans from PDB and 3D modeling
- BCSDB: bacterial carbohydrate structure database from the literature
- KEGG GLYCAN: CarbBank data cleaned up in terms of redundancy; computer algorithms for structure analysis
- Consortium for Functional Glycomics (CFG): CarbBank data plus some from the literature

Many databases now exist for glycoscience!

SOKA University

- JCGGDB, now ACGG-DB
  - Glycogenes, lectins, glycoproteins, glycan structures
- GlycoSuite
  - Curated glycans and glycoproteins from the literature
- GlycomeDB
  - Integrated database of seven major glycan structure databases
  - Now incorporated into GlyTouCan
- UniCarbKB
  - Curated glycans and glycoproteins from the literature
- GlyConnect
  - Glycan-related data connecting glycans with proteins and epitopes, etc.
- GlyTouCan: the international glycan structure repository

Integrating glycan databases

SOKA University

- The Database Center for Life Sciences (Japan), sponsored by the Japan Science and Technology agency (JST) and the National Bioscience Database Center (NBDC) has been funding projects to integrate life science databases since 2006
- Glycomics has been a part of this project since the start
- Genomics, Proteomics and many omics projects are participants
- Semantic Web technologies are used to integrate heterogeneous data

Integrated Database Project of Japan

SOKA University

- Sponsored by Japan Science and Technology Agency (JST) and the National Bioscience Database Center (NBDC)
- **Requirement: capability to provide data after funding ends!**
- http://biosciencedbc.jp/en/

Glycoscience Team Projects

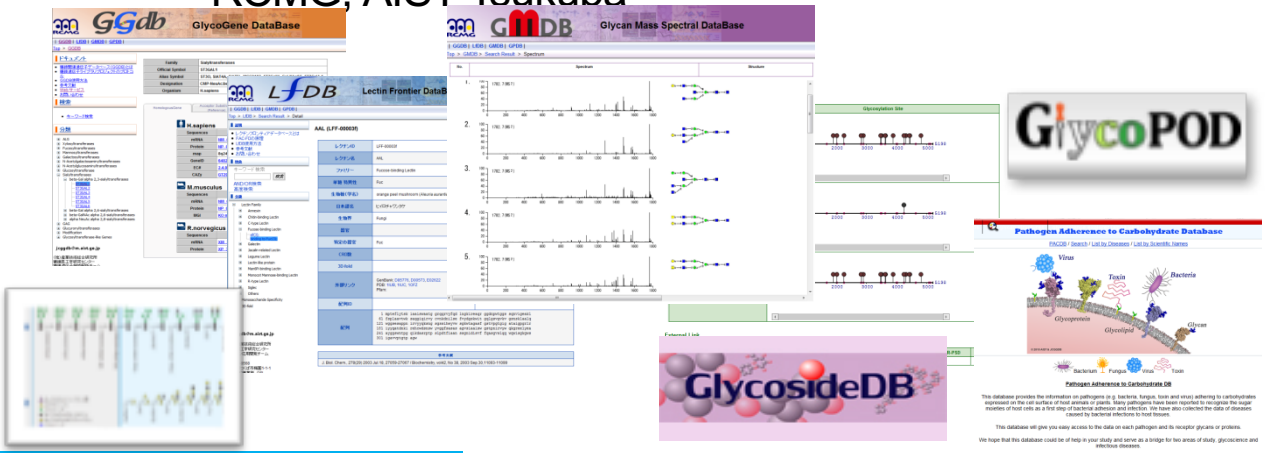
SOKA University

- Integration of glycoscience databases (PI: H. Narimatsu; April 2011-March 2014)
  - JCGGDB: Web services
  - GlycoRDF (http://glycoinfo.org/GlycoRDF)
- Development of an International Glycan Structure Repository and integration of glycoscience databases (PI: H. Narimatsu; Co-PI: K.F. Aoki-Kinoshita; April 2014-March 2017)
  - Based on Semantic Web technologies (RDF)
- Construction of a Glycoscience Portal (PI: K. F. Aoki-Kinoshita; April 2017-March 2022)
  - Integrating other –omics data with glycoscience through SemWeb technologies

JCGGDB Alliance databases (-2014)

SOKA University

RCMG, AIST Tsukuba



JCGGDB Report

Pathway database

Glycosidase database

GlycoGene Inhibitors database

RIKEN, Systems Glycobiology Research Group

N-glycan Conformation DB

Knockout Mouse DB

Nagoya Univ.

28FlyGlyco DB

Soka Univ.

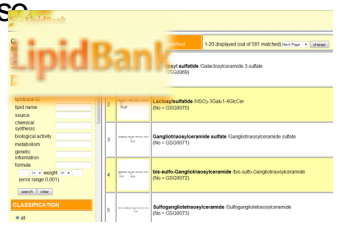
Patents database

GlycoGene DNA microarray database

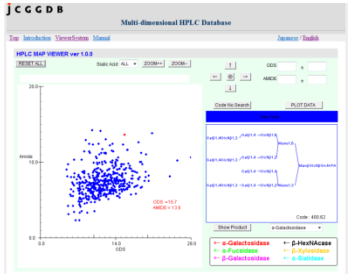
Noguchi Institute

Ritsumeikan Univ.

Executive Committee of Lipid Database



Kyoto Univ.



Nagoya City Univ.

Seikagaku Corp.

Worm database

Kyushu Univ.

2013

ACGG-DB (2014-present)

SOKA University

ACGG Asian Community of Glycoscience and Glycotechnology

Diseases

GDGDB is a database of glycan-related diseases and their responsible genes. We have compiled a database of diseases caused by mutations in glycan-related genes (e.g. glycosyltransferase, glycosidase, sugar-nucleotide synthase). Currently around 80 diseases with the responsible gene and pathosis are stored in this database.

PACDB provides the information on pathogens (e.g. bacteria, fungus, toxin and virus) adhering to carbohydrates expressed on the cell surface of host animals or plants. We have also collected the data of diseases caused by bacterial infections to host tissues. This database will give you easy access to the data on each pathogen and its receptor glycans or proteins.

Gene / Protein / Glycoprotein

GGdb

GlycoGene is a database which includes genes associated with glycan synthesis such as glycosyltransferase, sugar nucleotide synthases, sugar-nucleotide transporters, and sulfotransferases.

GlycoProtDB

GlycoProtDB is a glycoprotein database providing information of Asn (N)-glycosylated proteins and their glycosylated site(s), which were constructed by employing a bottom-up strategy using actual glycopeptide sequences identified by LC/MS-based glycoproteomic technologies. Glycoproteins identified from model organisms C.elegans , mouse (C57BL/6, male) and human.

LfDB

Lectin Frontier DataBase (LfDB) provides quantitative interaction data in terms of the affinity constants (Ka) of a series of lectins toward a panel of pyridylaminated (PA) glycans obtained by an automated frontal affinity chromatography with fluorescence detection (FAC-FD) system. Interaction data are shown in a bar graph format by either an actual measurement (V-V0) or an association constant (Ka). LfDB also provides a "1-parameter" function that helps to find a key structural element among the related glycan structures.

Ontologies / RDF

GDGDB provides information about the diseases of glycan synthesis and degradation that are caused by mutations in glycan-related genes. Genetic Glyco-Diseases Ontology (GGDonto): Ontology of genetic diseases known to be related to glycan metabolism.

Pathogen Adherence to Carbohydrate Ontology (PAConto): RDF representation of the data from the Pathogen Adherence to Carbohydrate Database (PACDB) and Ontology of infectious diseases known to be related to glycan binding.

GDGDB SemWeb version

SOKA University

ACGG-DB Genetic Glyco-Diseases Ontology (GGDonto) and Glyco-Disease Genes DataBase (GDGDB)

Genetic Diseases of Glycan Metabolism

120 List of Diseases

sorted by: labels; then by... • grouped as sorted

1. **Achondrogenesis type IB**
   [Aliases] ACG1B; Achondrogenesis, Fraccaro type
2. **ALG1-CDG**
   [Aliases] Congenital disorder of glycosylation, type Ik; CDG-Ik
3. **ALG11-CDG**
   [Aliases] Congenital Disorder of Glycosylation, Type Ip; CDG-Ip
4. **ALG12-CDG**
   [Aliases] Congenital disorder of glycosylation, type Ig; CDG-Ig
5. **ALG13-CDG**
   [Aliases] Congenital Disorder of Glycosylation, Type Is; CDG-Is
6. **ALG2-CDG**
   [Aliases] CDG-Ii; Congenital disorder of glycosylation, type Ii
7. **ALG3-CDG**
   [Aliases] Congenital disorder of glycosylation, type Id; CDG-Id
8. **ALG6-CDG**
   [Aliases] Congenital disorder of glycosylation, type Ic; CDG-Ic
9. **ALG8-CDG**
   [Aliases] Congenital disorder of glycosylation, type Ih; CDG-Ih
10. **ALG9-CDG**
    [Aliases] Congenital disorder of glycosylation, type IL; CDG-IL
11. **Alpha-mannosidosis**
    [Aliases] Alpha-mannosidase B deficiency
12. **Aspartylglucosaminuria**

Search

Databases

76 GDGDB
36 GGDB

Disease Types by Metabolic Pathways

74 Congenital Disorders of Glycosylation (CDGs)
46 Lysosomal Storage Diseases (LSDs)

Manifestation

1 Abdominal Pain
1 Adrenal Gland Diseases
9 Airway Obstruction

Ontology Tree

46 Genetic disorders of glycan degradation
46 Genetic disorders of glycan degradation (by systems, signs and symptoms)
74 Genetic disorders of glycan synthesis (by systems, signs and symptoms)
70 Genetic disorders of glycan synthesis (new nomenclature)

PacDB SemWeb version

SOKA University

ACGG-DB Pathogen Adherence to Carbohydrate Database (PACDB) and ontology

12 List of Microorganisms filtered from 446 originally (Reset All Filters)

sorted by: ScientificName and StrainName; then by... • grouped as sorted

Helicobacter pylori (12)

1. ***Helicobacter pylori***
   [Pathogen Adherence Molecules] bacteria adhesins:Adhesins, bacteria adhesins:BabA (blood group antigen binding adhesin), bacteria adhesins:HP0721, bacteria adhesins:HpaA, bacteria adhesins:Leb-binding adhesin, bacteria adhesins:Phe147 in HpaA (Helicobacter pylori adhesin A), bacteria adhesins:SabA (sialic acid-binding adhesin), bacterial glycan-binding proteins:Fibrillar hemmagglutinin, bacterial glycan-binding proteins:Neutrophil-activating proteins (150-kDa), and bacterial glycan-binding proteins:Unknown
2. ***Helicobacter pylori*** **CCUG 1139 (strain)**
   [Pathogen Adherence Molecules] bacterial glycan-binding proteins:35S-labeled or 125I-labeled proteins
3. ***Helicobacter pylori*** **CCUG 17874 (strain)**
   [Pathogen Adherence Molecules] bacteria adhesins:Adhesins, bacteria adhesins:Lectin-hemagglutinin adhesins, bacterial glycan-binding proteins:35S-labeled or 125I-labeled proteins, and bacterial glycan-binding proteins:Unknown
4. ***Helicobacter pylori*** **CCUG 17874 wt (strain)**
   [Pathogen Adherence Molecules] bacteria adhesins:SabA (sialic acid-binding adhesin)
5. ***Helicobacter pylori*** **CCUG 17875 (strain)**

Search

PACDB provides information about the diseases, in the pathogenesis of which the interaction of microbial glycan-binding proteins and glycans with host glycan ligands plays an important role.

(sys_mesh)	Classification of Diseases by Systems using MeSH (Medical Subject Headings) vocabulary
(sys)	Diseases recorded in PACDB classified by organ systems and organs
(pathog_mesh)	Classification of Diseases by Pathogens using MeSH
(pathog)	Diseases recorded in PACDB classified by infectious agents
(anim_mesh)	Classification of Animal Infectious Diseases using MeSH
(anim)	Animal Infectious Diseases recorded in PACDB

Diseases 1

- 100 (sys_mesh) Digestive System Diseases
- 96 (sys_mesh) Gastrointestinal Diseases
- 96 (sys_mesh) Gastroenteritis
- 12 **(sys_mesh) Peptic Ulcer**
- 2 (sys_mesh) Stomach Diseases

Species

12 Bacteria

GGdb GlycoGene Database

AIST ACGG-DB

The original data comes from this page. How is the Exhibit JSON data file.

Japanese

List of GlycoGene

Search

Family
- 18 (missing this field)
- 13 Fucosyltransferase
- 16 Galactosyltransferase
- 8 Glucosyltransferase
- 6 Glucuronyltransferase
- 5 Glucuronyltransferase N-

PathWay Class
- 13 glycolipid, ganglio series
- 10 glycolipid, globo series
- 27 glycolipid, lacto / neolacto series
- 19 Glycosaminoglycan, chondroitin sulfate
- 24 Glycosaminoglycan, heparan sulfate
- 16 Glycosaminoglycan, keratan sulfate

Keyword
- 80 (missing this field)
- 2 3'-phosphoadenosine 5'-phosphosulfate transporter
- 1 A blood type
- 1 a1,3-N-acetylgalactosaminyltransferase
- 2 alpha1,3-xylosyltransferase
- 1 asialo-GM2 (GA2)

Donor
- 45 (missing this field)
- 1 (not applicable)
- 1 ATP
- 2 CDP-Rbo
- 20 CMP-Neu5Ac
- 3 Dol-P-Glc

Expression
- 73 (missing this field)
- 1 aderenal gland
- 1 adipose
- 1 adrenal cortex
- 1 adrenal gland
- 31 adrenal glands

221 List of GlycoGenes

sorted by: GeneSymbol; then by... • ○ grouped as sorted

A(ABO)

Pathway Class	glycolipid, lacto/neolacto series
Keyword	A blood type, Histo-Blood group antigen, a1,3-N-acetylgalactosaminyltransferase
Designation	alpha 1,3-N-acetylgalactosaminytransferase
Donor	UDP-GalNAc
Modify Date	2017-03-27

A4GALT

Pathway Class	glycolipid, globo series
Keyword	CD77, Gb3, globotriaosylceramide, glycolipid, glycosphingolipid
Designation	alpha 1,4-galactosyltransferase
Donor	UDP-Gal
Modify Date	2017-03-27

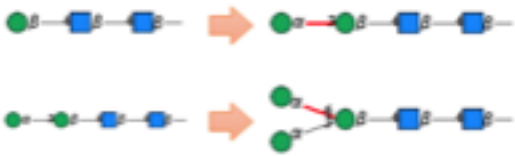
A4GNT

Pathway Class	O-glycan, mucin-type
Keyword	
Designation	alpha-1,4-N-acetylglucosaminyltransferase
Donor	UDP-GlcNAc
Modify Date	

ALG1

Pathway Class	N-glycan
Keyword	congenital disorders of glycosylation, lipid-linked oligosaccharide
Designation	GDP-Mannose:GlcNAc2-PP-dolichol beta-1,4-mannosyltransferase
Donor	GDP-Man
Modify Date	2016-08-10

ALG2



GlycoProtDB GlycoProtein DataBase AIST ACGG-DB

Top | Documents | Search | Contact Us

ADGRG1 Search

Search

Tissue/Cell line List

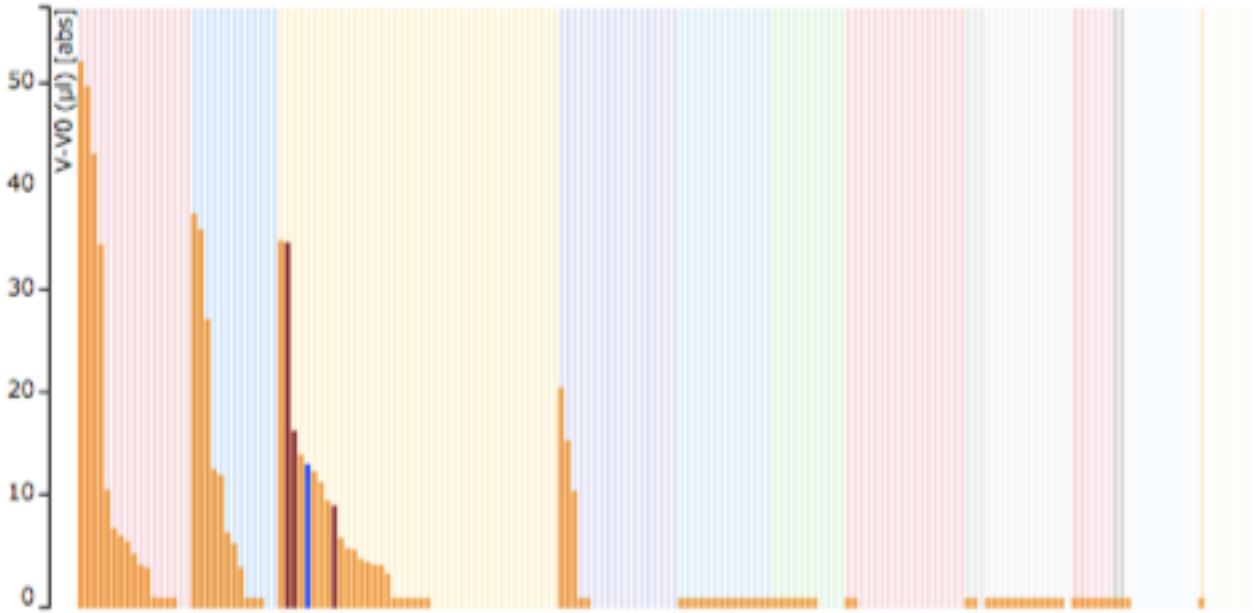
- Caenorhabditis elegans
  - whole body
- Homo sapiens
  - Ascites fluids
  - Ovary
  - Stomach
  - cell line
    - HepG2
    - HuH-7
  - serum
    - HCC patients
    - Healthy volunteers (cf HCC patients)
- Mus musculus
  - Brain
  - Colon
  - Heart
  - Kidney
    - WT
    - Fut9 KO
  - Liver
    - WT
    - b4Gal/T-I KO
  - Lung
  - Serum
  - Skeletal muscle
  - Stomach
  - Testis
- with Glycoform

Reset

Total:1227 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

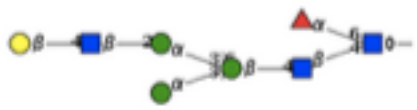
Assigned Protein						Sample(s) detected	
Gene Symbol	Protein Variant(s)					Organism	Tissue
	GlycoProtDB ID	UniProt	Protein Name	Length	Glycoform		
A1BG	GPDB0010987	M0R009	Alpha-1B-glycoprotein (Fragment)	305		Homo sapiens	cell line, serum
	GPDB0011359	P04217	Alpha-1B-glycoprotein	495		Homo sapiens	cell line, serum
	GPDB0011360	P04217-2	Isoform 2 of Alpha-1B-glycoprotein	373		Homo sapiens	cell line, serum
A2M	GPDB0010214	F5H1E8	Alpha-2-macroglobulin	90		Homo sapiens	cell line, serum
	GPDB0010355	F8W7L3	Alpha-2-macroglobulin (Fragment)	168		Homo sapiens	cell line, serum
	GPDB0011292	P01023	Alpha-2-macroglobulin	1474		Homo sapiens	cell line, serum
A2ML1	GPDB0008777	A8K2U0	Alpha-2-macroglobulin-like protein 1	1454		Homo sapiens	serum
AAT	GPDB0011981	Q2L9S7	Alpha-1-antitrypsin MBrescia variant (Fragment)	89		Homo sapiens	Ascites fluids, Ovary, cell line, serum
ACE2	GPDB0012545	Q9BYF1	Angiotensin-converting enzyme 2	805		Homo sapiens	cell line
	GPDB0012546	Q9BYF1-2	Isoform 2 of Angiotensin-converting enzyme 2	555		Homo sapiens	cell line
	GPDB0009237	B7Z6U3	Lysosomal acid phosphatase	162		Homo sapiens	Ovary, cell line
							Ovary

MornigaM (LfDB0154 - LPJ-0000bb)

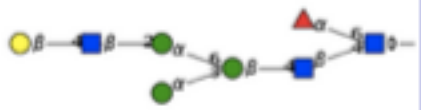
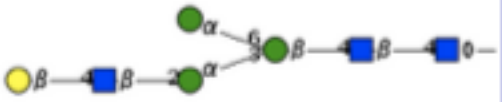


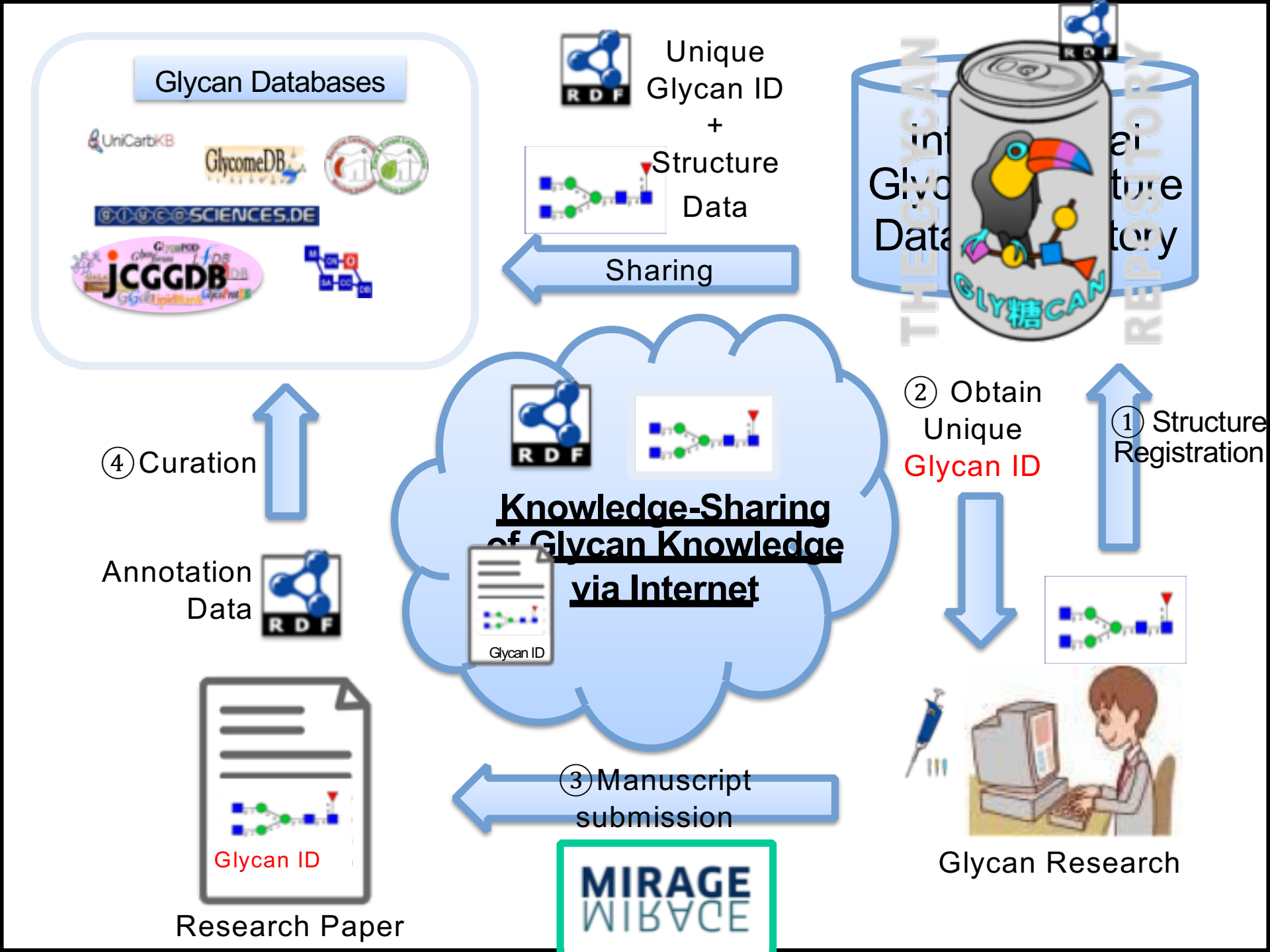
V-V0 | Ka | Sort values | Sort Family

Selected Glycan

Galactosylated JCGG-STR017541 GLYCAN ID : 402 V-V0 (μl) [abs]: 12.995	

Difference

Galactosylated JCGG-STR001131 GLYCAN ID : 401 V-V0 (μl) [abs] : 34.54	
Galactosylated JCGG-STR001142 GLYCAN ID : 302 V-V0 (μl) [abs] : 16.271	
Galactosylated JCGG-STR001368 GLYCAN ID : 404 V-V0 (μl) [abs] : 8.978	


Glycan Databases
UniCarbKB
GlycomeDB
SCIENCES.DE
JCGGDB
Unique
Glycan ID
+
Structure
Data
Sharing
GLYTOUCAN
① Structure
Registration
② Obtain
Unique
Glycan ID
Knowledge-Sharing
of Glycan Knowledge
via Internet
Glycan ID
④ Curation
Annotation
Data
Glycan ID
Research Paper
③ Manuscript
submission
MIRAGE
Glycan Research

Glycobiology, 2017, vol. 27, no. 10, 915–919
doi: 10.1093/glycob/cwx066
Glyco-Forum

Letter to Glycoforum

GlyTouCan: an accessible glycan structure repository

Michael Tiemeyer[2], Kazuhiro Aoki[2], James Paulson[3], Richard D Cummings[4], William S York[2], Niclas G Karlsson[5], Frederique Lisacek[6], Nicolle H Packer[7,8], Matthew P Campbell[7], Nobuyuki P Aoki[9], Akihiro Fujita[9], Masaaki Matsubara[2], Daisuke Shinmachi[9], Shinichiro Tsuchiya[8], Issaku Yamada[10], Michael Pierce[2], René Ranzinger[2], Hisashi Narimatsu[11], and Kiyoko F Aoki-Kinoshita[9,1]

1. Friedrich Altmann, University of Natural Resources and Life Sciences, Vienna, Austria
2. Antony Bacic, University of Melbourne, Australia
3. Christopher B. Barnett, University of Cape Town, South Africa
4. Julia Costa, Laboratory of Glycobiology, ITQB NOVA, Portal
5. Vivien J. Coulson-Thomas, University of Houston, USA
6. Tamara L. Doering, Washington University School of Medicine, USA
7. Nathan Edwards, Georgetown University, USA
8. Michiko Ehara, Asahi University, Japan
9. Tamao Endo, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan
10. Ten Feizi, Imperial College London, UK
11. Martin Frank, Biognos AB, Sweden
12. Morihisa Fujita, Jiangnan University, China
13. Koichi Fukase, Osaka University, Japan
14. Yuzuru Ikehara, AIST and Chiba University, Japan
15. Makoto Ito, Kyushu University, Japan
16. Yukishige Ito, RIKEN, Japan
17. Kenji Kadomatsu, Nagoya University Graduate School of Medicine, Japan
18. Osamu Kanie, Tokai University, Japan
19. Takane Katayama, Kyoto University, Japan
20. Toshisuke Kawasaki, Ritsumeikan University, Japan
21. Hiroto Kawashima, Chiba University, Japan
22. Carsten Kettner, Beilstein Institut, Germany
23. Kshitij Khatri, Boston University, USA
24. Yoshinobu Kimura, Okayama University, Japan
25. Hiroshi Kitagawa, Kobe Pharmaceutical University, Japan
26. Shinobu Kitazume, RIKEN, Japan
27. Yuriy A. Knirel, N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia
28. Kyoko Kojima-Aikawa, Ochanomizu University, Japan
29. Daniel Kolarich, Griffith University, Australia
30. Matthew R. Kudelka, Emory University, USA
31. Todd L. Lowary, Canadian Glycomics Network Scientific Director and University of Alberta, Canada
32. Thomas Luetteke, ITech Progress GmbH, Germany
33. Shino Manabe, RIKEN, Japan
34. David Matten, University of Cape Town, South Africa
35. Raja Mazumder, George Washington University, USA
36. Eiji Miyoshi, Osaka University, Japan
37. Antonio Molinaro, University of Napoli Federico II, Italy
38. Yasu S. Morita, University of Massachusetts Amherst, USA
39. Toni M. Mueller, University of Alabama at Birmingham, USA
40. Shunji Natsuka, Niigata University, Japan
41. Shoko Nishihara, Soka University, Japan
42. Sriram Neelamegham, State University of New York, USA
43. Tetsuya Okajima, Nagoya University School of Medicine, Japan
44. Shujiro Okuda, Niigata University, Japan
45. Noorjahan Panjwani, Tufts University School of Medicine, USA
46. Dayoung Park, University of California, Davis, USA
47. Serge Perez, France
48. Salomé S. Pinho, University of Porto and Institute for Research and Innovation in Health, Portugal
49. Melody Porterfield, University of Georgia, USA
50. Alka Rao, CSIR-Institute of Microbial Technology, Chandigarh, India
51. Celso A. Reis, University of Porto, Portugal
52. Rafael Ricci de Azevedo, University of Sao Paulo, Brazil
53. Nancy Schwartz, University of Chicago, USA
54. Siro Simizu, Keio University, Japan
55. Avadhesha Surolia, Indian Institute of Science, Bangalore, India
56. Naoyuki Taniguchi, RIKEN, Japan
57. Carlo Unverzagt, University of Bayreuth, Germany
58. Masahiro Wakao, Kagoshima University, Japan
59. Christopher M. West, University of Georgia, USA
60. Robert J. Woods, University of Georgia, USA
61. Ajit Varki, University of California, San Diego, USA
62. Yoshiki Yamaguchi, RIKEN, Japan
63. Kazuo Yamamoto, The University of Tokyo, Japan
64. Heng Yin, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, China
65. Joseph Zaia, Boston University, USA

Currently over 100,000 entries registered

- Data from GlyTouCan Partners as well as CHEBI, BCSDB and GlycoEpitope

BCSDB (5428)
CFG (6374)
Carbbank(CCSD) (14840)
GLYCOSCIENCES.de (15905)
GlycoChemExplorer (8021)
GlycoEpitope (169)
GlycoNAVI (91)
Glycobase (199)
GlycomeDB (39053)
JCGGDB (22039)
JCGGDB AIST (7867)
JMSDB (911)
KEGG (10134)
PDB (894)
PDBe CC (264)
PDBj CC (264)
PubChem CID (21356)
PubChem SID (21356)
RCSB PDB CC (264)
SugarBindDB (173)
UniCarb-DB (865)
UniCarbKB (2936)

GlyTouCan Data Overview

SOKA University

UniCarb-DB
The Glycomic MS Database and Repository

Plant & Fungal Carbohydrate Structure Database

Bacterial Carbohydrate Structure Database

UniCarbKB

THE GLYCAN REPOSITORY

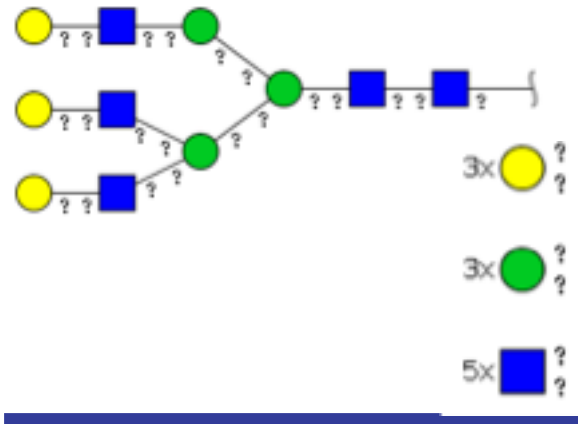
GLY糖CAN

PDBj
Protein Data Bank Japan

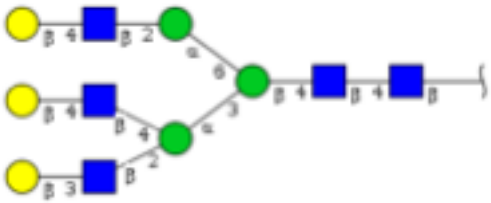
jCGGDB

GlycomeDB

PubChem

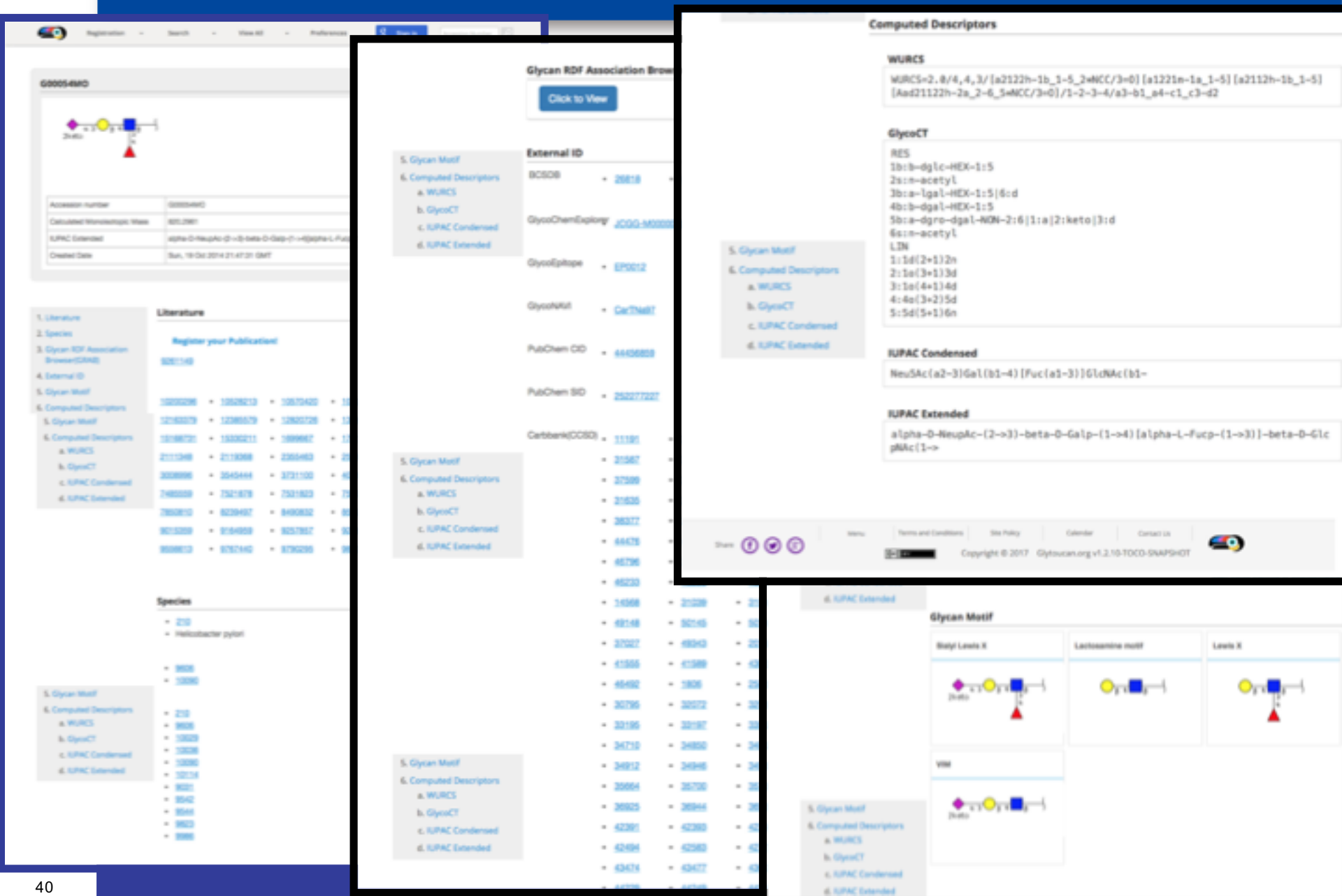


GRAB Tool



GlyTouCan Entry Attributes

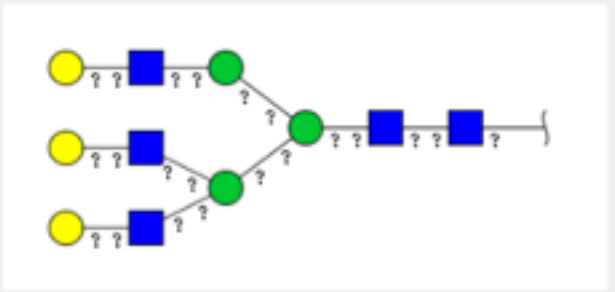
SOKA University



Glycan RDF Association Browser

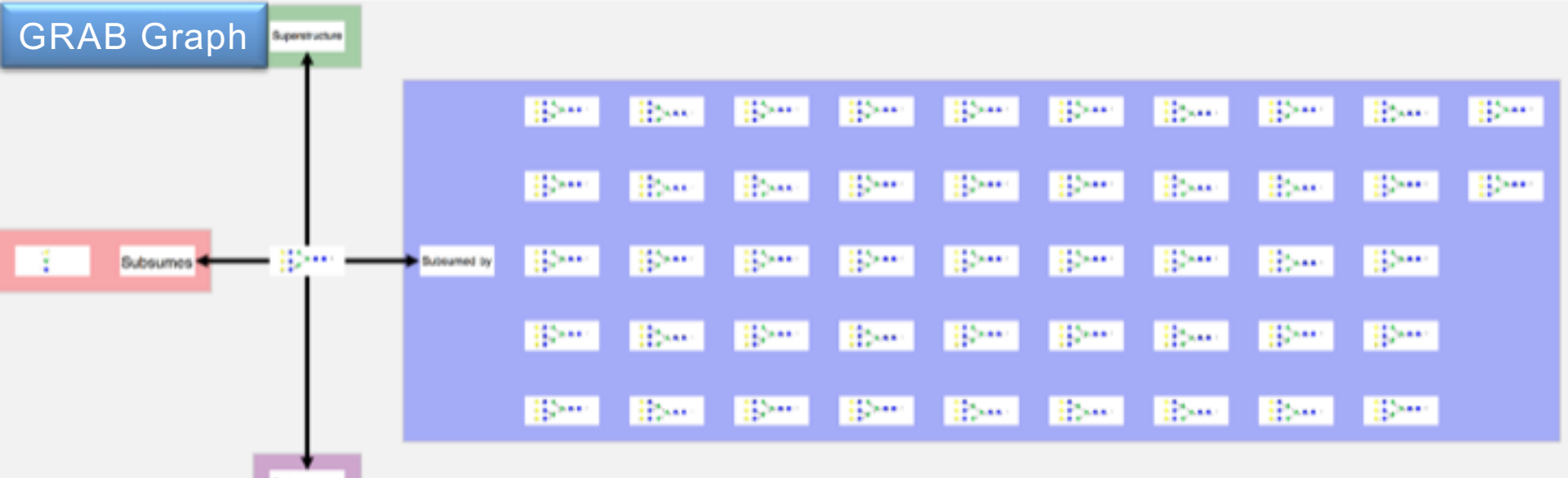
SOKA University

GRAB Tree



G91496ML

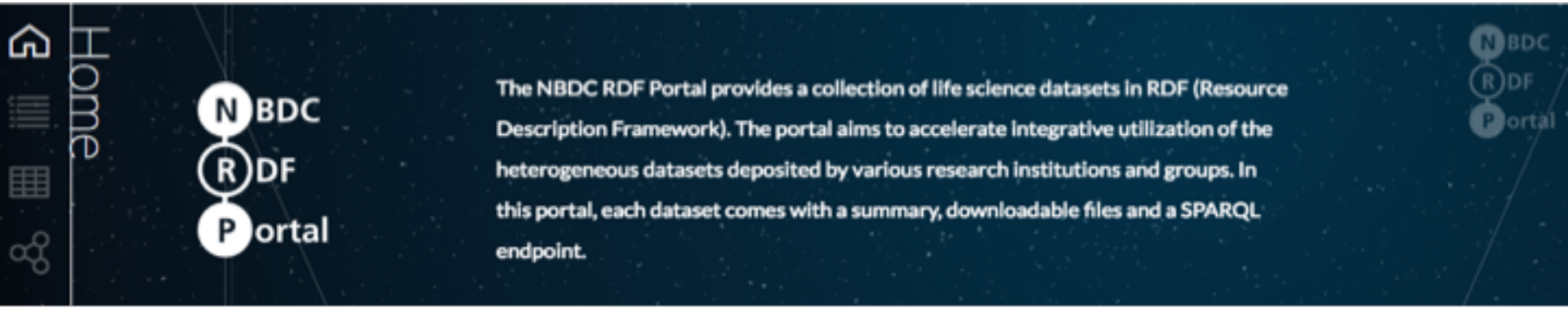
GRAB Graph



National Bioscience Database Center

SOKA University

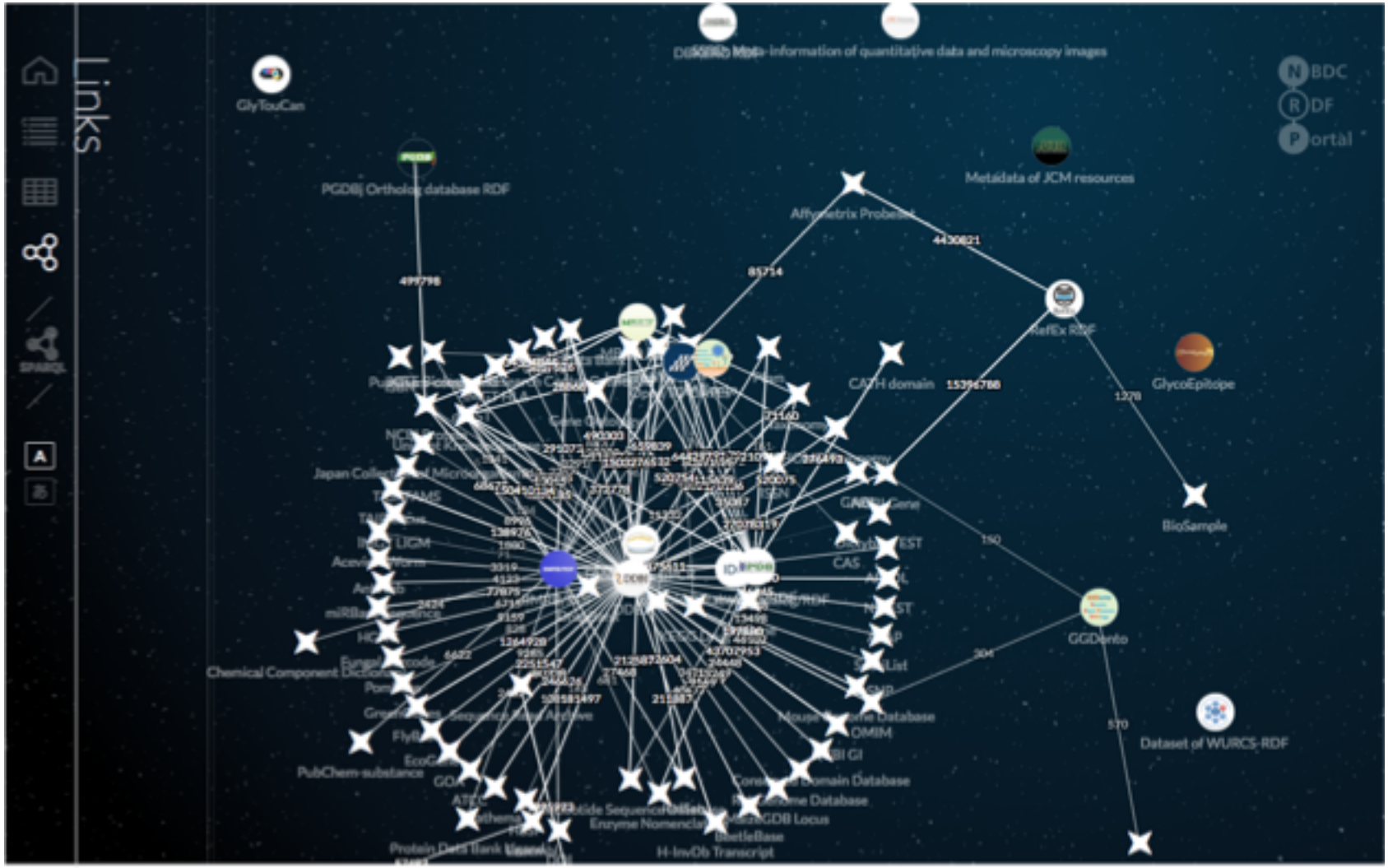
- Over ten years of life science database integration
- BioHackathons held yearly
  - http://biohackathon.org/10years/
  - Keywords: web services, workflows, RDF, ontology

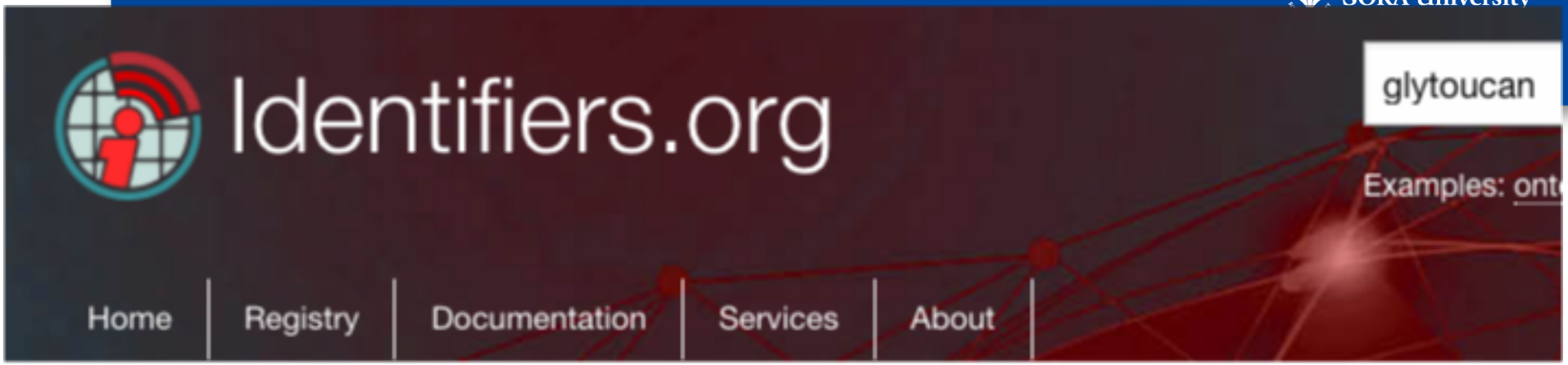


- https://biosciencedbc.jp/en/

The Linked Data in NBDC

SOKA University





2 results found.

GlyTouCan **[Namespace: glytoucan]**

GlyTouCan is the single worldwide registry of glycan (carbohydrate sugar chain) data.

The Glycan Repository

GlyTouCan is the single worldwide registry of glycan (carbohydrate sugar chain) data.

Home: https://glytoucan.org
Institution: Soka University, Hachioji, Tokyo
Location: Japan
Example: http://identifiers.org/glytoucan:G00054MO
Uptime: 100%

Overview of lecture

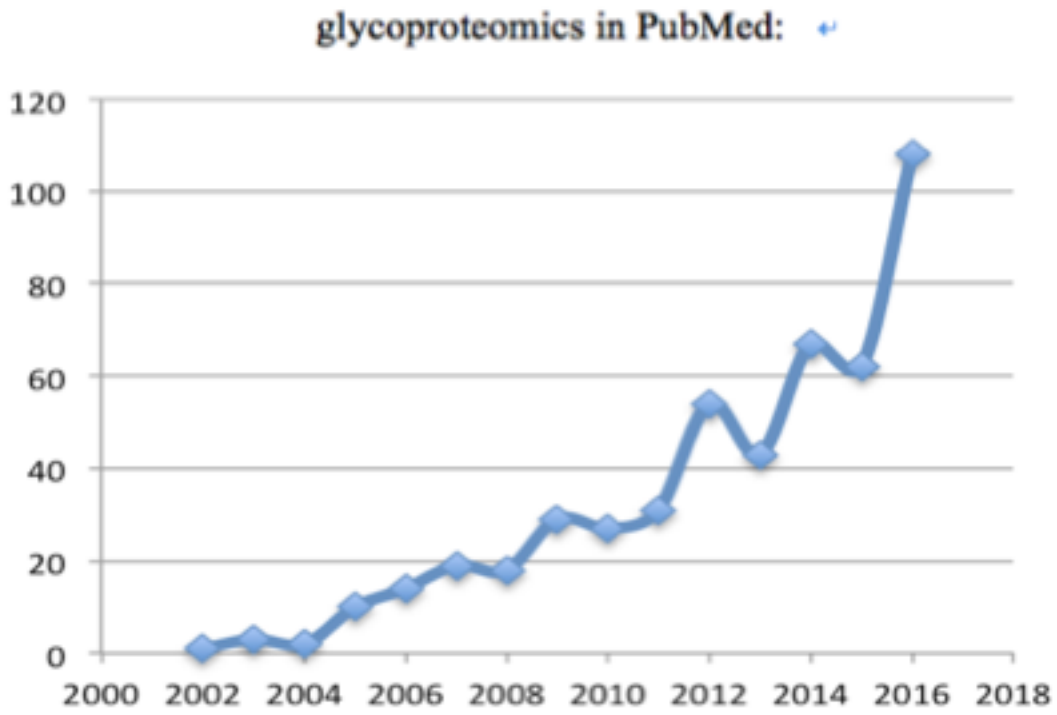
SOKA University

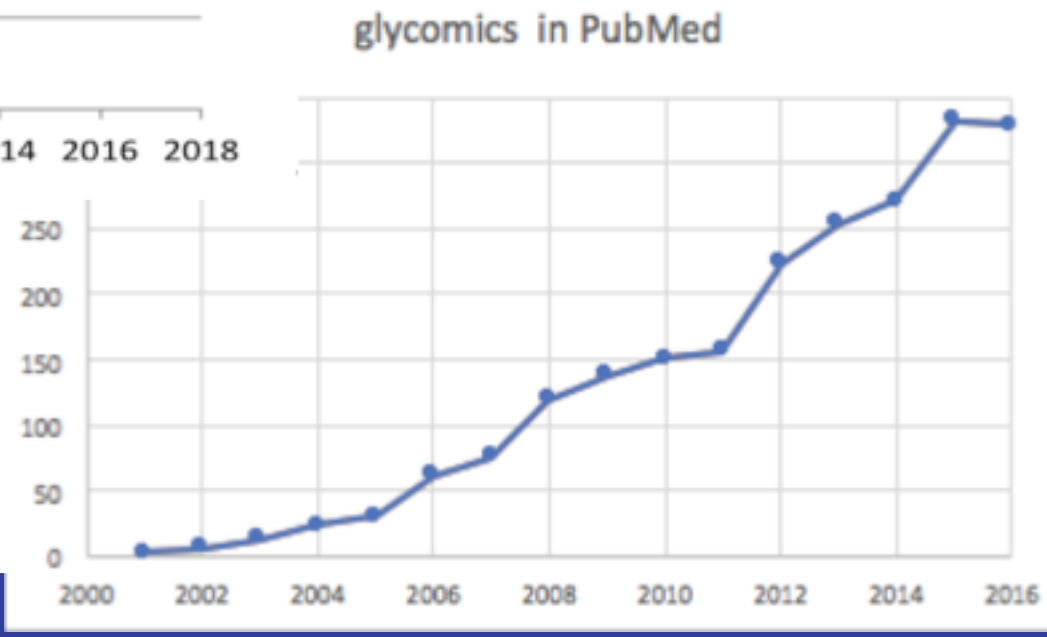
- The Semantic Web and Databases
- WURCS representation of carbohydrate structures
- Glyco-related database examples
- Applicability of the Semantic Web

GlyCosmos Portal
(April, 2017-March 2022)

Trends in glycoscience (PubMed)

SOKA University





jPOSTrepo

SOKA University

JPOST REPOSITORY

Repository | Submit | Help | Sign in | Sign up

About jPOSTrepo

jPOSTrepo (Japan ProteOme STandard Repository) is a new data repository of sharing MS raw/processed data. It consists of a newly-developed, high-speed file upload process, flexible file management system and easy-to-use interfaces. Users can release their "raw/processed" data via this site with a unique identifier number for the paper publication. Users also can suspend (or "embargo") their data until their paper is published. The file transfer from users' computer to our repository server is very fast (roughly ten times faster than usual file transfer) and uses only web browsers – it does not require installing any additional software.

Reference

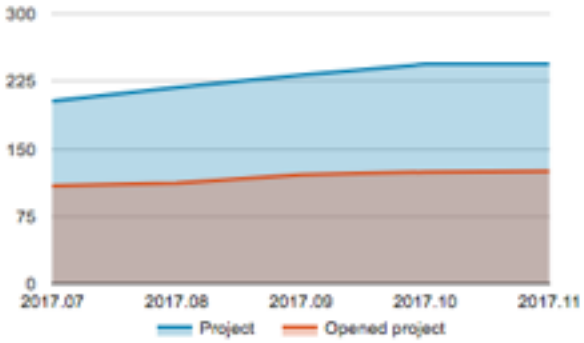
Okuda, S. et al. jPOSTrepo: an international standard data repository for proteomes. Nucl. Acids Res. 45 (D1): D1107-D1111 (2017). doi: 10.1093/nar/gkw1080
[pubmed]

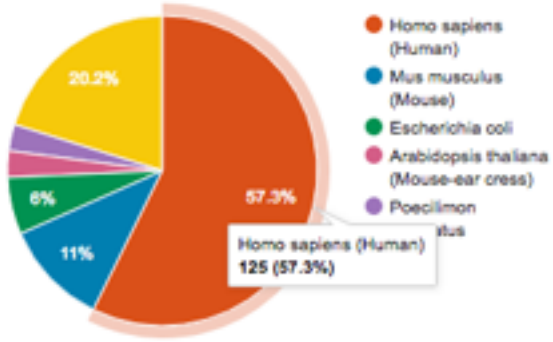
Statistics

244 projects are registered. **125** are opened.

27893 files amount to **6.4 TB**.

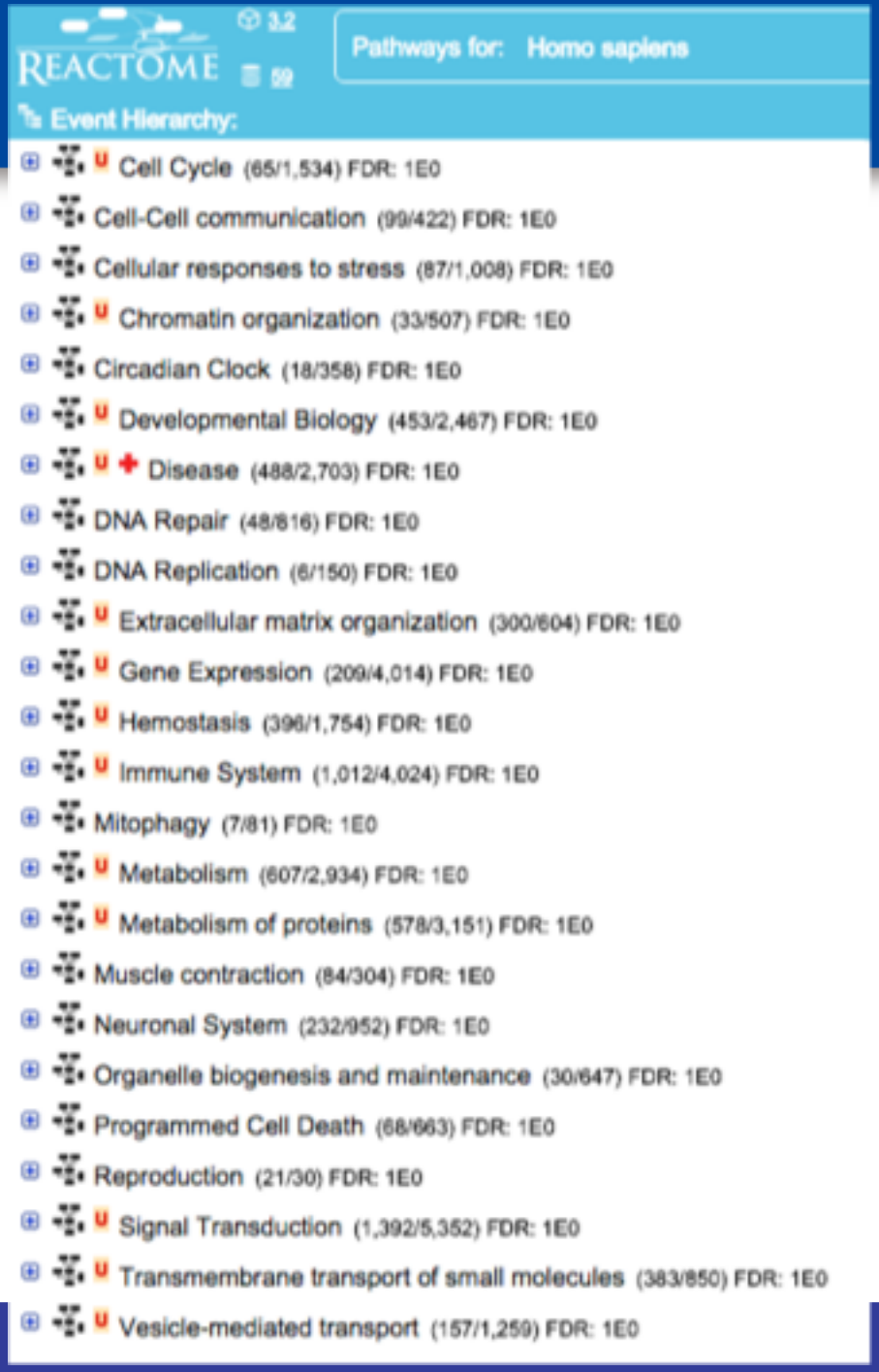
26 species.





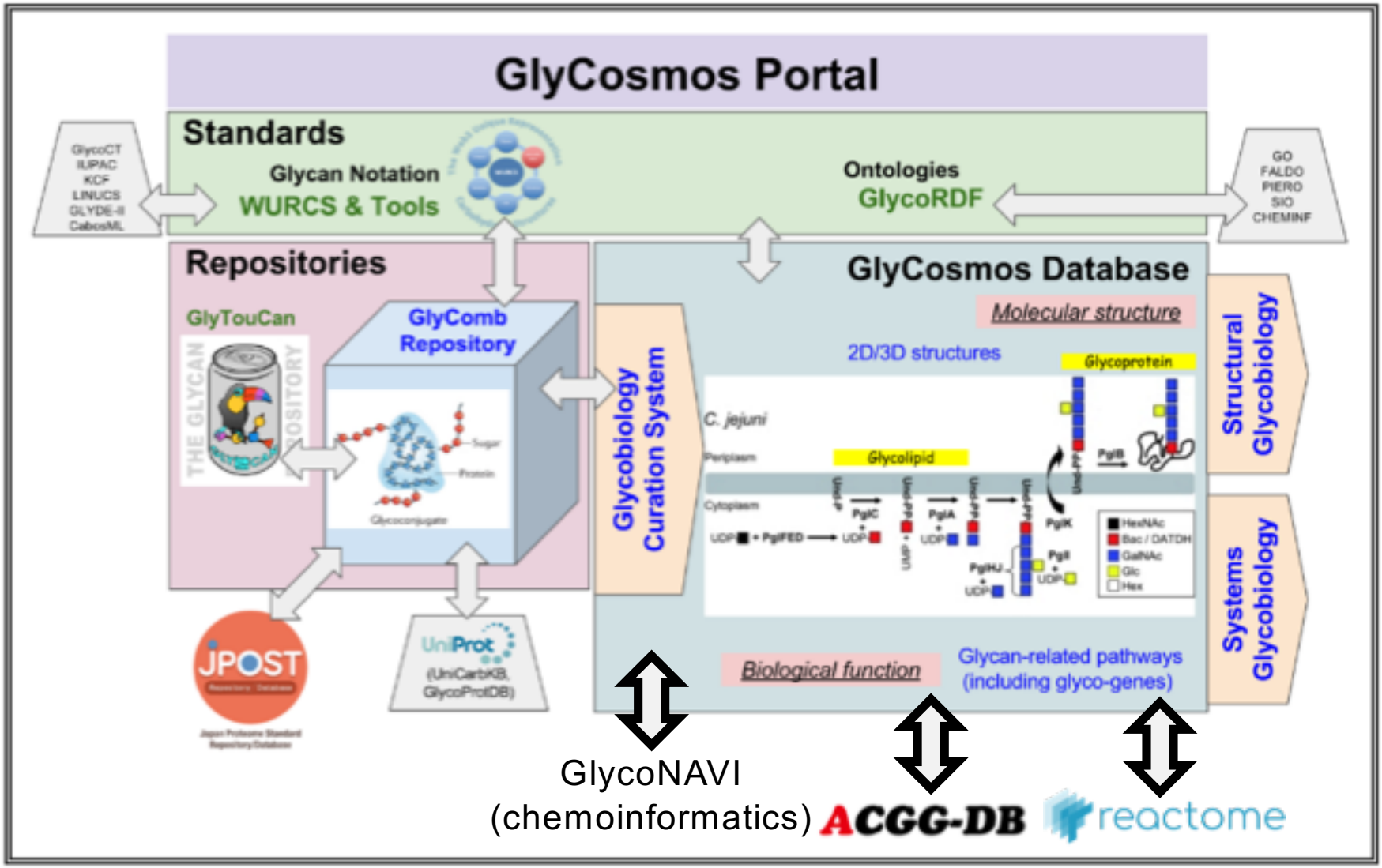
Glycan-related pathways

- Taking the 30,000 or so glycoproteins in UniProt (as of March, 2017), we ran them through Reactome's Analysis Tools and obtained 1660 hits in human.
- These human glycoproteins are involved in 6773 pathways, including signaling, immune system, extracellular matrix and development, in addition to glycan biosynthesis and metabolism.



GlyCosmos Overview

SOKA University



GlyCosmos Goals

SOKA University

- Taking advantage of the strengths of the Japanese carbohydrate research community, to collaborate with the glycoscience community to consolidate glycan-related knowledge into GlyCosmos.

- Moreover, by developing the GlyCosmos Database, it will be possible to integrate other omics research communities including proteomics, metabolomics, lipidomics, etc.

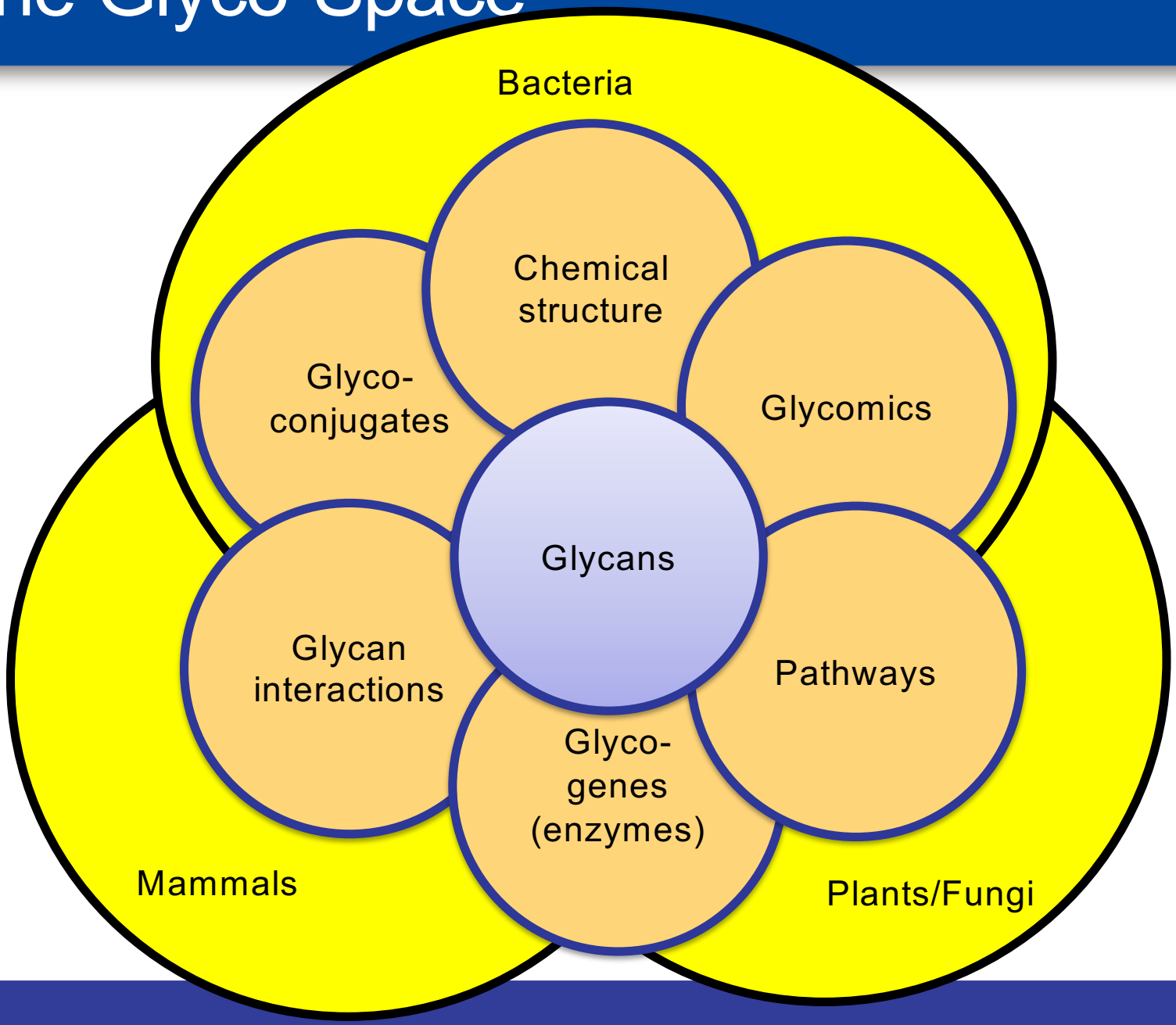
- In the end, any student or researcher with interest in glycoscience can access the portal and easily gain an understanding of glycans.
  - Glycogenes → Related pathways, diseases, glycan structures, 3D structures, glycoconjugates → Glycan function

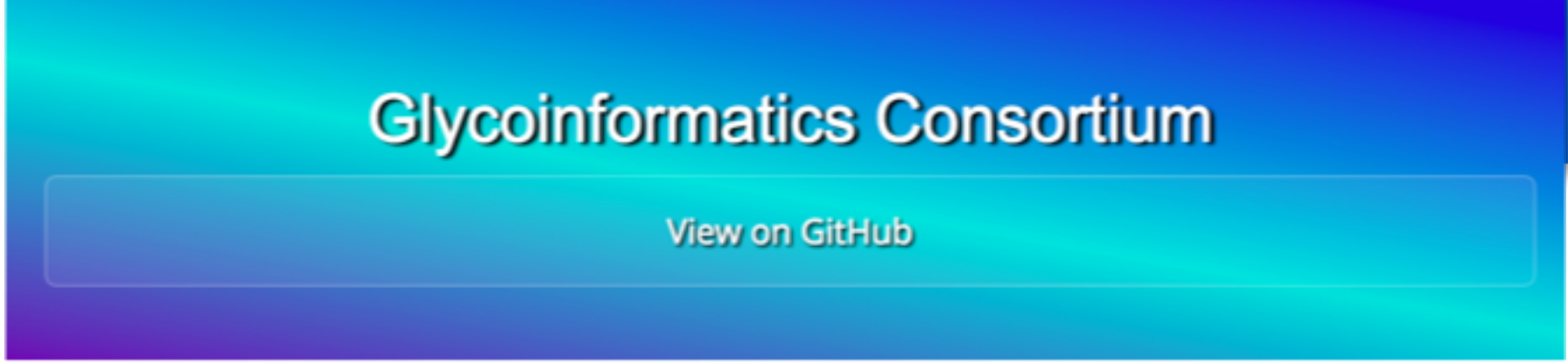
“True” integration

SOKA University

- GlyCosmos will not accumulate data from other resources and store them centrally
  - All other resources will stay in place (as long as they are publicly available and stable)
- GlyCosmos will provide **semantic integration** of other omics data by establishing:
  - Ontologies that define the biological/chemical relationships between glycans and other data
  - User interfaces that make these relationships clear to the user and easily searchable
- GlyCosmos is also cooperating with other major integration projects including GlyGen (US) and GlyConnect (SIB) to share work and data

The Glyco-Space





- Purpose:
  - To provide and maintain a centralized source software repository for the glycosciences
    - http://glic.glycoinfo.org
  - To provide a communication point for glycoscientists to contact glycoinformaticians
    - List of members and their software
    - “Wish list” for anyone to ask for specific tools they need developed or requests to improve/expand existing tools

Glycoinformatics Consortium GLIC

SOKA University

- Members:
  - Sanjay Agravat, Emory University
  - Davide Alocci, Swiss Institute of Bioinformatics
  - Nobuyuki Aoki, SPARQLite, LLC
  - Kiyoko F. Aoki-Kinoshita, Soka University
  - Matthew Campbell, Macquarie University
  - Joshua Klein, Boston University
  - Frédérique Lisacek, Swiss Institute of Bioinformatics
  - Thomas Luetteke, Justus-Liebig University Giessen
  - Julien Mariethoz, Swiss Institute of Bioinformatics
  - Masaaki Matsubara, Noguchi Institute
  - Rene Ranzinger, University of Georgia
  - Daisuke Shinmachi, SPARQLite, LLC
  - **Issaku Yamada, Noguchi Institute**
  - Joseph Zaia, Boston University

GLIC: Recent activities

SOKA University

- Development of glycoconjugate ontology: GlycoCoO
- GlyCosmos, UniCarbKB, GlyConnect and GlyGen are involved
- Currently preparing manuscript describing GlycoCoO

- Please contact me if you'd like to join us!
- kkiyoko@soka.ac.jp

Acknowledgements

Soka University

Masaaki Shiota

Tamiko Ono

Shinichiro Tsuchiya

Haruko Kitakaze

AIST & GL-i, Japan

Noriaki Fujita

Yoshinori Suzuki

Kiyohiko Angata

Hisashi Narimatsu

Niigata University

Shujiro Okuda

Yu Yamaguchi

The Noguchi Institute

Issaku Yamada

Nobuaki Miura

SPARQLite, LLC

Nobuyuki P. Aoki

Daisuke Shinmachi

Funding agencies:

Integrated Database Project (Japan Science and Technology Agency (JST) and Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT))

Discover your potential

SOKA University