

「統合化推進プログラム」(統合データ解析トライアル)

開発課題名:

タンパク質-糖鎖間の糖鎖結合部位の
解明のためのツール改良及び解析

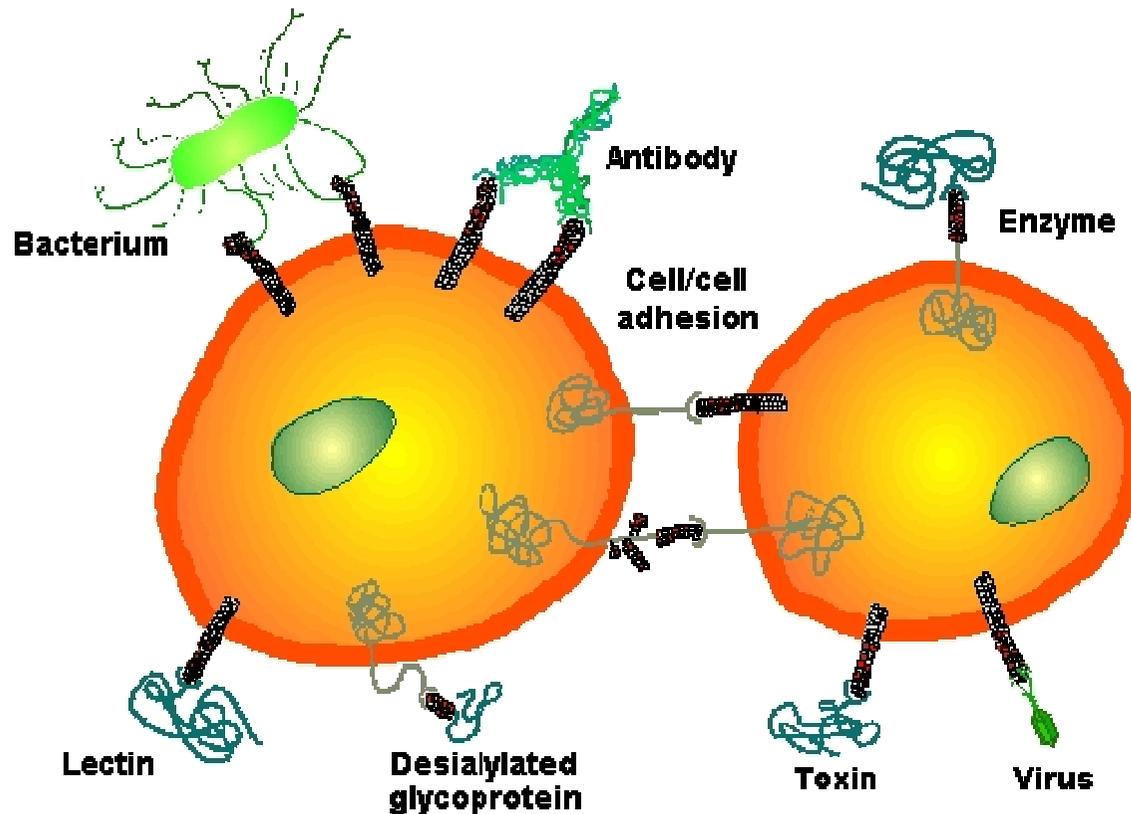
平成**26**年**3**月**2**日(研究終了報告)

研究代表者: 細田 正恵

創価大学大学院工学研究科生命情報工学専攻

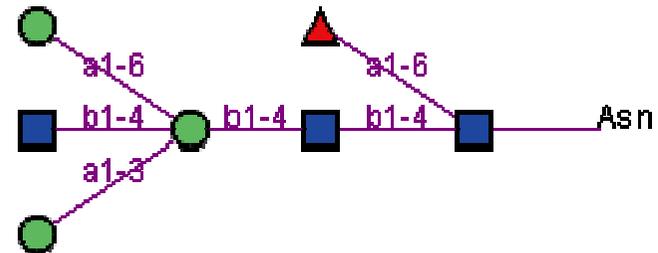
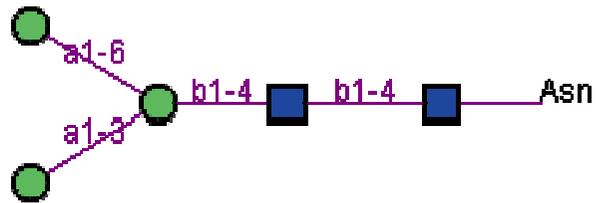
生体内での糖鎖機能

- 糖鎖は細胞表面のタンパク質や脂質に結合し細胞間接着、血液型の決定、ウイルス感染や抗原抗体反応などの生体反応で重要な役割を果たしている
- 生体分子が糖鎖を認識し結合することによりシグナル伝達や生体内機能調節を行っている

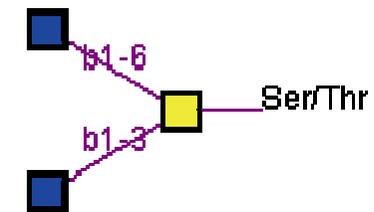
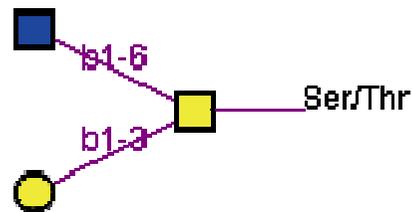
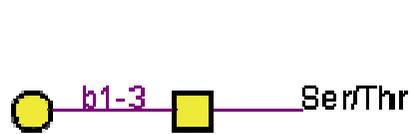


糖鎖構造 (コア構造)

○ N型糖鎖



○ O型糖鎖



○ グリコサミノグリカン



▲	フコース
●	ガラクトース
■	Nアセチルガラクトサミン
◇	グルクロン酸
■	Nアセチルグルコサミン
●	マンノース

糖鎖インフォマティクスについて

- 生体内の機能的役割の理解に糖鎖生物学の研究が重要視されプロテオグリカンやグリコシル化の研究がすすめられている
- 糖鎖の研究として多くの実験結果のデータや糖鎖構造の情報がデータベースに蓄積され公開されている
(JCGGDBやKEGG GLYCAN、Consortium for Functional Glycomics、Glycosciences.deなど)
- 多様な修飾や分岐構造を持つ糖鎖構造の複雑性により、プロテオミクスに用いられる構造解析ツールが糖鎖にすぐ対応できず、データ解析を行えるツールが多く存在しない
- データを利用した糖鎖の機能解析には糖鎖インフォマティクスの技術が必要になっている

本研究開発課題の目的

- タンパク質と糖鎖の結合の関係性を探る
- 糖鎖構造を解析するMCAWツールの完成
 - 糖鎖認識結合部位の特定をする

ツールの開発

- 複数の糖鎖構造から共通する構造パターンを可視化するWebツールMCAW (Multiple Carbohydrate Alignment with Weights) を開発した
- 糖鎖構造のアラインメントから共通する部分構造の割合を糖鎖構造プロフィールとして表し特徴を可視化する

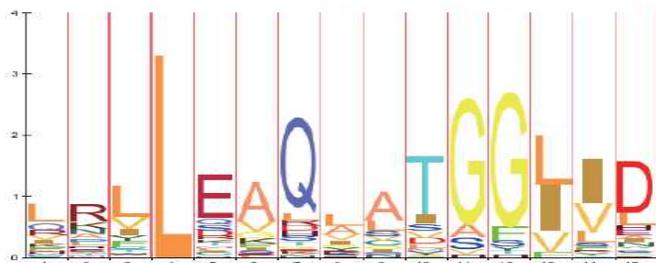
アミノ酸配列

```

1. proteinA_RAT      LRVLEAQAATGGLID
2. proteinA_MOUSE   QKLLEAQLATGGIID
3. proteinA_HUMAN   LRLLEAQLATGGIVL
4. proteinA_HORSE   LRLLEVQAATGGLVD
    
```

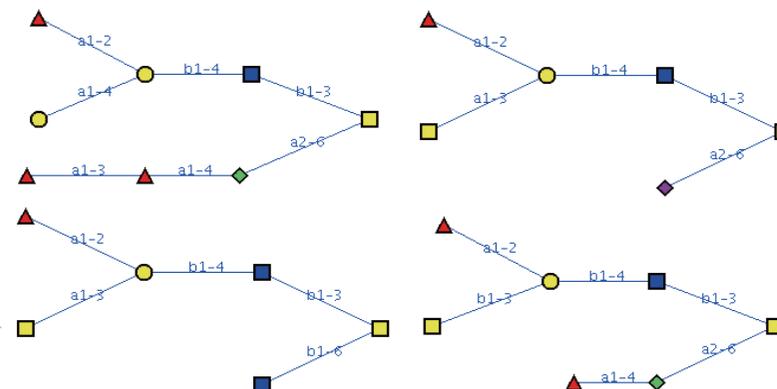


アミノ酸配列プロフィール

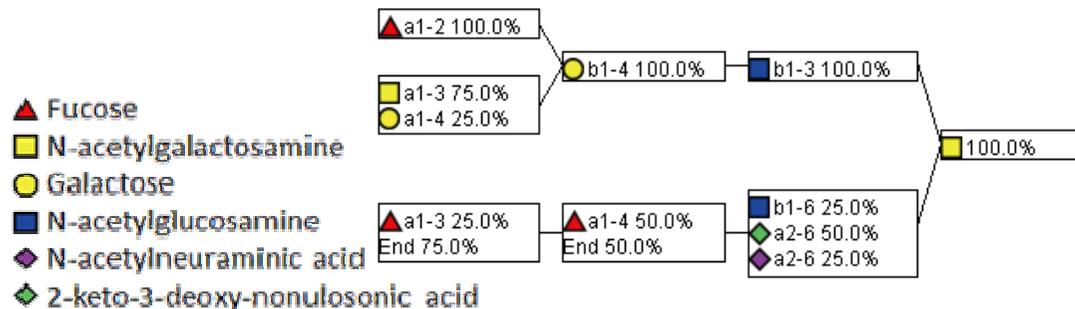


<http://pfam.sanger.ac.uk/>

糖鎖構造



糖鎖構造プロフィール



糖鎖のアライメント手順

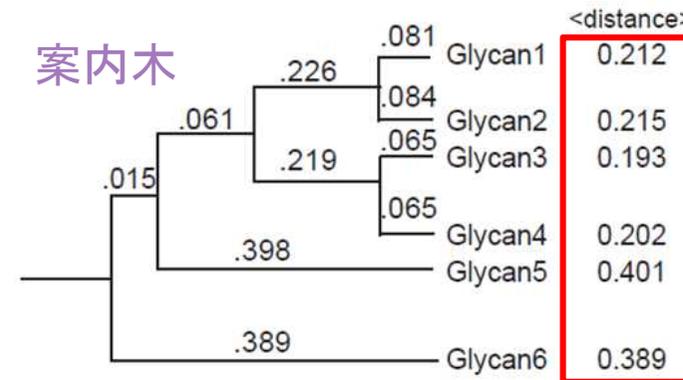
○ MCAWアルゴリズムはClustalWを参考に多重木アライメントを行う

1. 全糖鎖のペアワイズアライメント (KCaM*)を行い、距離行列を得る

Glycan	1	2	3	4	5
Glycan 1	-				
Glycan 2	.17	-			
Glycan 3	.59	.60	-		
Glycan 4	.59	.59	.13	-	
Glycan 5	.77	.77	.75	.75	-
Glycan 6	.81	.82	.73	.74	.80

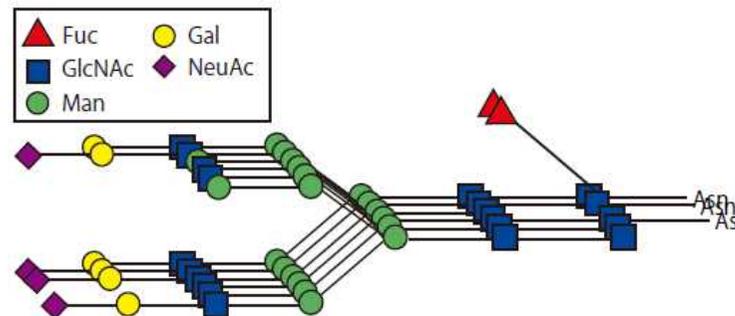
距離行列

2. 距離行列からFM法を用いて系統樹(案内木)を作成し、根からの距離を計算する



3. 案内木に従って糖鎖の多重木アライメントを行う

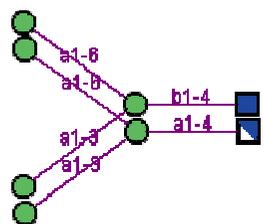
- アライメントのオーダー
- (1) Glycan1+Glycan2
 - (2) Glycan3+Glycan4
 - (3) (1)+(2)
 - (4) (3)+ Glycan5
 - (5) (4) + Glycan6



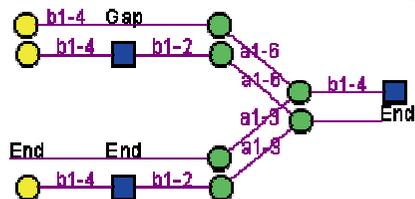
糖鎖構造のアライメント (糖鎖構造の比較)

- DP法を用いたKCaM*を参考にペアワイズアライメントの式から複数の糖鎖構造のアライメントの式に改良した

アライメント構造A



アライメント構造B



ダイナミックプログラミングアルゴリズム計算式

$$Q[u, v] = \max \left\{ \begin{array}{l} 0, \\ \max_{v_i \in \text{sons}(v)} \{Q[u, v_i] + d(v)\}, \\ \max_{u_i \in \text{sons}(u)} \{Q[u_i, v] + d(u)\}, \\ \frac{1}{|A||B|} \left\{ \sum_{n=1}^{|A|} \sum_{m=1}^{|B|} w(u_n, v_m) a_n b_m \right\} + \\ \max_{\psi \in M(u, v)} \left\{ \sum_{u_i \in \text{sons}(u)} Q[u_i, \psi(u_i)] \right\} \end{array} \right\}$$

$Q[u, v]$: 糖鎖アライメントのスコアを表す
 A_n, B_m 間の糖の類似スコア
 u と v : 比較するアライメントのポジション
 $\text{sons}(u)$ ($\text{sons}(v)$) : u (v)に結合する子ノード
 $d(u)$ ($d(v)$) : ギャップ

$w(u_n, v_m)$: u と v の位置にある糖鎖

a_n (b_m) : 糖鎖 A_n (B_m)の重み
 $M(u, v)$: 子ノードの組み合わせ

A (B) : u (v)が属するアライメントにおいての糖鎖の数(n, m)

MCAWツールの実装

○ RINGS(RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA)にて実装

<http://www.rings.t.soka.ac.jp/>

RINGS

Welcome to  RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

[User registration form](#) [Feedback search](#)

E-mail address Password

Search for display

RINGS is a web resource providing algorithmic and data mining tools to aid glycobiology research.
Many of the tools have been published in the literature;
RINGS provides free access to these methods for practical research.
Web services are also available via [BioMOBY](#) for programmers interested in incorporating these tools in custom programs.

Tools	Utilities
<ul style="list-style-type: none">• DrawRINGS: 2D Java-based glycan structure drawing tool for generating KCF format and/or querying the RINGS database, which current contains the glycan structures from GlycomeDB and an internally curated data set from the literature.• Glycan Miner Tool: implemented based on Hashimoto et al. 2008 for mining alpha-closed frequent subtrees from a set of glycan structures• Glycan Pathway Predictor (GPP) Tool: implemented based on Krambeck et al. 2005, which was later improved in Krambeck et al. 2009, for dynamically computing the <i>N</i>-glycan biosynthesis pathway from a given glycan structure• GlycomeAtlas: Visualization of glycome profiling data on human and mouse tissue samples, as described in Konishi et al. 2012.• Glycan Kernel Tool: Glycan structure classification tool, based on Jiang et al. 2011, for finding distinguishing glycan substructures in a target data set compared to a control• MCAW Tool: glycan structure multiple alignment tool. Align a set of glycan structures and visualize the result as a glycan profile. This tool can be used to find commonalities among a group of glycan structures.• Profile PSTMM Tool: implemented based on Aoki-Kinoshita et al. 2006, generate glycan profiles from glycan structure data which can be entered together with binding affinity data, for example, for a particular biological sample	<ul style="list-style-type: none">• GlycoCT to KCF: convert a glycan structure in GlycoCT format to KCF• GLYDE2 to KCF: convert a glycan structure in GLYDE2 to KCF• IUPAC to KCF: convert a glycan structure in IUPAC format to KCF• KCF to image: retrieve the image given a KCF• KCF to LinearCode: retrieve the LinearCode format given a KCF• KCF to LINUCS: retrieve the LINUCS format given a KCF• KCF to Mol: retrieve the chemical structure in MDL format given a glycan structure in KCF• KEGG GLYCAN ID to KCF: retrieve the KCF for a given KEGG GLYCAN ID• LinearCode to KCF: retrieve the KCF format given a LinearCode• LINUCS to KCF: convert a glycan structure in LINUCS format to KCF

MCAWツール(入力画面)

- http://www.rings.t.soka.ac.jp/cgi-bin/tools/MCAW/mcaw_index.pl



MCAW (Multiple Carbohydrate Alignment with Weights)

- [Home](#)
- [Help](#)
- [Feedback](#)

糖鎖構造
の入力

Data set name

Enter a glycan structure in KCF format:

```
ENTRY      G04845                      Glycan
COMPOSITION (Gal)3 (Glc)1 (GlcNAc)2 (LFuc)2 (Neu5Ac)1
MASS      1656.5
DBLINKS   CCSD: 23949
          GlycomeDB: 20420
          JCGGDB: JCGG-STR011245
NODE
1  Glc      0    0
2  Gal     -10   0
3  GlcNAc  -20   10
4  GlcNAc  -20  -10
5  Gal     -30   15
6  LFcuc   -30   5
7  LFcuc   -30  -5
8  Gal     -30  -15
9  Neu5Ac  -40   15
EDGE
8
1  2:b1  1:4
```

複数の糖鎖構造を
KCF形式で入力

入力はファイルを選
択することも可能

Or load KCF from a file: 選択されていません

Advanced weighting options

Gap penalty:

Monosaccharide:

- Linkage information -

Anomer:

Non reducing side carbon number:

Reducing side carbon number:

KCF 形式

ENTRY	G10652	Glycan	
NODE	6	X	Y
1	Asn	20	0
2	GlcNAc	10	0
3	GlcNAc	0	0
4	Man	-10	0
5	Man	-20	5
6	Man	-20	-5
EDGE	5		
1	2	1	
2	3:b1	2:4	
3	4:b1	3:4	
4	5:a1	4:6	
5	6:a1	4:3	

MCAWツール(入力画面)

- http://www.rings.t.soka.ac.jp/cgi-bin/tools/MCAW/mcaw_index.pl



MCAW (Multiple Carbohydrate Alignment with Weights)

- [Home](#)
- [Help](#)
- [Feedback](#)

糖鎖構造
の入力

Data set name

Enter a glycan structure in KCF format:

```
ENTRY      G04845      Glycan
COMPOSITION (Gal)3 (Glc)1 (GlcNAc)2 (LFuc)2 (Neu5Ac)1
MASS      1656.5
DBLINKS   CCSD: 23949
          GlycomeDB: 20420
          JCGGDB: JCGG-STR011245
NODE
  9
  1  Glc      0    0
  2  Gal     -10   0
  3  GlcNAc  -20  10
  4  GlcNAc  -20 -10
  5  Gal     -30  15
  6  LFuc    -30   5
  7  LFuc    -30  -5
  8  Gal     -30 -15
  9  Neu5Ac  -40  15
EDGE
  8
  1  2:b1  1:4
```

複数の糖鎖構造を
KCFで入力

入力はファイルを選
択することも可能

ギャップが入る時のペナルティスコア

Or load KCF from a file: 選択されていません

Advanced weighting options

Gap penalty:

Monosaccharide:

- Linkage information -

Anomer:

Non reducing side carbon number:

Reducing side carbon number:

単糖が一致したときのスコア

アノマーが一致したときのスコア

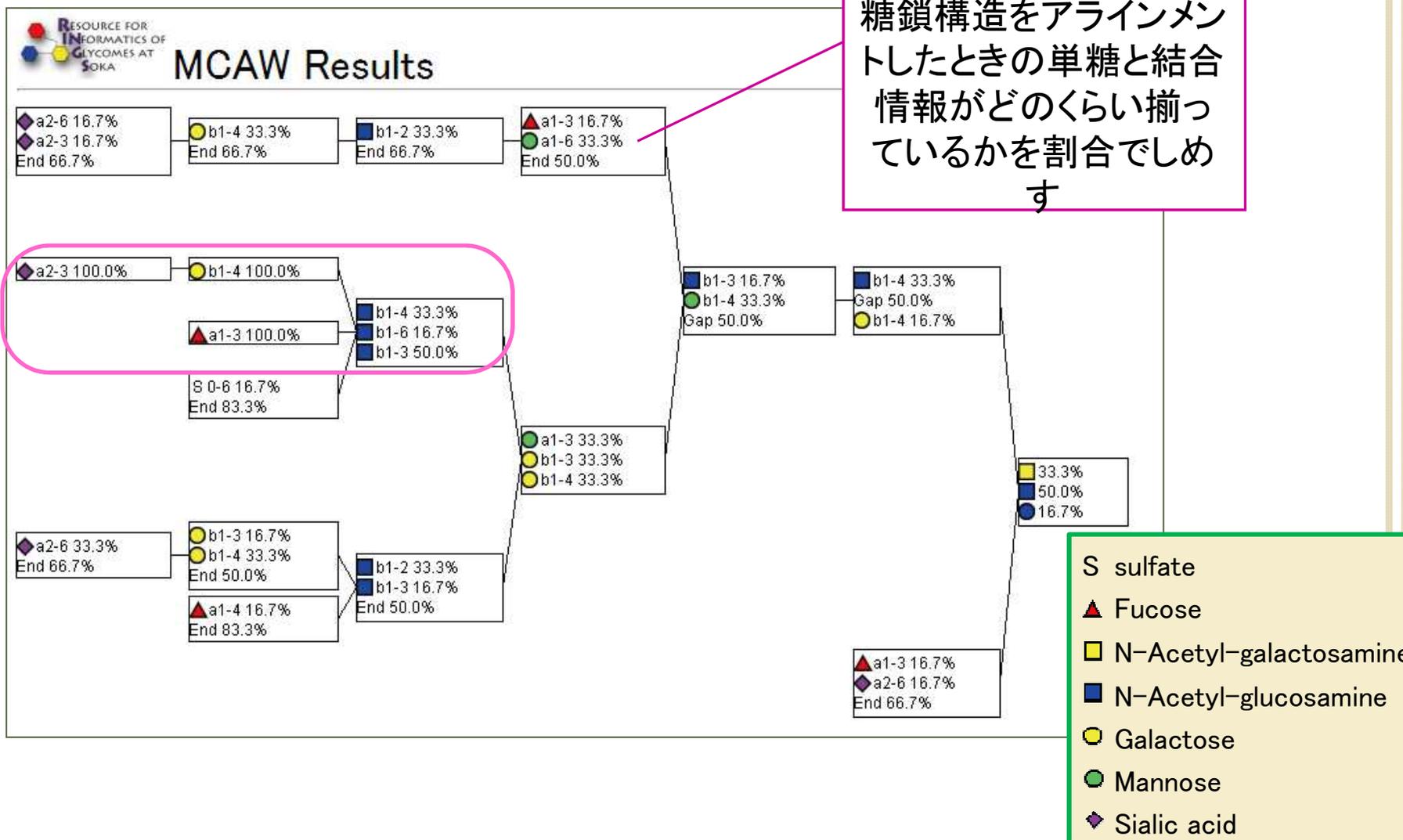
非還元末端側の炭素番号が一致
したときのスコア

還元末端側の炭素番号が一致
したときのスコア

アライメント時の
オプションスコア
の設定

実行

MCAWツール実行結果の例



統合データベースの活用について

- レクチンフロンティアデータベース
- レクチンに対する結合親和性を持つ糖鎖構造のデータ



LfDBとは

レクチンフロンティアデータベース(LfDB)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, Fig. 1)を用いて取得した各種レクチンとピリジルアミノ化糖鎖間の結合定数を含む定量的相互作用データを格納しています。本方法で取得した相互作用データの精度、信頼性は非常に高いことから、LfDBは今後、広く生物学研究において貴重な情報源になると考えられます。

Fig. 1. A photograph of the automated FAC-FD system

LfDBの仕様

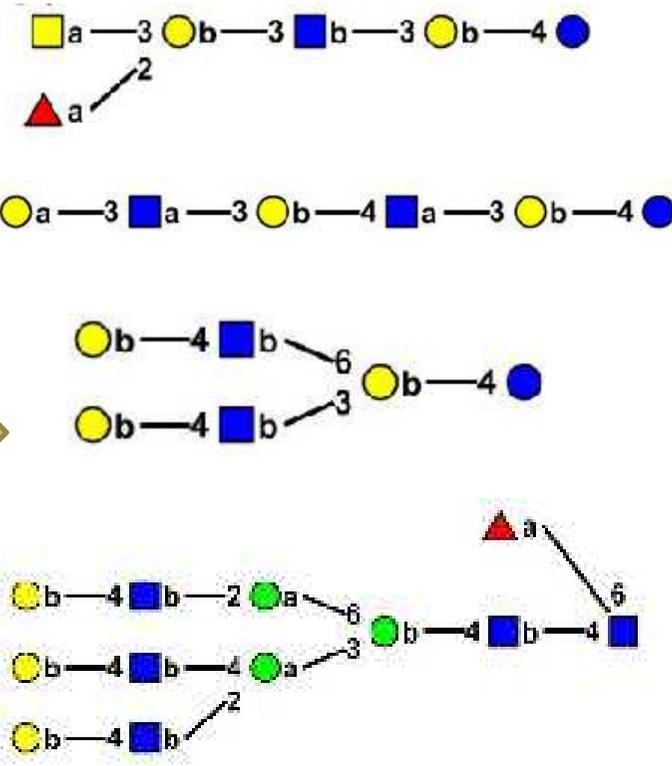
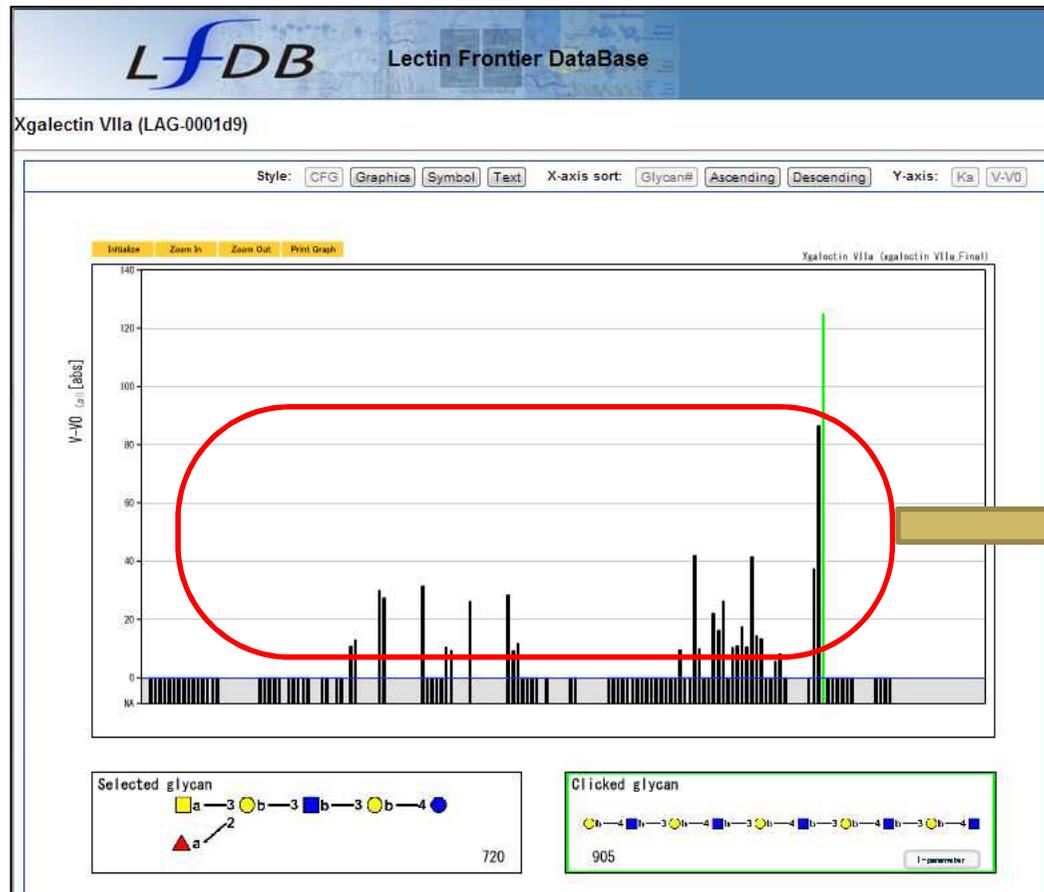
LfDBは、レクチンに関する基本情報を提供する「レクチン情報ページ」と、FAC-FDシステムによって得られた相互作用データを提供する「相互作用ページ」から構成されています。相互作用データは、レクチンのピリジルアミノ化糖鎖に対する実測値($V-V_0$)、及び結合定数(K_D)の棒グラフ形式で表示されます。また、LfDBはレクチンが認識する構造エレメントを見つける際に便利な、「1パラメータ差異検索」機能を有し

分類

- Lectin Family
 - Annexin
 - all(12);
 - binding to Others(6);
 - Chitin-binding Lectin
 - all(9);
 - binding to GalNAc(1);
 - binding to GlcNAc(8);
 - C-type Lectin
 - all(38);
 - binding to Fuc(5);
 - binding to Gal(6);
 - binding to GalNAc(3);
 - binding to GlcNAc(3);

統合データベースの活用について

- 結合定数が高い糖鎖構造を取得
- 取得した糖鎖構造の共通する部分構造を抽出



タンパク質と結合親和性を示す糖鎖構造の取得

LFDDB Lectin Frontier DataBase

検索条件

- キーワード: MornigaM
- 検索結果: 17件

レクチン名	生物種	ファミリー	単糖特異性	相互作用
BanLec	<i>Musa acuminata</i>	Jacalin-related Lectin	Man, Glc	
Conarva	<i>Convolvulus anensis</i>	Jacalin-related Lectin		
CRLL(前)	<i>Cycas revoluta</i>	Jacalin-related Lectin	Man, Glc	
PALa	<i>Phlebotium aureum</i>	Jacalin-related Lectin	Man, Glc	
Oryzata	<i>Oryza sativa</i>	Jacalin-related Lectin	Man, Glc	
PALb	<i>Phlebotium aureum</i>	Jacalin-related Lectin	Glc, Man	
CRLL	<i>Cycas revoluta</i>	Jacalin-related Lectin	Man	
CCA	<i>Castanea crenata</i>	Jacalin-related Lectin	Man, Glc	
MPIA	<i>Macleura pomifera</i>	Jacalin-related Lectin	Gal	
Jacalin	<i>Artocarpus integrifolia</i>	Jacalin-related Lectin	Gal	
MornigaM	<i>Morus nigra</i>	Jacalin-related Lectin	Man, Glc	
Parkia platycephala agglutinin	<i>Parkia platycephala</i> (ブラジル綿)	Jacalin-related Lectin	Man, Glc	
HTL	<i>Helianthus tuberosus</i>	Jacalin-related Lectin	Man	
Calsepa	<i>Caystegia sepium</i>	Jacalin-related Lectin	Man, Glc	
MornigaG	<i>Morus nigra</i>	Jacalin-related Lectin	Gal	
Helluba	<i>Helianthus tuberosus</i>	Jacalin-related Lectin	Man, Glc	
KM+ (Artocarpin)	<i>Artocarpus integrifolia</i>	Jacalin-related Lectin	Man, Glc	

LFDDB Lectin Frontier DataBase

GGDB | LFDDB | GMDB | GlycoProtDB | English

Top > LFDDB > Search Result > Detail

MornigaM (LPJ-0000bb)

説明

- レクチンフロンティアデータベースとは
- FAC-FDBの原理
- LFDDB使用方法
- 参考文献
- お問い合わせ

検索

キーワード検索

AND/OR検索

高度検索

分類

- Lectin Family
 - Annexin
 - Chitin-binding Lectin
 - C-type Lectin
 - Fucose-binding Lectin
 - Galectin
 - Jacalin-related Lectin
 - all(17);
 - binding to Fuc(0);
 - binding to Gal(3);
 - binding to Man(13);
 - binding to Glc(11);
 - Legume Lectin
 - Lectin-like protein
 - Mann6P-binding Lectin
 - Monocot Mannose-binding Lectin
 - R-type Lectin
 - Siglec
 - Others
 - Monosaccharide Specificity
 - 3D-fold

レクチンID	LPJ-0000bb
レクチン名	MornigaM
ファミリー	Jacalin-related Lectin
単糖 特異性	Man, Glc
生物種(学名)	Black Mulberry (<i>Morus nigra</i>)
日本語名	ブラックマルベリー
生物界	Plant
器官	bark
特定の器官	
CRD数	1
3D-fold	β -prism 1
外部リンク	GenBank: AY048577, AAL10685, AAL09163 PDB: Pfam: PF01419
配列ID	
配列	1 MAGISTINIQT IGTISQIVEVGLWGGPFGGNW DDGSGYTGIRE INLHSHDAIG AFSVIYDLNG 61 QPFTGTPHFG NEFSFKTVKI TLDFFNEFLV SVSGYIGVLP RLAIKGDVIR SLTFKTKNKI 121 YGPGKKEGT FFSLSPIENGL IVGFKGRSGF VVDAIGVHLS L

分子構造

3D structure Binding Site

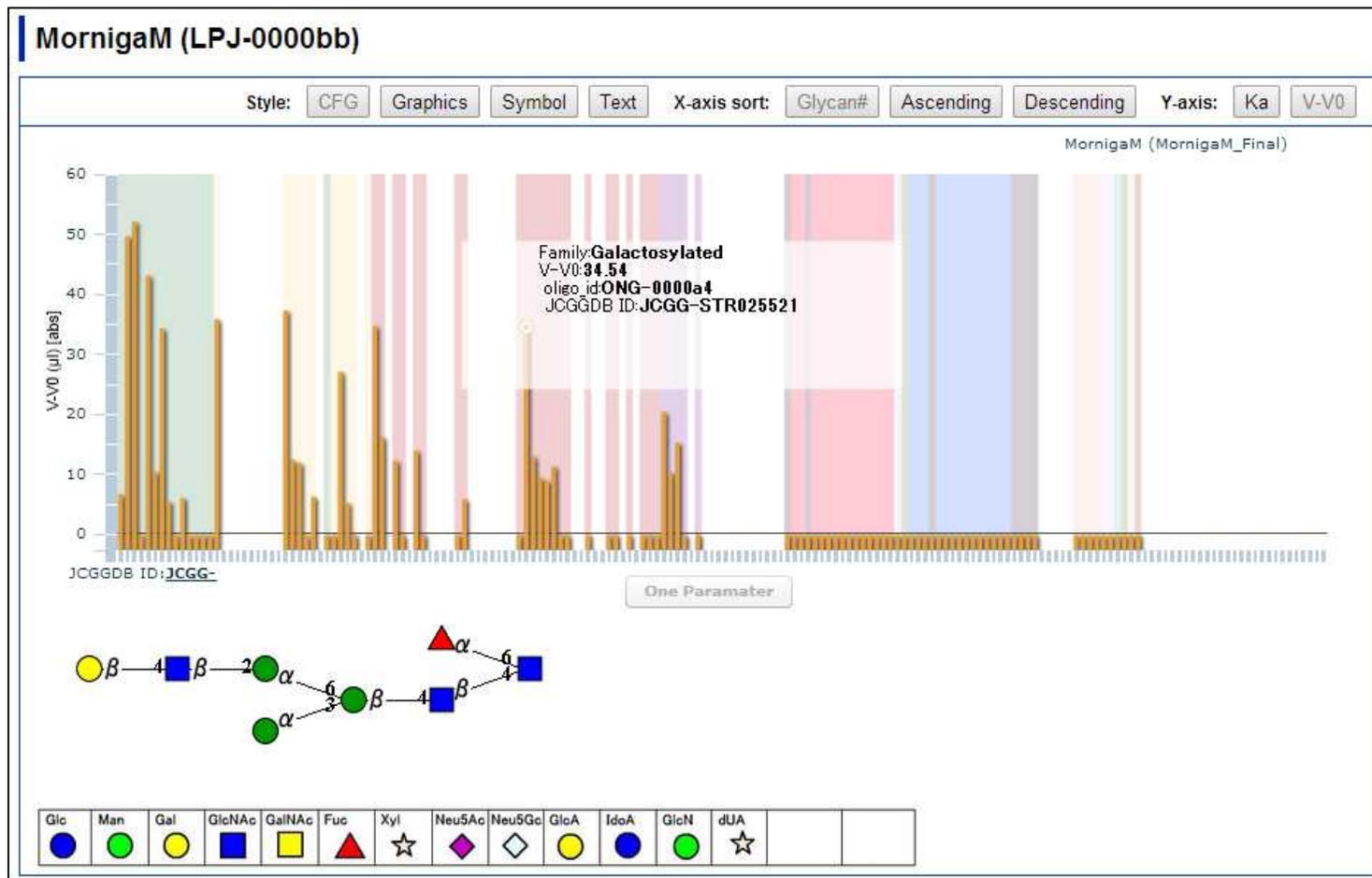
no data

Interaction Graph [Viewer] [GlycanList]

参考文献

データベースから糖鎖構造の取得

- 親和性を持つ糖鎖構造をツールで扱える形式に変換した



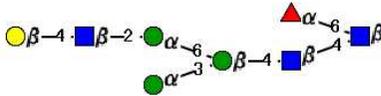
データベースから糖鎖構造の取得

JCGGDB English 日本語

JCGGDB Home キーワード 糖鎖構造 Glycoscience Explorer (α version)

JCGGDB ID **JCGG-STR025521** version 2

Glycan Information



Composition

Descriptors Computed from Structure

Mass

Glycan Motifs

ID History Deleted Version

External Links

JCGG JGD-000451

GlycomeDB (Structure) [4566](#)

KEGG (Structure) [G10773](#)

GMDB (MSⁿ) [ONG-0000a4](#)

Glycan Number: 401

RicinB [RCA1](#) [RCA120](#)

Unclassified [ELySL](#) [EFiaL](#) [HypninA1](#) [ECA](#) [ECL](#) [CRLL](#) [HypninA3](#) [Ecafl](#) [BRL](#) [C16](#) [CG-16](#) [TxLCI](#) [BRA2](#) [HypninA2](#) [PPL](#) [EW29-Ch](#) [AOL](#) [Conarva](#) [Orysata](#) [Ecri](#) [ECrL](#) [MornigaM](#) [EvesL](#) [AMA](#) [ECorL](#) [Calsepa](#) [Heltuba](#) [KM+](#) (Artocarpin) [AAL](#)

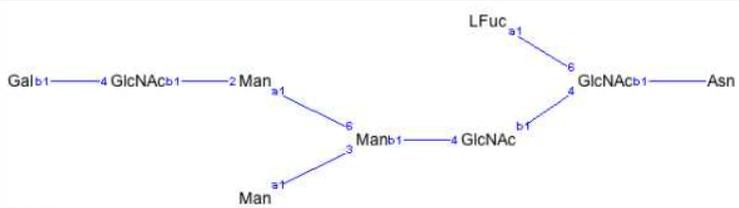
KEGG GLYCAN: G10773 Help

Entry G10773 Glycan

Composition (Gal)1 (GlcNAc)3 (LFuc)1 (Man)3 (Asn)1

Mass 1404.3 (Asn)

Structure



G10773
KCF file KcAM KegDraw

Class Glycoprotein; N-Glycan
[BRITE hierarchy](#)

Reference 1 [PMID:12054586] Oiczak K, Watorek W. Structural analysis of N-glycans from human neutrophil azurocidin. Biochem. Biophys. Res. Commun. 293 (2002) 213-9.

Other DBs CCSD: 5091 7793 7794 7795
GlycomeDB: 4566
JCGGDB: JCGG-STR025521

KCF data

ENTRY NODE	G10773	Glycan
1	Asn	30 3
2	GlcNAc	20 3
3	LFuc	10 8
4	GlcNAc	10 -2
5	Man	0 -2
6	Man	-10 3
7	Man	-10 -7
8	GlcNAc	-20 3
9	Gal	-30 3
EDGE		
1	2:b1	1
2	3:a1	2:6
3	4:b1	2:4
4	5:b1	4:4
5	6:a1	5:6
6	7:a1	5:3
7	8:b1	6:2
8	9:b1	8:4
/// KCFデータ		

糖鎖構造の形式変換

- インターネット上で公開しているRINGSでMCAWツールで扱える形式に変換した
- 糖鎖構造の描画から糖鎖構造情報のデータに変換できるツール (DrawRINGS)を利用した

E-mail address Password

Welcome to  RESOURCE FOR INFORMATICS OF GLYCOMES AT SOKA

[User registration form](#) [Feedback search](#)

Search for display [Advanced Search](#)

RINGS is a web resource providing algorithmic and data mining tools to aid glycobiology research.
Many of the tools have been published in the literature;
RINGS provides free access to these methods for practical research.
Web services are also available via [BioMOBY](#) for programmers interested in incorporating these tools in custom programs.

Tools	Utilities
<ul style="list-style-type: none">• DrawRINGS: 2D Java-based glycan structure drawing tool for generating KCF format and/or querying the RINGS database, which current contains the glycan structures from Glycome-DB and an internally curated data set from the literature.• Glycan Miner Tool: implemented based on Hashimoto et al., 2008 for mining alpha-closed frequent subtrees from a set of glycan structures• Glycan Pathway Predictor (GPP) Tool: implemented based on Krambeck et al., 2005 which was later improved in Krambeck et al., 2009, for dynamically computing the <i>N</i>-glycan biosynthesis pathway from a given glycan structure• GlycomeAtlas: Visualization of glycome profiling data on human and mouse tissue samples.• Glycan Kernel Tool: Glycan structure classification tool for finding distinguishing glycan substructures in a target data set compared to a control.• MCAW Tool: glycan structure multiple alignment tool. Align a set of glycan structures and visualize the result as a glycan profile. This tool can be used to find commonalities among a group of glycan structures.• Profile PSTMM Tool: implemented based on Aoki-Kinoshita et al., 2006, generate glycan profiles from glycan structure data which can be entered together with binding affinity data, for example, for a particular biological sample	<ul style="list-style-type: none">• GlycoCT to KCF: convert a glycan structure in GlycoCT format to KCF• GLYDE2 to KCF: convert a glycan structure in GLYDE2 to KCF• IUPAC to KCF: convert a glycan structure in IUPAC format to KCF• KCF to image: retrieve the image given a KCF• KCF to LinearCode: retrieve the LinearCode format given a KCF• KCF to LINUCS: retrieve the LINUCS format given a KCF• KCF to Mol: retrieve the chemical structure in MOL format given a glycan structure in KCF• KEGG GLYCAN ID to KCF: retrieve the KCF for a given KEGG GLYCAN ID• LinearCode to KCF: retrieve the KCF format given a LinearCode• LINUCS to KCF: convert a glycan structure in LINUCS format to KCF

Documentation

- [Help](#): brief users manual for the tools and utilities provided
- [What's new?](#): latest updates

Downloads



Many of the methods behind the tools provided here (and more) are described in detail in [Glycome Informatics: Methods and Applications](#)



[Kinoshita Laboratory](#)
[糖鎖インフォマティクス 若手の会HP](#)
[Faculty of Engineering](#)
[Soka University](#)

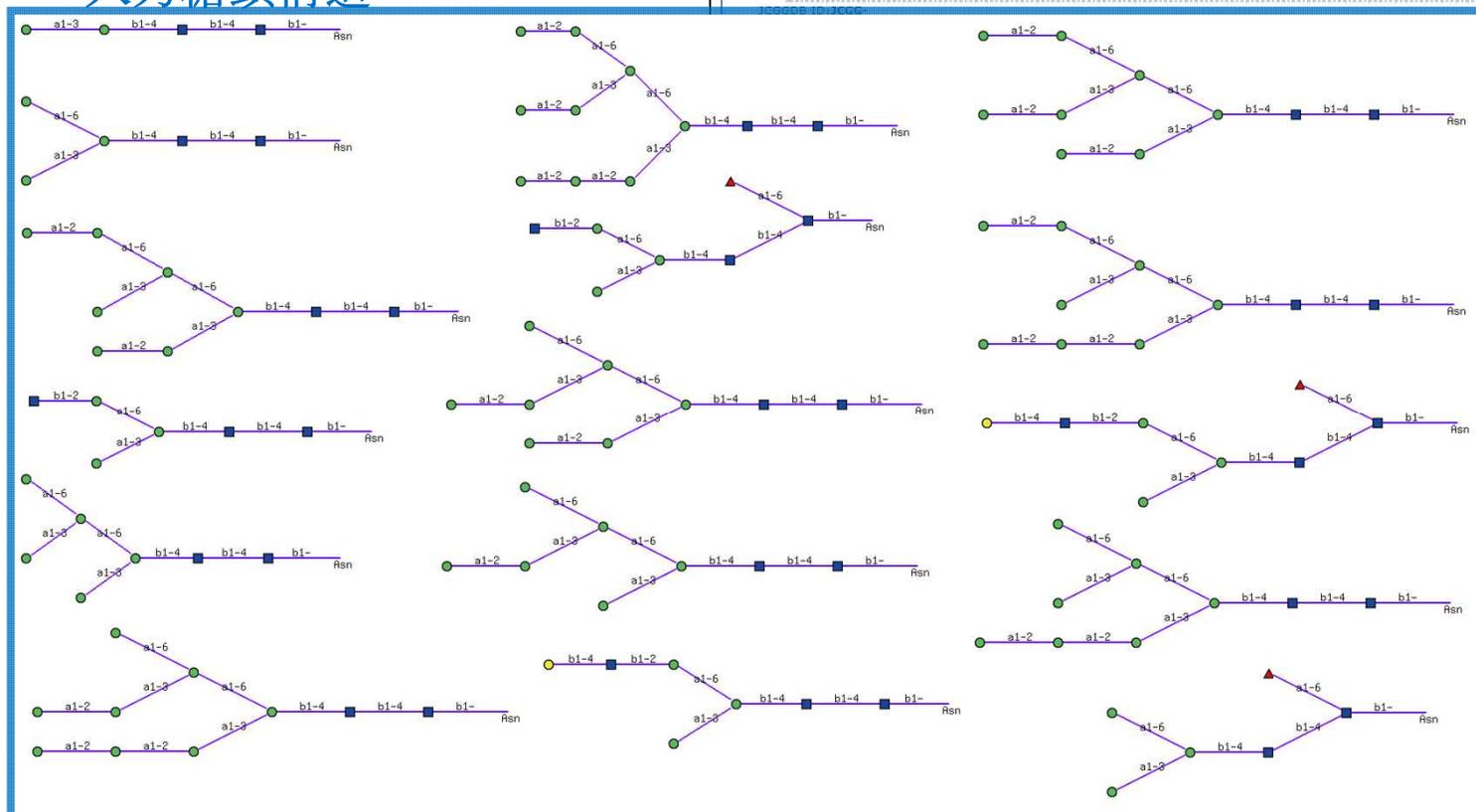
This site has been supported by Grant-In-Aid for Young Scientists (A), KAKENHI (20016025), the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT).

Last modified: August 2, 2012 You are visitor number 006438 !!

MCAWツールの活用

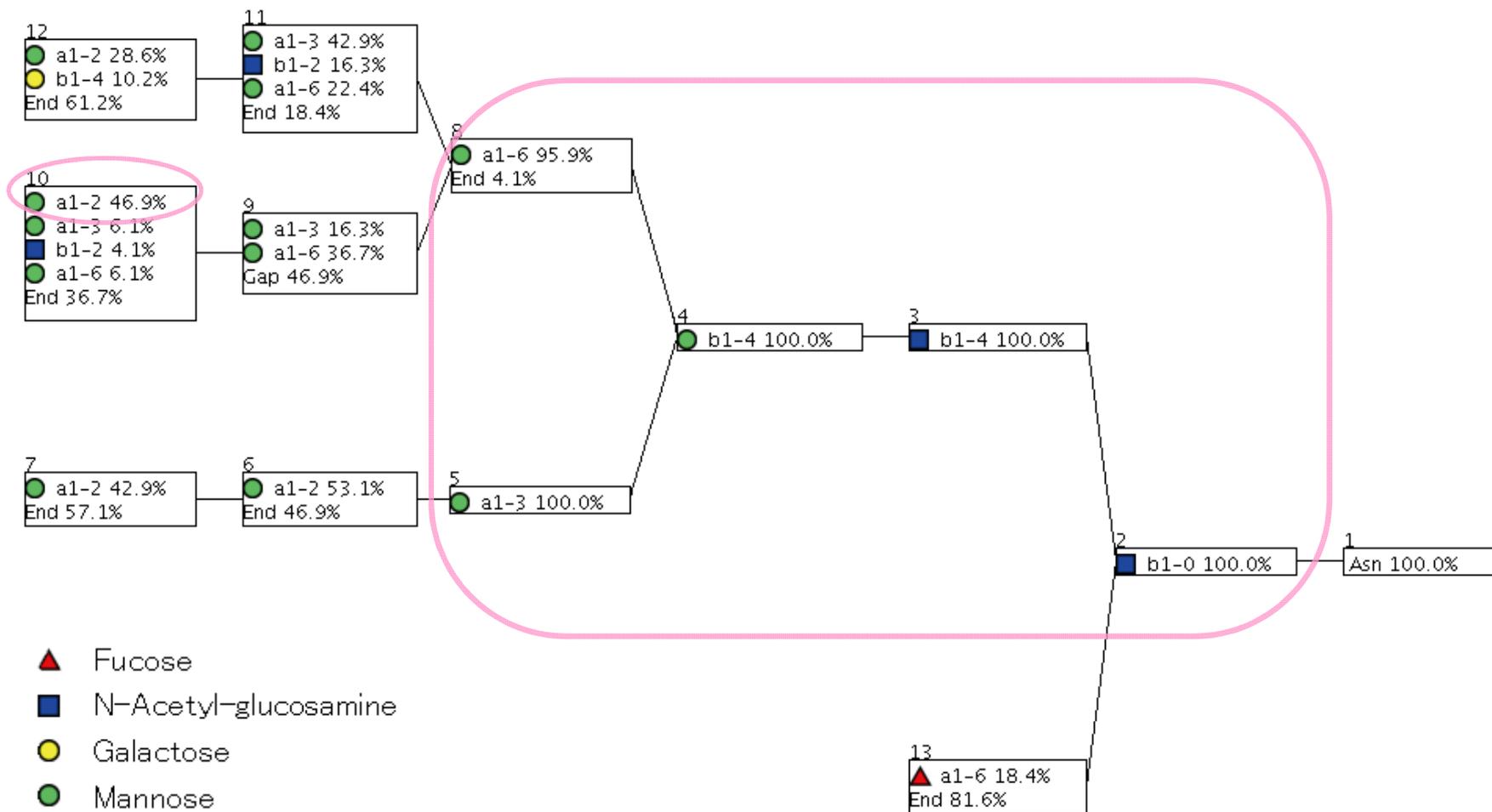
- CRLL (ソテツの葉)
jacalin関連レクチン
マンノース特異性

入力糖鎖構造



MCAWツール実行 結果

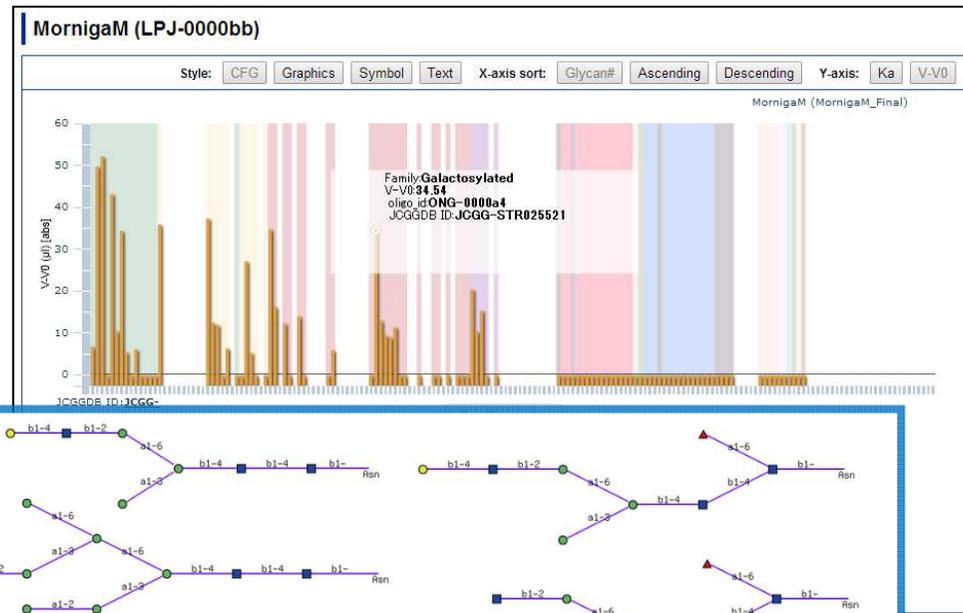
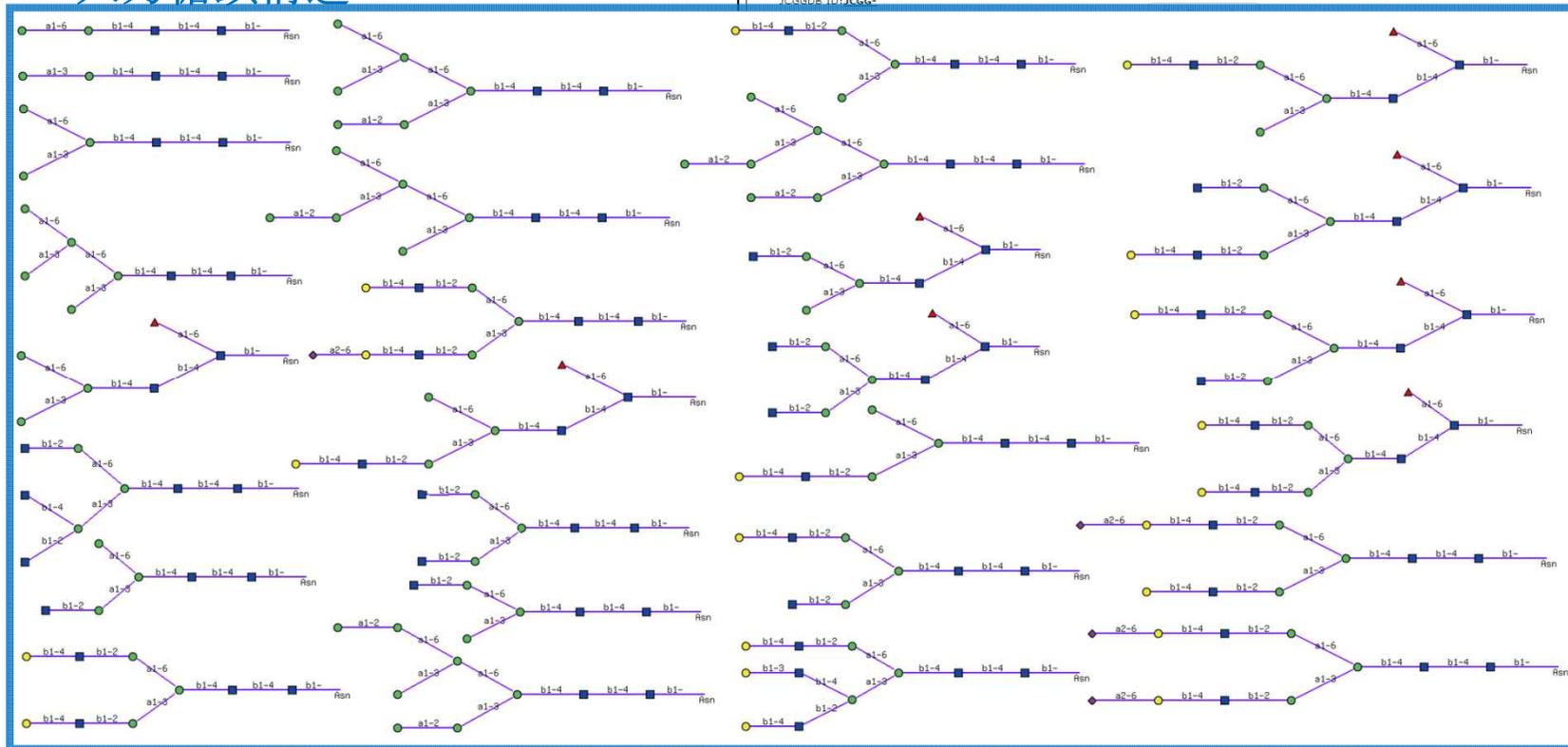
- 単糖の結合構成もタンパク質結合に関与する可能性が考えられる



MCAWツールの活用

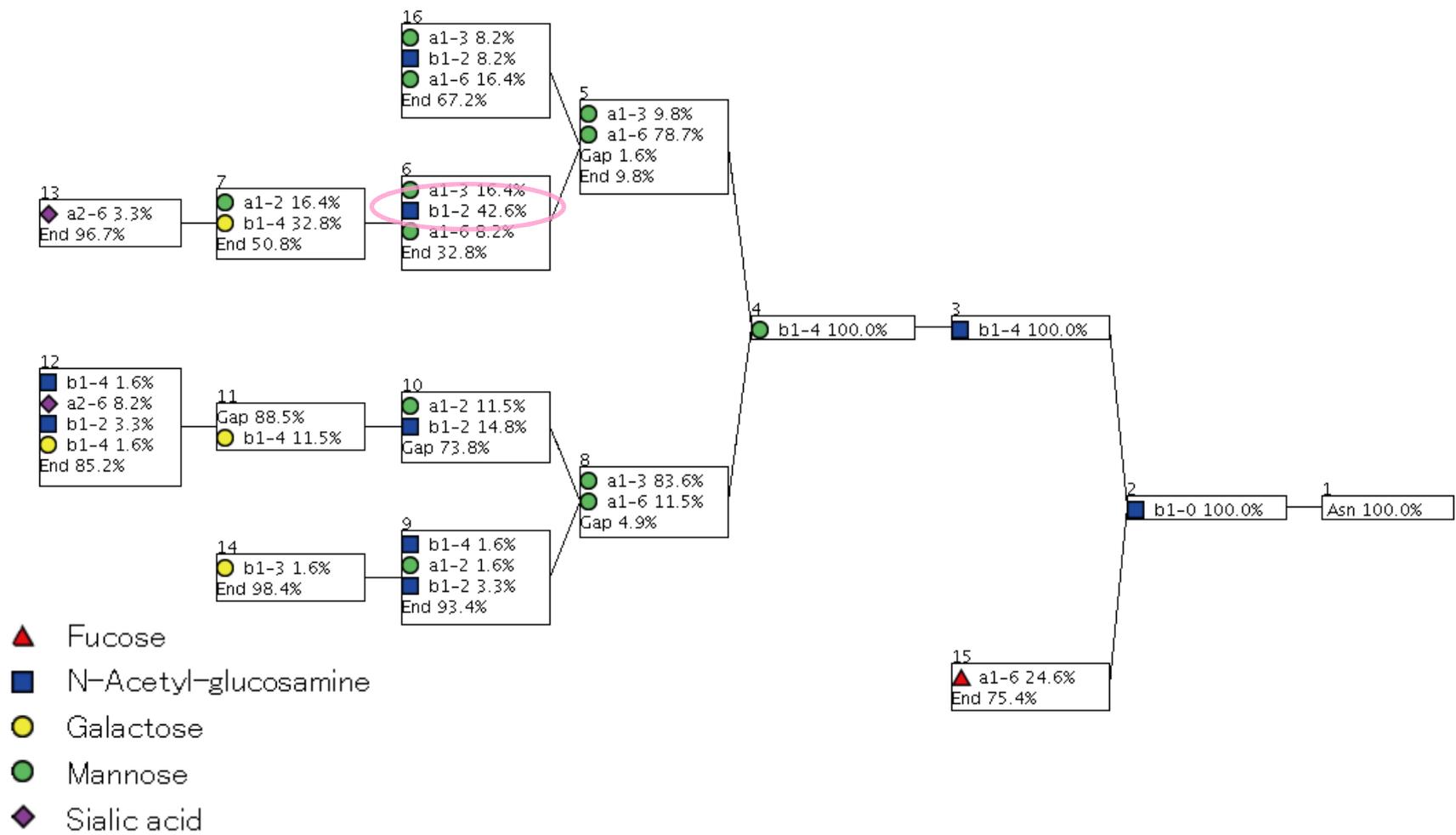
- MornigaM (ブラックマルベリー) jacalin 関連レクチン

入力糖鎖構造



MCAWツール実行 結果

- タンパク質との結合に既知以外の部位の関与が可能性として考えられるアラインメントを得ることができた



今後の本研究の将来性

- 糖鎖が持つ新たな共通構造の発見
- 機能特定のための糖鎖モチーフ生成
- 単糖同士の類似度を表すスコア行列の生成
- さらに精度の高いアラインメントを計算できるなど、解析の幅が広がると考えられる