



Glyco-Databases in Japan

Kiyoko F. Aoki-Kinoshita

Bioinformatics Department

Faculty of Engineering

Soka University



Carbohydrate-related databases in Japan

- Glycan structure databases
 - KEGG GLYCAN (Kyoto University)
 - Glycoconjugate Data Bank (Hokkaido University)
- Glyco-epitope database (Ritsumeikan University)
- GGDB Glyco-gene database (AIST)



KEGG GLYCAN

Glycome informatics resource integrating genomics and chemistry

KEGG2

DISEASE

DRUG

GLYCAN

COMPOUND

REACTION

Organisms

Glycosyltransferases

Functional information of genes and proteins is organized in KEGG as ortholog groups, called KEGG Orthology (KO) groups, to cover all organisms. The KO groups for glycosyltransferases are finely classified ortholog groups distinguishing known substrate specificity, which can be viewed as functional hierarchies of the [KEGG BRITE](#) database.

- [Glycosyltransferases](#)
- [Glycosyltransferase reactions](#)

Metabolic Pathway Maps

The [KEGG PATHWAY](#) database is a collection of manually drawn KEGG pathway maps representing current knowledge on molecular interaction networks, including glycan biosynthesis and metabolism.

- [Glycan biosynthesis and metabolism](#)
- [Glycosyltransferases along the pathways](#)

Glycan Structure Maps

Gene Microarray to Glycan Structures

GECS (Gene Expression to Chemical Structure) is a new software package for linking genomic or transcriptomic contents of genes to chemical structures of biosynthetic substances, especially glycans.

- [N-glycan prediction server](#) *New!*

Glycan Binding Proteins

Different types of glycan binding proteins are organized as KO groups and classified as a functional hierarchy of [KEGG BRITE](#).

- [Glycan binding proteins](#)

Regulatory Pathway Maps

The [KEGG PATHWAY](#) database contains various pathway maps involving glycan-binding proteins, such as:

- [Signaling molecules and interaction](#)
- [Cell communication](#)

■ <http://www.genome.jp/kegg/glycan/>



Gly

GlycoEpitope

[Home](#) | [List all epitopes](#) | [List all antibodies](#)

About this database

Carbohydrate chains occupy truly significant positions in various fields of life sciences and biotechnology. Recently, the wide-ranging involvement of carbohydrate chains in life sciences has been extended to such diverse functions as cell to cell recognition and communication in neuronal tissues and immune systems, pathogen recognition, sperm-egg recognition and fertilization, regulating hormonal half-lives in the blood, directing embryonic development and differentiation, and directing distribution of various cells and proteins throughout the body. A large number of polyclonal or monoclonal antibodies have been used as very important tools for analyzing expression of various carbohydrate chains and their functions. In this database, useful information on these carbohydrate antigens, i.e. glyco-epitopes, and antibodies has been assembled as a compact encyclopedia.

[Epitope Search](#)

[Antibody Search](#)

- ◆ If you have any questions or comments, please contact us by [Feedback Form](#).
- ◆ GlycoEpitope Database is provided by Kawasaki laboratory. The database server is located in Research Center for Glycobiotechnology, Ritsumeikan University.

■ <http://www.glyco.is.ritsumeikan.ac.jp/epitope/>

Document

[About GlycoGene DataBase](#)
[Protocols of GlycoGene Library Project](#)
[How to use GGDB](#)
[References](#)
[Contact Us](#)

Text Search

GG Family

- [-] ALG
 - [-] Doner Synthesis
 - [-] Dol-P-Man Synthesis
 - [DPM1](#)
 - [DPM2](#)
 - [DPM3](#)
 - [-] Dol-P-Glc Synthesis
 - [ALG5](#)
 - [-] Cytoplasmic ALG(GnT)
 - [ALG14](#)
 - [DPAGT1](#)
 - [GLT28D1](#)
 - [-] Cytoplasmic ALG(ManT)
 - [ALG1](#)
 - [ALG11](#)
 - [ALG2](#)
 - [-] ER Lumen ALG(ManT)
 - [ALG12](#)
 - [ALG3](#)
 - [ALG9](#)
 - [-] ER Lumen ALG(GlcT)
 - [ALG10](#)
 - [ALG6](#)
 - [ALG8](#)
 - [-] Glucosyltransferases
 - [UGCG](#)
 - [UGCG1](#)

About GlycoGene DataBase

Significance of GGDB

GlycoGene includes genes associated with glycan synthesis such as glycosyltransferase, sugar nucleotide synthases, sugar-nucleotide transporters, sulfotransferases, etc. At present, over 180 human glycoGenes were identified, cloned and characterized. In "Construction of GlycoGene Library Project" (April, 2001 - March, 2004), we collected and compiled the data on such glycoGenes as GlycoGene Database (GGDB), which is the first database to store information on substrate specificity. GGDB provides necessary information for the analysis of glycoGenes.

Present Status of GGDB

The purpose of GlycoGene Database (GGDB, <http://jcgddb.jp/ggdb>) is to provide users with easy access to the information on glycoGenes via website. In GGDB, the following property information of each glycoGene are stored in XML format: gene names(gene symbols), enzyme names, DNA sequences, tissue distribution(gene expression), substrate specificities, homologous genes, EC numbers, and external links to various databases. It graphically shows the information such as substrate specificities, etc.



Family	Fucosyltransferases
Official Symbol	FUT9
Other Alias	Fuc-TIX
Recognition	alpha(1,3)-fucosyltransferase 9
Organism	Homo sapiens

Sequence	Accession
mRNA	NM_010681 (GDS)
Protein	NP_038572 (GDS)
map	6q16



Japan-based Glycome analysis tools

■ KEGG GLYCAN

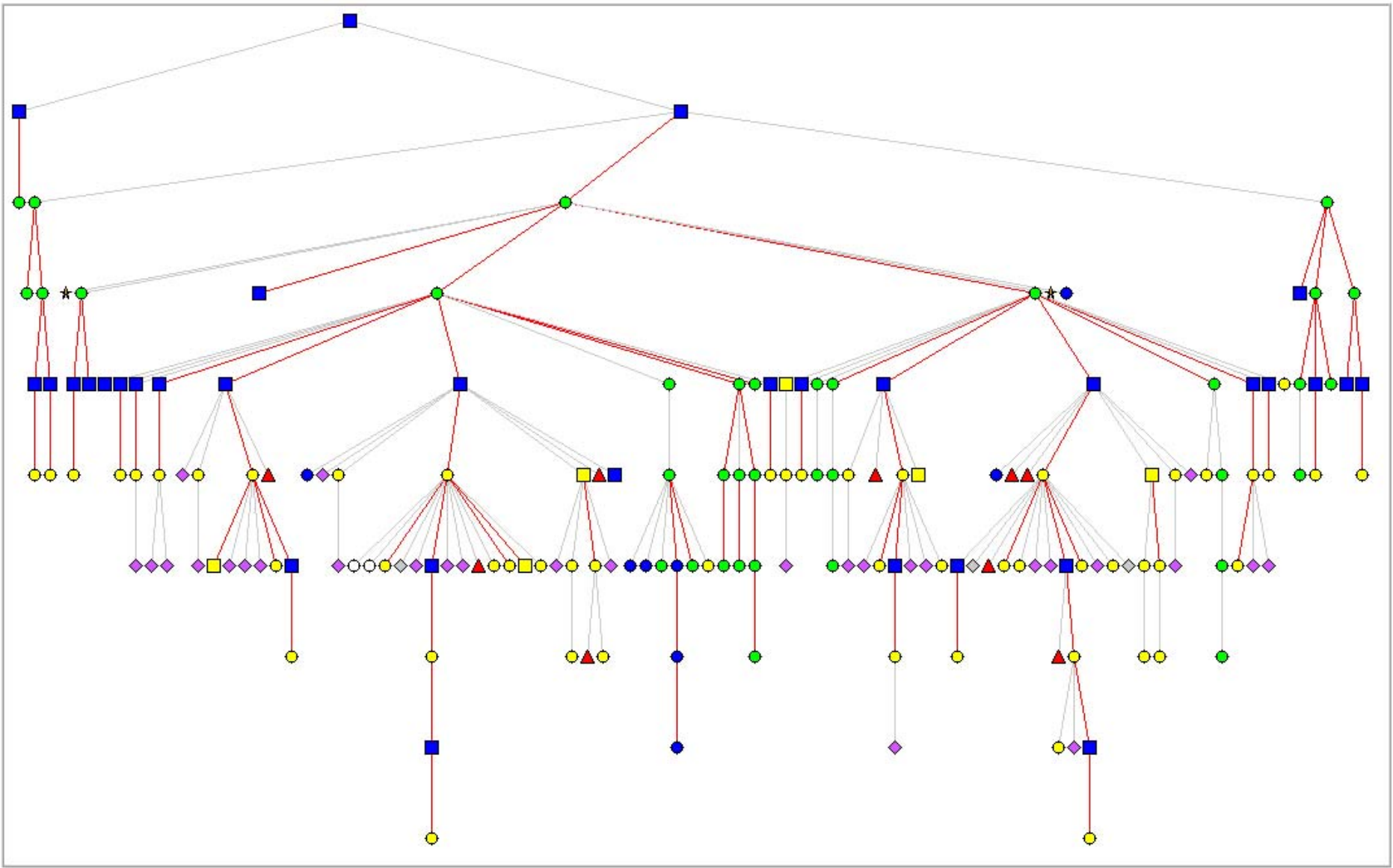
- Prediction of glycan structures from glyco-gene expression data
- Composite Structure Map (CSM)

■ RINGS

- Utilities for glycan data format conversion
- Glycan structure mining tools
- <http://rings.t.soka.ac.jp/>

Next2: -
 Next1: -
 Root:

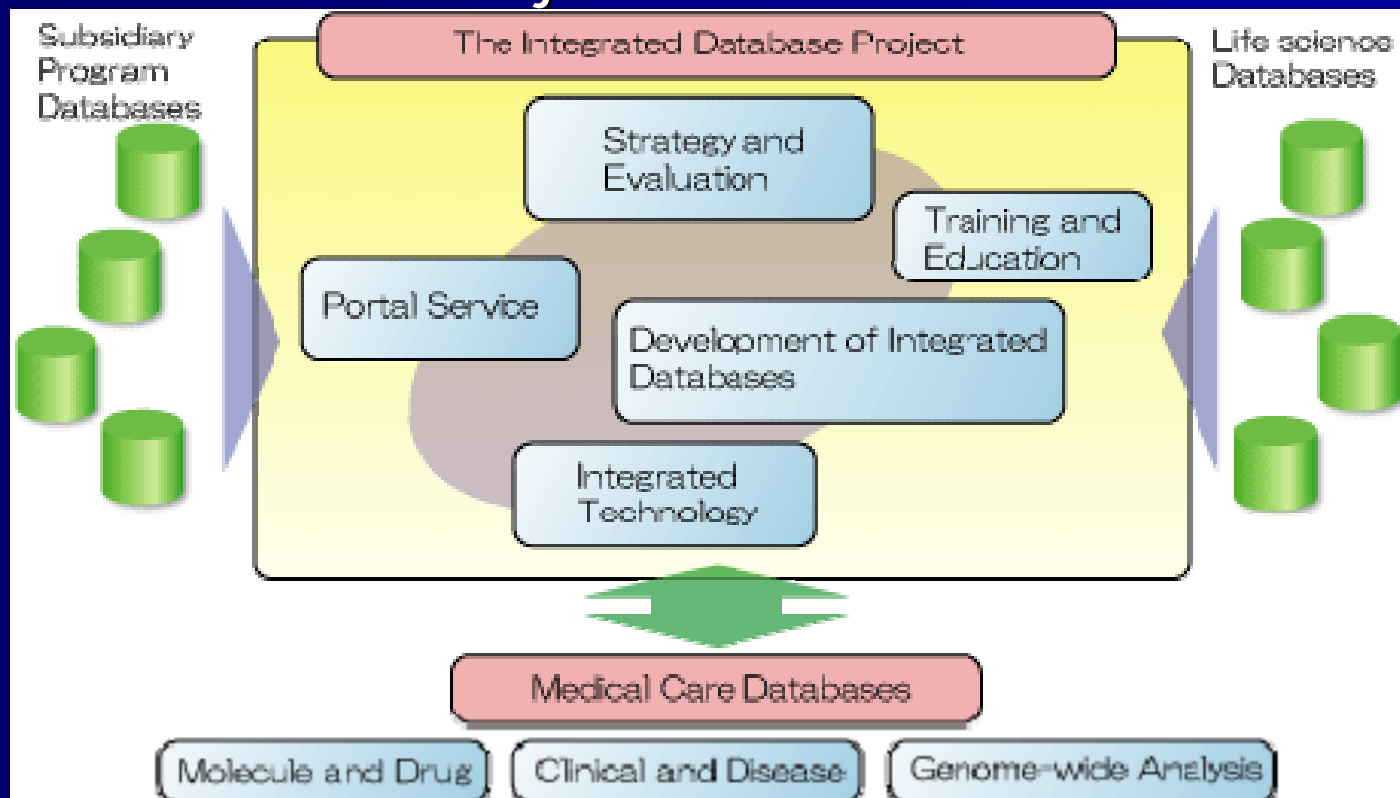
Threshold: Position: Node size:





Integrated Database Project

- Since 2006, Ministry of Education, Culture, Sports, Science, and Technology (MEXT) started the Integrated Database Project for life sciences





Up to last year

- Survey of databases, evaluation, and drafting of a strategy
- Development and maintenance of a portal site
- Research and development on integrated technology



Towards 2010

- Promotion by establishing a central organization (DBCCLS: Database Center for Life Science)
- Development and maintenance of integrated databases
- Combination of published literature and data annotations
- Development of new databases and research into their utilization
- Acceptance of useful existing databases that have been poorly maintained
- Training and education for database development



What's needed?

- A “GenBank” (GlycoBank?) for a central world-wide resource for glyco-data
- The Integrated Database Project is mainly focused on Japanese databases, but the glyco-community may be more involved
 - KEGG
 - Glyco-epitope
 - Glycoconjugate Data Bank



A possible alternative

- Web services and workflows:
 - Direct method for “integrating” databases that are customizable for individual researchers
 - Both KEGG and RINGS provide web services for their tools, although they are as of yet still limited in scope
 - Development of large libraries of web services may be a start