

「統合データベースプロジェクト」補完課題

「糖鎖修飾情報とその構造解析データの統合」

(糖鎖科学統合データベースの構築)

21年度 研究成果報告書

平成22年3月

独立行政法人産業技術総合研究所

成松 久

本報告書は、文部科学省の平成21年度科学技術試験研究委託事業による委託業務として、独立行政法人産業技術総合研究所が実施した平成21年度「糖鎖修飾情報とその構造解析データの統合（糖鎖科学統合データベースの構築）」の成果を取りまとめたものです。

内容

1. 委託業務の目的	4
2. 平成21年度（報告年度）の実施内容	4
2.1 実施計画	4
2.2 実施内容（成果）	7
2.3 成果の外部への発表	13
2.4 活動（運営委員会等の活動等）	15
2.5 実施体制	16
2.6 整備実績一覧	16
別表1 平成21年度に於ける実施体制	16
別表2 整備実績一覧	20
DB名：GGDB	20
DB名：LfDB	22
DB名：GMDB	25
DB名：GlycoProtDB	27
DB名：GlycoEpitope	31
DB名：LipidBank	33
DB名：MDHPLCDB	34
DB名：PACDB	37
DB名：糖鎖関連疾患とその原因遺伝子のデータベース	40
DB名：JCGGDB Reports	42
DB名：JCGGDB GlycoScience Protocol Online Database	44

1. 委託業務の目的

糖鎖業界に散在するデータベースを産業技術総合研究所・糖鎖医工学研究センター（以下、糖鎖医工学研究センター）に設置予定のデータベースサーバに集約し糖鎖科学統合データベースを構築する。まずは、糖鎖医工学研究センターが構築した5種類のデータベース（糖鎖関連データベース、糖転移酵素特異性に関するデータベース、レクチンデータベース、糖タンパク質データベース、および質量分析データベース）について、統合を行う。

また、我が国に存在する糖鎖関連データベースを統合するために、糖鎖医工学研究センターは糖鎖関連データベースを保有している研究機関と交渉し、了解を得た上でデータの提供元となるよう促す役割を行う。そのために、各データ提供機関が糖鎖医工学研究センターに円滑にデータを提供できるように糖鎖医工学研究センターは必要な支援を行う。複数の機関から得られたデータを標準化した上でデータベースに格納する。

糖鎖医工学研究センターはデータベースに格納した専門的な情報を直感的に理解できるインターフェースを開発し、糖鎖研究分野以外の研究者等にも理解可能な情報として公開する。

各種様々なデータをいくつかのカテゴリに分類した上で、統合に必要な情報を中核機関である情報・システム研究機構に受け渡し、情報・システム研究機構の「DB ポータル」等の検索方法と連携できるように開発を行う。

最終的には情報・システム研究機構等とのデータの統合を目指し、統合検索を共同で開発する業務を行う。

2. 平成21年度（報告年度）の実施内容

2.1 実施計画

糖鎖医工学研究センターが平成19年度から構築している横断検索を基盤にし、国内に散在する糖鎖関連データベースを集約しコンテンツの充実を図る。国内に存在する、より多くの糖鎖関連データベースを統合するために、平成21年度も引き続き糖鎖関連データベースを保有している各研究機関との交渉を行い、当プロジェクトと連携するためのAPI開発の支援および新規データベース構築の支援を行っていく。

①運営と開発体制の準備

既に公開しているポータルサイトおよび横断検索サイトの内容を充実させるための活動並びにデータベース普及のための活動をしていく。データを保有している研究機関に統合データベースの活動を説明し、平成22年度（次年度）の協力機関としての参画を促す。そして理解が得られた機関を含めた平成22年度（次年度）の業務計画を立てる。

最終目標である統合検索システムを実現させるために、開発体制と情報基盤を整え、統合検索システムのプロトタイプ版の開発を開始する。

参加募集の活動以外にも、糖鎖に関連するライフサイエンス分野の研究者に、統合検索システムの活用方法を広めていく普及活動を行う。

②データ提供機関との交渉

データ統合の活動をしやすくする為に、国内の数多くの主要な研究機関・大学・企業が集結している最大の糖鎖の団体である「日本糖鎖科学コンソーシアム（JCGG）」に協力を要請し、統合検索システム構築の活動を JCGG の枠組みの中の活動として公認を得た。既に糖鎖関連データベースを保有している機関に協力機関として参画するように促す。データベース化していないデータを保有している機関には糖鎖科学統合データベースの枠組みの中で構築し統合するように促す。

データの提供時に発生する作業に関して、個々の研究機関・企業・大学と交渉し、より良い手段で解決し、円滑にデータを統合できるように実務レベルの話し合いを進め、業務を遂行する。

③統合データベース構築

糖鎖医工学研究センターが平成20年度に構築したデータベースを基盤として、糖鎖科学統合データベースのコンテンツと検索機能を拡充する。大きく分けて主に2つの開発を行う。まずは、それぞれの研究機関のDBのデータの属性から研究領域に分類し、振り分けた研究領域の中で融合できる統合検索システムを構築する。もう一つは、研究領域毎に細かく統合した検索システムを1つにまとめるための技術開発を行うことと、中核機関のライフサイエンス統合DBとの連携するための作業を行う。

③-1 研究領域毎の統合データベース構築

i 「構造解析と検出に関するカテゴリー」（構造の横断検索）

平成20年度に構築した「構造解析と検出に関するカテゴリー」（構造の横断検索）に理化学研究所の「糖鎖コンフォメーションデータベース」が参加することになりAPIの開発を行い、構造による横断検索システムで連携を開始する。

ii オンラインプロトコルの統合検索システム構築

「糖鎖大量合成」研究領域の中で、財団法人野口研究所と共同で有機化学による糖鎖合成とその後末端構造にバラエティを持たせるための酵素合成法の2つを融合させて糖鎖標準品の合成を支援するシステムを2年かけて開発する。外部の研究者が反応経路と合成に関わるノウハウを登録できるシステムを開発し、糖鎖の標準品の合成法を誰でも検索できるよう目指す。これ以外のプロトコルに関して、糖鎖研究の初心者でも糖鎖の解析を行えるように各種様々な解析に関するオンラインプロトコルを整備する。それら

を実現するために研究機関や企業など幅広く募集を行い、より多くのプロトコルの蓄積を目指す。

iii 糖鎖関連疾患とその糖鎖関連遺伝子のデータベース

名古屋大学の協力を得てノックアウトマウスのデータベースを継続してデータ更新を行う。国内の研究者が保有しているノックアウトマウスの情報を研究者が自ら登録できる仕組みを構築する。更には登録したリソースの管理（配布方法・問合せ先等の情報も管理）ができるシステムも構築する。これに関連して、創価大学のショウジョウバエの糖鎖関連遺伝子データベース（以後、ショウジョウバエの糖鎖関連遺伝子データベースを FlyGlycoDB と略す）と連携するために API の構築を行う。そして、平成 21 年度内に糖鎖関連遺伝子のノックアウトマウス DB、創価大学の FlyGlycoDB 及び産総研の GGDB の 3 機関の DB で連携（左記、以後 3 機関連携 GGDB と略す）を行う。また、新規に糖鎖関連遺伝子が原因で引き起こされる疾患のデータベースを構築し、3 機関連携 GGDB と連携することを目標にする。

iv 糖鎖の機能に関する情報のデータベース

糖鎖の役割の重要性を広く知ってもらうために、一般の方から専門の方まで理解しやすい易しい読み物を作成して、データベース化していく予定とする。蓄積した文章を利用しデータマイニング技術を駆使して、糖鎖の専門用語と遺伝子名などを結びつける情報を抽出できる技術の確立を目指し、中核機関と連携しながら作業を進める。そのための糖鎖の専門用語集作成などを 2 年かけて整備する。

v 経産省側の統合 DB プロジェクトとの連携

糖鎖医工学研究センターのデータベースに遺伝子多型やタンパク質ドメイン構造等の H-invDB の情報を表示できるように、平成 20 年度には産業技術総合研究所・メディシナル情報研究センター（旧・生物情報解析研究センター）の代表者と協議を行い API の開発を行った。平成 21 年度はお互いの DB の画面に補完してメリットのあるデータを表示させる（経産省側との DB 連携）。

③-2 統合データベース構築に向けた開発と中核機関との連携

i 検索用インデックスの自動更新システムの開発

これまでに構築した外部 DB と連携するための API を活用し、横断検索のインデックスを自動更新するシステムの開発を行う。

ii 直観的なポータルサイト実現のための技術開発

平成 20 年度に完成した API を利用し統合検索システムを構築し始める。平成 21 年度開発する API が完成した場合には統合検索システムに組み込む。そして平成 21 年度から統合検索システム構築のために、中核機関と連携し、平成 22 年度末までに実現可能になるように情報の整理と web インターフェースの開発を 2 年かけて行う。

2. 2 実施内容(成果)

糖鎖医工学研究センターが平成19年度から構築している横断検索を基盤にし、国内に散在する糖鎖関連データベースを集約し糖鎖科学のポータルサイト(JCGGDB)を運用しているが、そのコンテンツの充実を進めるとともに、平成21年度も引き続き糖鎖関連データベースを保有している研究機関と交渉を行い、当プロジェクトと連携するためのAPI開発の支援および新規データベース構築の支援を行った。具体的な作業とその結果は以下のとおりである。

①運営と開発体制の準備

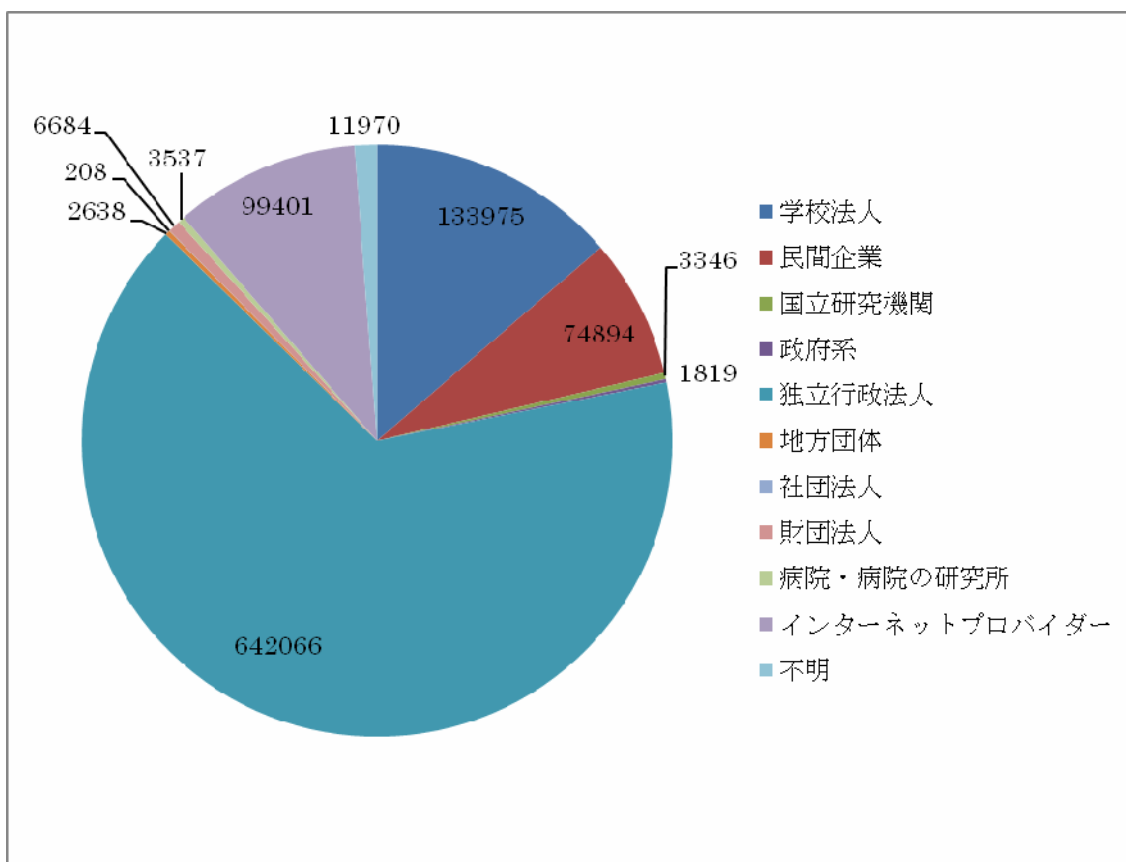
既に公開しているポータルサイトおよび横断検索サイトの内容を充実させるための活動並びにデータベース普及のための活動を行なった。データを保有している研究機関に統合データベースの活動を説明し、平成22年度の協力機関としての参画を依頼した。そして理解が得られた機関を含めた平成22年度の業務計画を立てた。最終目標である統合検索システムを実現させるために、開発体制と情報基盤を整え、統合検索システムのプロトタイプ版の開発を開始した。

参加募集の活動以外にも、糖鎖に関連するライフサイエンス分野の研究者に、統合検索システムの活用方法を広めていく普及活動を行った。日本糖質学会に於いて、データベース展示ブースを設置し、使い方の説明やプロジェクトに参加している研究機関のデータベースの解説などを行った。

下記にJCGGDBの各国のページビューを記載した。アジア(日本、中国、韓国)を中心に米国、英国が主にアクセスしている。

国名(上位5ヶ国)	年間ページビュー
JAPAN	980538
UNITED STATES	111119
CHINA	16200
KOREA, REPUBLIC OF	14623
UNITED KINGDOM	3261

日本の利用者の内訳

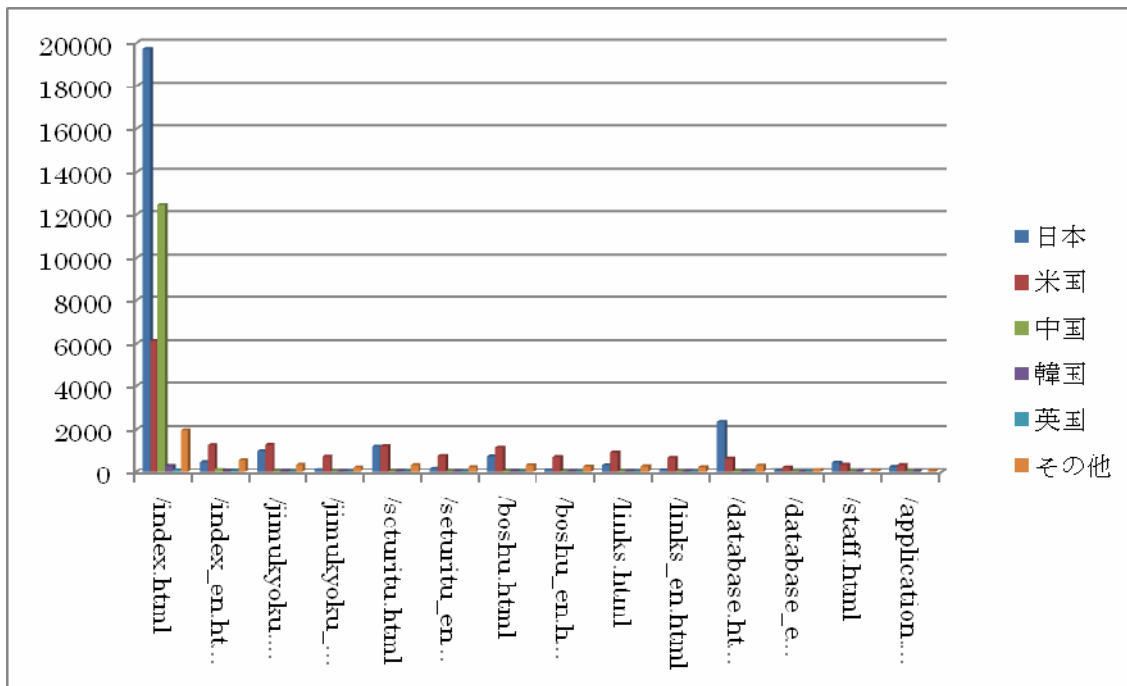


分類	アクセス数	パーセント
学校法人	133975	13.66
民間企業	74894	7.64
国立研究機関	3346	0.34
政府	1819	0.19
独立行政法人	642066	65.48
地方団体	2638	0.27
社団法人	208	0.02
財団法人	6684	0.68
病院・病院の研究所	3537	0.36
インターネットプロバイダー	99401	10.14
不明	11970	1.22
合計	980538	

(2010年3月31日現在)

国別のページアクセス数 (HTMLレベル)

日本からのアクセスでは、全体的に TOP ページの参照数が圧倒的に多い。その次にデータベース・サービス、設立、事務局の説明の順位多い。英語のページに関して、米国からのアクセスは事務局、設立、トップページ、募集、データベースの順に多い。



	日本	米国	中国	韓国	英国	その他
/index.html	19738	6131	12437	318	93	1965
/index_en.html (英語)	485	1245	133	45	41	571
/jimukyoku.html	981	1258	29	31	8	371
/jimukyoku_en.html (英語)	140	736	7	21	7	248
/seturitu.html	1182	1200	26	30	8	357
/seturitu_en.html (英語)	188	769	5	21	11	259
/boshu.html	742	1139	27	38	8	354
/boshu_en.html (英語)	117	718	9	20	7	290
/links.html	355	914	25	20	7	302
/links_en.html (英語)	100	673	12	16	8	259
/database.html	2339	643	20	19	16	331
/database_en.html (英語)	55	253	5	9	4	136
/staff.html	468	372	1	30	0	67
/application.html	272	357	2	31	0	47

サービス毎のアクセス数

キーワードによる横断検索

合計	日本	米国	中国	韓国	英国	その他
17366	10835	4681	52	181	120	1497

キーワード検索では日本からのアクセスが最多である。日本からのアクセスが多い理由としては日本のユーザ数が多く、日本語のドキュメントも検索対象になっている理由と思われる。

糖鎖構造検索

合計	日本	米国	中国	韓国	英国	その他
56382	22886	27836	34	2763	162	2701

構造検索に関して日本より米国からのアクセスの方が多い。

構造の詳細画面参照数

合計	日本	米国	中国	韓国	英国	その他
134754	3490	57744	17	1997	11	71495

JCGGDBでは各機関の構造名などを集約しているため Google などから構造名でヒットして直接ページにアクセスしてくるためと推定できる。構造検索から構造の詳細画面に遷移する数が少ない。恐らく使い方が解らずエントリーにヒットしていないためと想像できる。

病原体と糖鎖のデータベース

合計	日本	米国	中国	韓国	英国	その他
37141	5878	21334	19	4263	5	5642

米国が一番興味を持っている。日本、韓国の順に多い。公開から3, 4ヶ月しかたっていないが比較的アクセス数が多い。

②データ提供機関との交渉

データ統合の活動をしやすくする為に、国内の数多くの主要な研究機関・大学・企業が集結している最大の糖鎖の団体である「日本糖鎖科学コンソーシアム(JCGG)」に協力を要請し、統合検索システム構築の活動を JCGG の枠組みの中の活動として公認を得ている。既に糖鎖関連データベースを保有している機関に協力機関として参画するように依頼した。データベース化していないデータを保有している機関には糖鎖科学統合データベースの枠組みの中で構築し統合するようにした。

データの提供時に発生する作業に関して、個々の研究機関・企業・大学と交渉し、より良い手段で解決し、円滑にデータを統合できるように実務レベルの話し合いを進め、業務を遂行した。

③統合データベース構築

糖鎖医工学研究センターが平成20年度に構築したデータベースを基盤として、糖鎖科学統合データベースのコンテンツと検索機能を拡充した。大きく分けて主に2つの開発を行った。まずは、それぞれの研究機関のDBのデータの属性から研究領域に分類し、振り分けた研究領域の中で融合できる統合検索システムを構築した。例えば感染と疾患の関連や病原菌と糖鎖の結合情報をデータベース化してPACDBと名付けた。他にも糖鎖関連遺伝子の変異によって起こる病気の情報を収集（Glyco-Disease Genes Database と命名しGDGDBを略す）し遺伝子のデータベース（GDGDB）やショウジョウバエの糖鎖関連遺伝子のデータベース（FlyGlycoDB）との連携を行った。FlyGlycoDB がまだ公開されていないこともありGDGDBとGDGDBのクロスリンクレベルの連携のみ利用できる。もう一つは、研究領域毎に細かく統合した検索システムを1つにまとめるための技術開発を行ない、そのDBと中核機関のライフサイエンス統合DBとの横断検索連携に必要なインデックス化を行った。

③-1 研究領域毎の統合データベース構築

i 「構造解析と検出に関するカテゴリー」（構造の横断検索）

平成20年度に構築した「構造解析と検出に関するカテゴリー」（構造の横断検索）に理化学研究所の「糖鎖コンフォメーションデータベース」が参加することになりAPIの開発が完了し、構造による横断検索システムで連携できる状況になった。この連携は、糖鎖コンフォメーションデータベースの論文で発表した後にこのデータベースを公開することになっているので、JCGGDBとの連携部分の公開が遅れている。

ii オンラインプロトコルの統合検索システム構築

「糖鎖大量合成」研究領域の中で、財団法人野口研究所と共同で有機化学による糖鎖合成とその後末端構造にバラエティを持たせるための酵素合成法の2つを融合させて糖鎖標準品の合成を支援するシステムを2年かけて開発しているところである。野口研究所以外の研究機関や大学の方も巻き込んでより多くのデータが集まるように計画した。糖鎖の標準品の合成法を誰でも検索できるように、外部の研究者が反応経路と合成に関わるノウハウを登録できるシステムを開発した。糖鎖研究の初心者でも糖鎖の解析を行えるように各種様々な解析に関するオンライン実験プロトコルの整備をするために研究機関や企業など幅広く募集を行った。その結果、35名の著名な研究者に協力して頂き、54種類の実験プロトコルを電子化してDBに登録した。3月末日時点では未

公開であるが平成22年度の初めに著作権などの問題を解決してから公開する予定と
している。最終年度には合計で約100以上のプロトコルを公開することを目指してい
る。

iii 糖鎖関連疾患とその糖鎖関連遺伝子のデータベース

名古屋大学の協力を得てノックアウトマウスのデータベースを継続してデータ更新
を行った。国内の研究者が保有しているノックアウトマウスの情報を研究者が自ら登録
できる仕組みを構築した。更には登録したリソースの管理（配布方法・問合せ先等の情
報も管理）ができるシステムも構築した。これに関連して、創価大学のショウジョウバ
エの糖鎖関連遺伝子データベース（以後、ショウジョウバエの糖鎖関連遺伝子データベ
ースをFlyGlycoDBと略す）と連携するためにAPIの構築を行った。そして、糖鎖関連
遺伝子のノックアウトマウスDB、創価大学のFlyGlycoDB及び産総研のGGDBの3
機関のDBで相互リンクによる連携（左記、以後3機関連携GGDBと略す）を行った。

iv 糖鎖の機能に関する情報のデータベース

糖鎖の役割の重要性を広く知ってもらうために、一般の方から専門の方まで理解しや
すい易しい読み物を作成して、データベース化した。蓄積した文章を利用しデータマイ
ニング技術を駆使して、糖鎖の専門用語と遺伝子名などを結びつける情報を抽出できる
技術の確立を目指し開発しているところである。そのための糖鎖の専門用語集作成など
を2年かけて整備しているところである。まずは英語の文章から専門用語を抽出するプ
ログラムを開発しており、Enjuを用いた構文解析器による用語の収集を行った。用語
を抽出する時の品詞の並びのパターンファイルを作成しているところである。この方法
で収集した専門用語を用いてJCGGDBで保有している各種データベースへの自動
リンクを行えるシステムを開発しており、遺伝子名や糖鎖構造名などの単語とその情報
が記載されているリンク先の情報をデータベース化している。この作業と同時に文章中
にJCGGDBの各DBへ、リンクできる情報がないか検索するプログラムの開発を行
っている。

v 経産省側の統合DBプロジェクトとの連携

糖鎖医工学研究センターのデータベースに遺伝子多型やタンパク質ドメイン構造等
のH-invDBの情報を表示できるように、平成20年度には産業技術総合研究所・メディ
シナル情報研究センターの代表者と協議を行いAPIの開発を行った。平成21年度は
お互いのDBの画面に補完してメリットのあるデータを表示させた（経産省側とのDB
連携）。産総研・糖鎖センターが保有しているGGDBには遺伝子産物のアイソフォー
ムの情報を表示し、産総研・バイオメディシナル情報研究センターのH-InvDBにはGGDB
へのリンクを表示できるように協力した。また、糖鎖センターが保有するGlycoProtDB

に API を設置し、H-InvDB から GlycoPortDB にリンクできるようになった。マウスやヒトの糖タンパク質の情報を公開したときに、API の改良せずと同様にリンクできるように API の設置とデータ提供に協力した。H-invDB に GlycoProtDB の糖鎖修飾の情報を表示できるように協力した。

③-2 統合データベース構築に向けた開発と中核機関との連携

i 検索用インデックスの自動更新システムの開発

これまでに構築した外部DBと連携するためのAPIを活用し、横断検索のインデックスを自動更新するシステムの開発を行った。

ii 直観的なポータルサイト実現のための技術開発

平成20年度に完成したAPIを利用し統合検索システムを構築しているところである。③-1ivの技術を利用し専門用語の共出現頻度を利用しながら専門用語間の関連性を導きながらビューワで表示して直感的に検索や閲覧ができるシステムを開発しており、現在、ポータルサイトでアルファ版を公開している。このビューワで利用する単語の頻度を計測するには、まず「iv糖鎖の機能に関する情報のデータベース」に記載の専門用語を整備する必要があり、この作業が終わらない限り開発が進まない。引き続き平成22年度も継続して開発をしていく予定である。

2.3 成果の外部への発表

プレス発表、取材対応 (該当なし)

学会等への口頭発表

発表した成果(発表題目、口頭・ポスター発表の別)	発表者氏名	発表した場所(学会等名)	発表した時期	
○JCGGDBの概要と利用方法(データベース講習会)	鹿内俊秀	第3回糖鎖産業技術フォーラム(GLIT) 産業技術総合研究所・臨海副都心センター	2009年5月12日	国内
○日本糖鎖科学統合データベース(ポスター発表)	鹿内俊秀	第3回糖鎖産業技術フォーラム(GLIT) 産業技術総合研究所・臨海副都心センター	2009年5月12日	国内

◎日本糖鎖科学コンソーシアム データベース ～JCGGDB～ (ポ スター発表)	鹿内 俊秀	データベースが拓くこれか らのライフサイエンス・東京 大学・武田ホール (根津)	2009年6月 12日	国 内
◎日本糖鎖科学コンソーシアム のデータベース (ポスター発表)	鹿内 俊秀	第29回日本糖質学会・飛 騨・世界生活文化センター (岐阜県高山市)	2009年9月8 日～12日	国 内
◎日本糖鎖科学コンソーシアム のデータベース (ポスター発表)	鹿内 俊秀	第82回日本生化学会大会・ (神戸ポートアイランド)	2009年10月 24日	国 内
◎糖鎖構造に関連した実験デー タの横断検索(JCGGDB) (ポスタ ー発表)	鈴木 芳典	第82回日本生化学会大会・ (神戸ポートアイランド)	2009年10月 24日	国 内
○日本糖鎖科学コンソーシアム のデータベース構築 (口頭発表)	鹿内 俊秀	バイオジャパン 2009	2009年10月 7日	国 内
◎JCGGDBの公開したサー ビスと最終年度の計画 (口頭発 表)	鹿内 俊秀	第7回日本糖鎖科学コンソ ーシアム年会 (大阪)	2009年12月 8日	国 内
◎Japan Consortium for Glycobiology and Glycotechnology DataBase (JCGGDB) (口頭発表)	鹿内 俊秀	1st Asia communication of glycobiology and glycotechnology	2009年10月 31日	国 外
◎Japan Consortium for Glycobiology and Glycotechnology DataBase (ポ スター発表)	鹿内 俊秀	IUBMB (International Union of Biochemistry and Molecular Biology) (上海)	2009年8月2 日 ～8日	国 外
◎Japan Consortium for Glycobiology and Glycotechnology DataBase (JCGGDB) (口頭発表)	鹿内 俊秀	20th International Symposium of Glycoconjugates (プエルト リコ)	2009年11月 28日 ～12月6日	国 外
◎糖鎖構造の解析データの横断 検索(JCGGDB) (ポスター発表)	鈴木 芳典	産総研ライフサイエンス関 連分野と産技連ライフサイ エンス部会・バイオテクノ ロジー分科会の合同発表会 LS-BT	2010年2月4 日 ～2月5日	国 内

◎病原体が接着する糖鎖と疾患のデータベース (PACDB) (ポスター発表)	鹿内俊秀	産総研ライフサイエンス関連分野と産技連ライフサイエンス部会・バイオテクノロジー分科会の合同発表会 LS-BT	2010年2月4日 ~2月5日	国内
〇ここ10年の糖鎖研究の進歩とそのデータベース化によるさらなる飛躍 (口頭発表)	成松久	学術会議シンポジウム「メタボロミクス研究の最前線とメタボライトデータベースの役割」	2010年1月15日	国内

雑誌等への論文投稿 (該当なし)

特許の申請 (該当なし)

2.4 活動 (運営委員会等の活動等)

完成した成果物をより多くの糖鎖分野の研究者へ広く知って頂くためにできるだけ多くの学会に参加し発表した。そして、完成した成果物の使い方を知って頂くため糖鎖産業フォーラムの勉強会で使い方とDBから得られる有益な情報の取得の仕方などについて説明した。

国内だけではなく、中国やヨーロッパの研究者との交流を行い、日本の糖鎖統合データベースの普及に努めた。

ブース	氏名	発表した場所 (学会等名)	発表した時期	
〇DBの使い方・ニーズを把握するためのブースを設置・ポスターの時間にDBの説明をする公開セッションを展開	成松 久 新間 陽一 鹿内 俊秀 藤田 典昭 協力者： 山田 一作 川崎 敏祐 八杉 悦子	第29回日本糖質学会・飛騨・世界生活文化センター (岐阜県高山市)	2009年9月8日~12日	国内

2.5 実施体制

別表1を参照のこと。

2.6 整備実績一覧

別表2を参照のこと。

別表1 平成21年度に於ける実施体制

研 究 項 目	担当機関等	研究担当者
---------	-------	-------

1. 運営と開発体制の準備	産業技術総合研究所 糖鎖医工学研究センター	◎○成松 久
	産業技術総合研究所 糖鎖医工学研究センター	新聞 陽一
	産業技術総合研究所 糖鎖医工学研究センター	鹿内 俊秀
2. データ提供機関との交渉	産業技術総合研究所 糖鎖医工学研究センター	○新聞 陽一
	産業技術総合研究所 糖鎖医工学研究センター	成松 久
	産業技術総合研究所 糖鎖医工学研究センター	鹿内 俊秀
3. 統合に向けたデータベース構築	産業技術総合研究所 糖鎖医工学研究センター	梶 裕之
	産業技術総合研究所 糖鎖医工学研究センター	亀山 昭彦
	産業技術総合研究所 糖鎖医工学研究センター	成松 久
	産業技術総合研究所 糖鎖医工学研究センター	平林 淳
	産業技術総合研究所 糖鎖医工学研究センター	新聞 陽一

	産業技術総合研究所 糖鎖医工学研究センター	舘野 浩章
	産業技術総合研究所 糖鎖医工学研究センター	梶谷内 晶
	産業技術総合研究所 糖鎖医工学研究センター	佐藤 隆
	産業技術総合研究所 糖鎖医工学研究センター	○鹿内 俊秀
	産業技術総合研究所 糖鎖医工学研究センター	鈴木 芳典
	産業技術総合研究所 糖鎖医工学研究センター	藤田 典昭
	産業技術総合研究所 糖鎖医工学研究センター	平林 淳
	産業技術総合研究所 糖鎖医工学研究センター	伏見 美峰
平成19年度以降、業務推進のために データの統合に関するご意見とご協力 頂いた方々	立命館大学	川崎 敏祐 奥田修二郎
	名古屋市立大学 大学院薬学研究科	加藤 晃一 高橋 禮子 山口 芳樹

研究項目	担当機関等	研究担当者
<p>平成20年度から、業務推進のためにデータの統合に関するご意見とご協力頂いた方々</p> <p>平成21年度から、業務推進のためにデータの統合に関するご意見とご協力頂いた方々</p>	国立医薬品食品衛生研究所	西島 正弘
	東京大学大学院 医学系研究科	八杉 悦子
	名古屋大学 大学院医学系研究科	古川 鋼一
	理化学研究所 フロンティア研究システム（名市大から移籍）	山口 芳樹
	理化学研究所 フロンティア研究システム	谷口 直之 吉田圭一 加藤 雅樹
	野口研究所	白井 孝 水野 真盛 山田 一作
	創価大学 生化学工業株式会社	西原祥子 木下 F. 聖子 萩谷恵里子

注1. ◎：課題代表者、○：サブテーマ代表者

注2. 本業務に携わっている方は、全て記入。

別表2 整備実績一覧

DB名：GGDB

(1) 保有データ情報

(1-1) データの種類

①「生物種	ヒトの糖鎖関連遺伝子。ヒトの遺伝子に類似した他の生物種の遺伝子情報（ホモログ）。H. sapiens：190件（実験データと論文のデータ）、M. musculus：183件 R. norvegicus：172件 C. elegans：63件 A. thaliana：13件 S. cerevisiae：18件 D. melanogaster：74件
②試料・ライブラリー等の種類、数	ヒト試料、ヒト cDNA
③測定方法	リコンビナント酵素の酵素活性測定
④データの内容	リコンビナント糖転移酵素を用いて糖鎖構造を合成した時の結果。ドナー基質・アクセプター・プロダクト情報
⑤その他、特記事項	経産省・NEDO プロジェクト通称「GG project」と「SG Project」の成果物。

(1-2) データソース

①現在のデータ量	95 KB (Oracle)、3.09 MB (eXist) 計 3.18 MB
②データ区分	■自前 ■文献データ
③将来の増加の見込み	新規の遺伝子が発見されるたびに追加
④権利関係	所有者（産業技術総合研究所） 公開（ <input checked="" type="checkbox"/> 可 <input type="checkbox"/> その他 []）
⑤その他、特記事項	http://riodb.ibase.aist.go.jp/rcmg/ggdb/

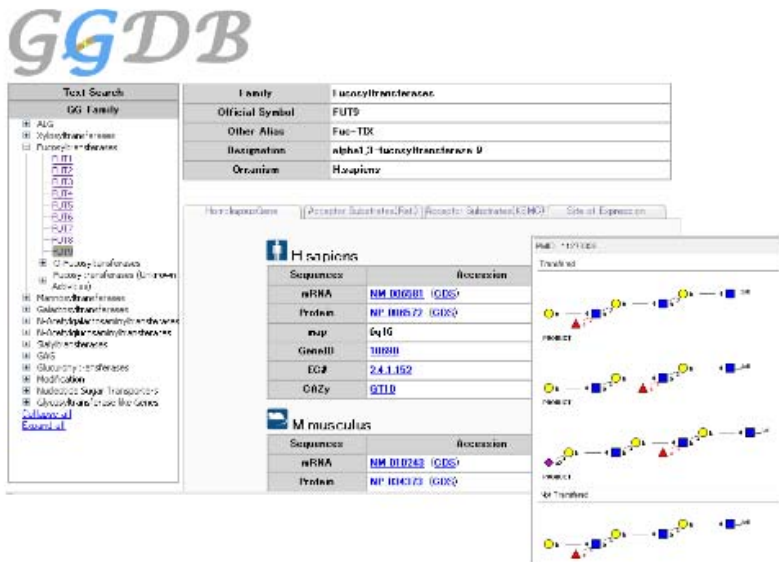
(1-3) データの管理状況

①更新頻度等の管理状況、体制	新規の遺伝子が発見されるたびに追加
②その他、特記事項	

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

①DB管理者数	4名
②キュレータ・	5名

アノテータ数	
③データ構造	XML ファイル、RDB
④DB 管理ソフト	Oracle 10g(10.2.0) 、XML データベース eXist
⑤サーバの OS	Linux
⑥サーバ規模	不明
⑦DB へのアクセス数	平成 20 年度：11,000 アクセス（月平均） 平成 21 年度：321,644 アクセス/年間
⑧独立 IP 数	
⑨その他、特記事項	 <p>ヒトの遺伝子を中心に活性をまとめた。 個々の遺伝子の活性が簡単に分かるように Viewer を開発し整備した。</p>

(2) データ（又はDB）の連結、統合化整備（※試験的、限定的公開済みのものも含む。）

通番	データ（又はDB）の名称 ※URL があれば記述	公開／未公開	概要（データの種類（生物種）・数量（kB等）、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述）
1	http://riodb.ibase.aist.go.jp/rcmg/ggdb/	公開	平成 19 年度に作成し、セキュリティチェックなどを行い平成 20 年に公開した。

			月平均 11000 アクセスがある。 平成21年年度：321,644 アクセス/年間
2	キーワードによる横断検索 http://jcgddb.jp/search/search.cgi?method=3&keyword=&lang=jp	公開	公開日：平成20年8月20日 平成20年8月～3月末日：約1万5千アクセス

(3) DB基盤システム、ツール等開発成果物の整備（※試験的、限定的公開済みのものも含む。）（該当なし）

(4) その他の成果物（(2)、(3)に該当しないもの）（該当なし）

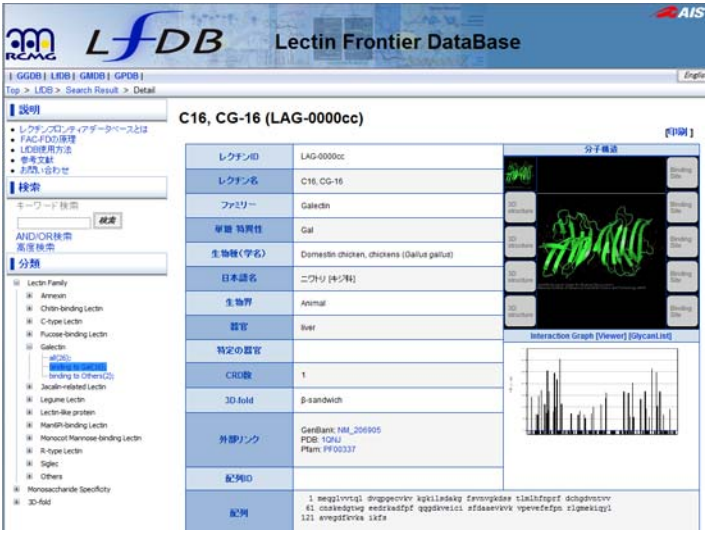
DB名：LfDB

(1) 保有データ情報

※ 貴機関グループ内で保有するデータに関して、以下の内容を記述して下さい。

(1-1) データの種類

①生物種	ヒト～植物まで
②試料・ライブラリー等の種類、数	糖鎖とレクチンの相互作用情報：主に結合活性が安定した植物レクチン
③測定方法	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)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。
④データの内容	レクチンの詳細情報（遺伝子・配列・ファミリー分類・立体構造）、糖鎖とレクチンの相互作用情報 R-type Lectin 18件、Annexin 12件、RicinB 1件、Siglec 11件、Legume 17件、Jacalin-related Lectin16件、Legume Lectin 19件、Chitin-binding 4件、Monocot 1件、Man6Pi-binding Lectin 3件、Fucose-binding Lectin 3件、Galectin 26件、Chitin-binding Lectin 5件、Monocot Mannose-binding Lectin 14件、C-type Lectin 38件、Jacaline-related Lectin 1件、Others 28件 相互作用の件数 86件

⑤ その他、特記事項	
------------	--

(1-2) データソース

現在のデータ量	96.9 MB (Oracle)、2.09 MB (eXist)、計 98.99 MB
② データ区分	<input checked="" type="checkbox"/> 自前 <input checked="" type="checkbox"/> 第三者 <input checked="" type="checkbox"/> 文献データ <input type="checkbox"/> 計算結果等の二次データ <input type="checkbox"/> その他（下欄に詳細を記述） ※複数選択可。二次データのみの保有は不可。
③ 将来の増加の見込み	論文リリース後に公開する。
④ 権利関係	所有者（産業技術総合研究所・レクチン提供機関） 公開（ <input checked="" type="checkbox"/> 可） ※既に公開している場合は URL を「⑤その他、特記事項」に記述 ※権利関係が未解決で、プロジェクト期間内に解決の見込みがある場合は、解決のための手立て等を「⑤その他、特記事項」に詳述
⑤ その他、特記事項	http://riodb.ibase.aist.go.jp/rcmg/glycodb/LectinSearch

(1-3) データの管理状況

① 更新頻度等の管理状況、体制	3ヶ月に1度更新
② その他、特記事項	

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

① DB 管理者数	4名
-----------	----

②キュレータ・アノテータ数	3名
③データ構造	XML ファイル、RDB
④DB 管理ソフト	Oracle 10g(10.2.0) 、XML データベース eXist
⑤サーバの OS	Linux
⑥サーバ規模	1台
⑦DB へのアクセス数	平成20年度：月平均1694アクセス 平成21年度：20265アクセス（年間）
⑧独立 IP 数	1つ
⑨その他、特記事項	※DB への検索メニュー（画面コピーの別紙添付でも可）、使用しているオントロジーがあれば記述

(2) データ（又はDB）の連結、統合化整備（※試験的、限定的公開済みのものも含む。）

通番	データ（又はDB）の名称 ※URL があれば記述	公開／未公開	概要（データの種類（生物種）・数量（kB 等）、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述） ※ 公開している場合は、開始年月、利用状況（平均利用者数、アクセス数、ダウンロード数等の数値的指標で記述） ※ 必要に応じて画面コピー等の図表添付可
	キーワードによる横断検索 http://jcgddb.jp/search/search.cgi?method=3&keyword=&lang=jp	公開	公開日：平成20年8月20日 8月～3月末日：約1万5千アクセス
	Lectin Frontier Database http://riodb.ibase.aist.go.jp/rcmg/glycodb/LectinSearch	公開	公開日：平成20年8月20日 月平均1694アクセス

(3) DB 基盤システム、ツール等開発成果物の整備（※試験的、限定的公開済みのものも含む。）（該当なし）

(4) その他の成果物 ((2)、(3)に該当しないもの) (該当なし)

DB名 : GMDB

(1) 保有データ情報

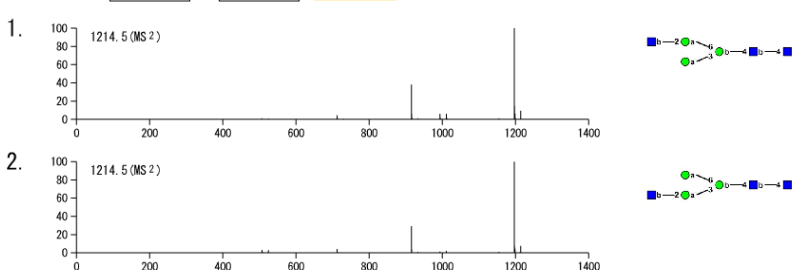
※ 貴機関グループ内で保有するデータに関して、以下の内容を記述して下さい。

(1-1) データの種類

①生物種	ヒト (ヒトと共通する糖鎖構造を持った生物種)
②試料・ライブラリー等の種類、数	ヒト型標準糖鎖
③測定方法	<p>MSⁿ スペクトル測定の手順 この説明は、MALDI-QIT-TOF 型質量分析計 (AXIMA-QIT; 島津製作所) の使用を前提としています。GMDB 内のスペクトルイメージを利用して糖鎖構造を推定する場合には、下記にある CID スペクトル取得手順に沿って MSⁿ スペクトルを測定してください。</p> <p>装置: MALDI-QIT-TOF MS (マトリクス支援レーザー脱離イオン化四重極イオントラップ飛行時間型質量分析計) CID スペクトル測定手順:</p> <ol style="list-style-type: none"> ターゲットプレートにサンプル水溶液の 0.5Lμ を乗せ、乾固します。乾いたサンプルにマトリクス水溶液 0.5L (10μ mg/mL of 2,5-dihydroxybenzoic acid in 30% ethanol) を乗せ、再度、乾固してください。最後に、99.5%エタノールの 0.15Lμ を、ターゲットプレートのマトリクス検体混晶に添加して再結晶化させます。 MS スペクトルを正イオンモードで測定します。 次に、糖鎖の [M + Na]⁺ イオンを前駆イオンとして MS² スペクトルを測定します。この際、前駆イオンのシグナルがほとんど消失する程度に CID エネルギーを調整してください。前駆イオンが完全に無くなるか、もしくは最大ピークの 15% 以上の強度で前駆イオンが残る場合には、そのスペクトルを廃棄し、再度、測定してください。また、測定にあたっては、レーザーショットパターンに従って自動的にレーザー照射位置を変える自動取得機能を使用してください。 MS³ スペクトルの測定についても、上の説明と同様に CID エネルギーを調整してスペクトルを取得します。
④データの	※記録しているデータ項目 (例えば、試料番号、遺伝子名、発現

内容	データ（画像）等
⑤その他、特記事項	経産省・NEDO プロジェクトの成果物

(1-2) データソース

①現在のデータ量	構造数 176 件 スペクトルの枚数 1013 96.9 MB (Oracle)、2.09 MB (eXist)、計 98.99 MB
②データ区分	<input checked="" type="checkbox"/> 自前 <input type="checkbox"/> 第三者 <input type="checkbox"/> 文献データ <input type="checkbox"/> 計算結果等の二次データ <input type="checkbox"/> その他（下欄に詳細を記述） ※複数選択可。二次データのみ の保有は不可。
③将来の増加の見込み	増える可能性はあるが、具体的な数は言えない。
④権利関係	所有者（産業技術総合研究所） 公開（ <input checked="" type="checkbox"/> 可 <input type="checkbox"/> 否 <input type="checkbox"/> その他 [])
⑤その他、特記事項	http://riodb.ibase.aist.go.jp/rcmg/glycodb/Ms_ResultSearch The displayed spectrum is the one based on the peak list. X-axis range from <input type="text" value="168.72972"/> to <input type="text" value="1216.96964"/> Y-axis range from <input type="text" value="0"/> to <input type="text" value="100"/> <input type="button" value="Redraw"/> 

(1-3) データの管理状況

①更新頻度等の管理状況、体制	論文がリリースされた後で公表。
②その他、特記事項	

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

①DB 管理者数	4 名
②キュレータ・アナデータ数	3 名
③データ構造	
④DB 管理ソフト	XMLDB, Oracle

⑤サーバの OS	Linux
⑥サーバ規模	1台
⑦DB へのアクセス数	平成20年度：931 アクセス（月平均） 平成21年度：13174 アクセス（年間）
⑧独立 IP 数	
⑨その他、特記事項	※DB への検索メニュー（画面コピーの別紙添付でも可）、 使用しているオントロジーがあれば記述

(2) データ（又はDB）の連結、統合化整備（※試験的、限定的公開済みのものも含む。）

通番	データ（又はDB）の名称 ※URL があれば記述	公開 / 未公開	概要（データの種類（生物種）・数量（kB 等）、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述） ※ 公開している場合は、開始年月、利用状況（平均利用者数、アクセス数、ダウンロード数等の数値的指標で記述） ※ 必要に応じて画面コピー等の図表添付可
	G M D B （Glycan Mass Spectral Database） http://riodb.ibase.aist.go.jp/rcmg/glycodb/Ms_ResultSearch	公開	公開日：平成20年12月2日 月平均931 アクセス
	構造による横断検索 http://jcggdb.jp/ibase	未公開	糖鎖構造による横断検索から検索できるようにした。 平成21年5月に公開予定。

(3) DB基盤システム、ツール等開発成果物の整備（※試験的、限定的公開済みのものも含む。）（該当なし）

(4) その他の成果物（(2)、(3)に該当しないもの）（該当なし）

DB名：GlycoProtDB

(1) 保有データ情報

※ 貴機関グループ内で保有するデータに関して、以下の内容を記述して下さい。

(1-1) データの種類

①生物種	線虫
②試料・ライブラリー等の種類、数	線虫 (Strain N2)
③測定方法	<p>N結合型糖タンパク質の同定とその糖鎖修飾位置の決定は、液体クロマトグラフィー (LC) と質量分析法 (MS) を連結した LC/MS 法を基礎としたプロテオーム解析技術を用いて行った。簡潔には、</p> <p>(1) 生体材料を 6M の塩酸グアニジンを含む緩衝液中で均質化 (溶解) する。</p> <p>(2) タンパク質のジスルフィド結合を還元し、アルキル化した後に、トリプシンなどのプロテアーゼで分解し、ペプチド混合物を調製する。</p> <p>(3) 得られたペプチド混合物を固定化レクチンカラムに通し、レクチンの特異性によって共通の糖鎖モチーフ (例えばハイマンノース型糖鎖) をもつ一群の糖ペプチドを捕集する。</p> <p>(4) 精製した糖ペプチドを安定同位体標識された水 ($H_2^{18}O$) の中で、N-グリコナーゼ処理し、糖鎖を切除すると同時に、その糖鎖修飾位置を ^{18}O で標識する (この酵素で糖鎖を切除すると、糖鎖が結合していたアスパラギン残基はアスパラギン酸に変換され、このとき溶媒より ^{18}O が取り込まれる)。</p> <p>(5) 標識された (酵素処理前は糖鎖修飾されていた) ペプチドを LC/MS ショットガン法で分析する。</p> <p>(6) MS/MS スペクトル情報を基に、検索エンジンソフトウェアを用いて公共の配列データベースを検索し (MS/MS-イオンサーチ法)、ペプチドがどのタンパク質に由来したかを同定する。</p> <p>我々はこの手法を糖鎖付加位置特異的安定同位体標識法 (Isotope-coded Glycosylation site-specific Tagging: IGOT 法) と名付けた。この一連の分析法では、一度の実験で数百のペプチドが同定できる。</p>
	N結合型糖鎖の修飾位置のデータベース 付加位置の数 1031 件 遺伝子シンボル数 761 件
⑤その他、特記事項	

(1-2) データソース

①「在」のデータ量	96.9 MB (Oracle)、2.09 MB (eXist)、計 98.99 MB
②データ区分	■自前 ■第三者
③将来の増加の見込み	マウスやヒトの組織・患者の組織毎のデータを公開する予定

④権利関係	所有者（産業技術総合研究所・首都大学） 公開（ <input checked="" type="checkbox"/> 可 <input type="checkbox"/> 否 <input type="checkbox"/> その他 [])												
⑤その他、特 記事項	<p>Glycosylation Site</p> <table border="1"> <thead> <tr> <th>Sample</th> <th>Lectin</th> <th>Glycosylation Site</th> </tr> </thead> <tbody> <tr> <td>ORGAN:mixed STRAIN:N2 AGE: Sex: Note:mixed populations Method:IGOT</td> <td>ConA</td> <td></td> </tr> <tr> <td>ORGAN:mixed STRAIN:N2 AGE: Sex: Note:mixed populations Method:IGOT</td> <td>WGA</td> <td></td> </tr> <tr> <td>ALL</td> <td></td> <td></td> </tr> </tbody> </table> <p>http://riodb.ibase.aist.go.jp/rcmg/glycodb/Glc_ResultSearch</p>	Sample	Lectin	Glycosylation Site	ORGAN:mixed STRAIN:N2 AGE: Sex: Note:mixed populations Method:IGOT	ConA		ORGAN:mixed STRAIN:N2 AGE: Sex: Note:mixed populations Method:IGOT	WGA		ALL		
Sample	Lectin	Glycosylation Site											
ORGAN:mixed STRAIN:N2 AGE: Sex: Note:mixed populations Method:IGOT	ConA												
ORGAN:mixed STRAIN:N2 AGE: Sex: Note:mixed populations Method:IGOT	WGA												
ALL													

(1-3) データの管理状況

①更新頻度 等の管理 状況、体制	論文発表後新しい情報を公開。
②その他、特	

記事項	
-----	--

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

①DB 管理者数	4 名
②キュレータ・アナレータ数	2 名
③データ構造	
④DB 管理ソフト	Oracle
⑤サーバの OS	Sun
⑥サーバ規模	1 台
⑦DB へのアクセス数	平成 21 年度 : 14072 アクセス (年間)
⑧独立 IP 数	
⑨その他、特記事項	※DB への検索メニュー (画面コピーの別紙添付でも可)、使用しているオントロジーがあれば記述

(2) データ (又は DB) の連結、統合化整備 (※試験的、限定的公開済みのものも含む。)

通番	データ (又は DB) の名称 ※URL があれば記述	公開 / 未公開	概要 (データの種類 (生物種)・数量 (kB 等)、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述)
1	キーワードによる横断検索 http://jcgddb.jp/search/search.cgi?method=3&keyword=&lang=jp	公開	公開日 : 平成 20 年 8 月 20 日 8 月 ~ 3 月末日 : 約 1 万 5 千アクセス
2	GlycoProtDB (糖タンパク質データベース) http://riodb.ibase.aist.go.jp/rcmg/glycodb/Glc_ResultSearch	公開	公開日 : 平成 20 年 8 月 20 日 月平均 812 アクセス

(3) DB 基盤システム、ツール等開発成果物の整備 (※試験的、限定的公開済みのものも含む。)(該当なし)

(4) その他の成果物 ((2)、(3) に該当しないもの) (該当なし)

DB 名 : GlycoEpitope

(1) 保有データ情報

※ 貴機関グループ内で保有するデータに関して、以下の内容を記述して下さい。

(1-1) データの種類

①生物種	糖鎖に関するエピトープ構造を持っている生物種 (ヒト・ヒツジ・ウシ・ネズミ～ハエ～酵母～植物など)
②試料・ライブラリー等の種類、数	抗糖鎖抗体に関する論文
③測定方法	論文から抽出
④データの内容	抗体が反応する糖鎖構造 (GlycoEpitope)、抗体名、抗体の由来生物種、キャリア情報 (タンパク質・脂質など)、エピトープのレセプター情報、生物種、発現器官・組織、糖鎖の機能についての情報
⑤その他、特記事項	

(1-2) データソース

①現在のデータ量	エピトープ数 : 161 構造 上記エピトープに反応する抗体数 : 512 個
②データ区分	<input type="checkbox"/> 自前 <input checked="" type="checkbox"/> 第三者 <input checked="" type="checkbox"/> 文献データ <input type="checkbox"/> 計算結果等の二次データ <input type="checkbox"/> その他 (下欄に詳細を記述) ※複数選択可。二次データのみの保有は不可。
③将来の増加の見込み	年間新規抗体 100-150 件程度、エピトープ 20-30 件程度
④権利関係	所有者 (立命館大学・川寄敏祐) 公開 (<input checked="" type="checkbox"/> 可 <input type="checkbox"/> 否 <input type="checkbox"/> その他 []) ※既に公開している場合は URL を「⑤その他、特記事項」に記述 ※権利関係が未解決で、プロジェクト期間内に解決の見込みがある場合は、解決のための手立て等を「⑤その他、特記事項」に詳述
⑤その他、特記事項	

(1-3) データの管理状況

①更新頻度等の管理状況、体制	半年に1回程度情報更新
②その他、特記事項	

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

①DB 管理者数	4人
②キュレータ・アナレータ数	8人
③データ構造	RDBMS
④DB 管理ソフト	PostgreSQL-8.2.3
⑤サーバのOS	Linux (Red Hat Enterprise Linux AS release 3)
⑥サーバ規模	NEC Express 5800 (1台)、SGI Altix 450 (1台)
⑦DB へのアクセス数	47775 (H19) 39063 (H20) 平成21年度：(管理者が海外研修中のため不明)
⑧独立 IP 数	不明
⑨その他、特記事項	※DB への検索メニュー (画面コピーの別紙添付でも可)、使用しているオントロジーがあれば記述

(2) データ (又はDB) の連結、統合化整備 (※試験的、限定的公開済みのものも含む。)

通番	データ (又はDB) の名称 ※URL があれば記述	公開 / 未公開	概要 (データの種類 (生物種) ・数量 (kB 等)、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述)
1	キーワードによる横断検索 http://jcgdb.jp/search/search.cgi?method=3&keyword=&lang=jp	公開	公開日：平成20年8月20日 8月～3月末日：約1万5千アクセス

(3) DB 基盤システム、ツール等開発成果物の整備 (※試験的、限定的公開済みのものも含む。) (該当なし)

(4) その他の成果物 ((2)、(3) に該当しないもの) (該当なし)

DB名 : LipidBank

(1) 保有データ情報

(1-1) データの種類

①生物種	生物種問わず論文報告のある脂質構造をデータベース化
②試料・ライブラリー等の種類、数	中でも、糖脂質に関するデータが Glycolipid (696)。内訳として Glycosphingolipid (581)、Glycoglycerolipid and others (115)の構造情報とそれらが記載された論文を明記している。
③測定方法	文献
④データの内容	lipid名 (common name, IUPAC), spectral information (mass, UV, IR, NMR、その他)、literature information
⑤その他、特記事項	

(1-2) データソース

①現在のデータ量	構造数 461 件
②データ区分	<input type="checkbox"/> 自前 <input checked="" type="checkbox"/> 第三者 <input checked="" type="checkbox"/> 文献データ <input type="checkbox"/> 計算結果等の二次データ <input type="checkbox"/> その他 (下欄に詳細を記述) ※複数選択可。二次データのみの保有は不可。
③将来の増加の見込み	なし
④権利関係	所有者 (Japanese Conference on the Biochemistry of Lipids (JCBL)) 公開 (<input checked="" type="checkbox"/> 可 <input type="checkbox"/> 否 <input type="checkbox"/> その他 [])
⑤その他、特記事項	クロスリンクレベルで統合することを条件にプロジェクトへの協力が得られた。 http://lipidbank.jp/

(1-3) データの管理状況

①更新頻度等の管理状況、体制	なし
②その他、特記事項	Version2 ができる可能性あり。

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

①DB 管理者数	1 名
②キュレータ・アナレータ数	1 名

③データ構造	RDBMS
④DB 管理ソフト	Oracle Ver. 8.0.3
⑤サーバの OS	Solaris Ver. 2.6
⑥サーバ規模	Sun Enterprise 450
⑦DB へのアクセス数	不明
⑧独立 IP 数	
⑨その他、特記事項	※DB への検索メニュー（画面コピーの別紙添付でも可）、使用しているオントロジーがあれば記述

(2) データ（又はDB）の連結、統合化整備（※試験的、限定的公開済みのものも含む。）

通番	データ（又はDB）の名称 ※URL があれば記述	公開／未公開	概要（データの種類（生物種）・数量（kB 等）、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述） ※ 公開している場合は、開始年月、利用状況（平均利用者数、アクセス数、ダウンロード数等の数値的指標で記述） ※ 必要に応じて画面コピー等の図表添付可
1	構造による横断検索 http://jcgddb.jp/idb	未公開	糖鎖構造による横断検索から検索できるようにした。 平成21年5月に公開予定。

(3) DB 基盤システム、ツール等開発成果物の整備（※試験的、限定的公開済みのものも含む。）（該当なし）

(4) その他の成果物（(2)、(3)に該当しないもの）（該当なし）

DB 名 : MDHPLCDB

(1) 保有データ情報

(1-1) データの種類

①生物種	ヒト型糖鎖
②試料・ライブラリー等の種類、数	様々な生物種由来のサンプルから抽出されたN結合型糖鎖
③測定方法	糖鎖を酵素的に蛋白質から切り離し、糖鎖の還元末端を 2-アミノピリジンで蛍光ラベルする。この PA 糖鎖を、3種類の分離モードの HPLC カラムを用いて連続的に分離すると同時に、その溶出位置から構造を推定するという仕組み。
④データの内容	グリコシダーゼ消化後の ODS カラムとアミドカラムの溶出位置をデータベース化
⑤その他、特記事項	

(1-2) データソース

①「現在のデータ量	構造数 564 件
②データ区分	<input type="checkbox"/> 自前 <input checked="" type="checkbox"/> 第三者 <input type="checkbox"/> 文献データ <input type="checkbox"/> 計算結果等の二次データ <input type="checkbox"/> その他（下欄に詳細を記述）
③将来の増加の見込み	なし
④権利関係	所有者（名古屋市立大学・加藤晃一・高橋禮子） 公開（ <input checked="" type="checkbox"/> 可 <input type="checkbox"/> 否 <input type="checkbox"/> その他 []）
⑤その他、特記事項	http://www.glycoanalysis.info/galaxy2/index.jsp

(1-3) データの管理状況

①更新頻度等の管理状況、体制	毎月メンテナンス
②その他、特記事項	

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

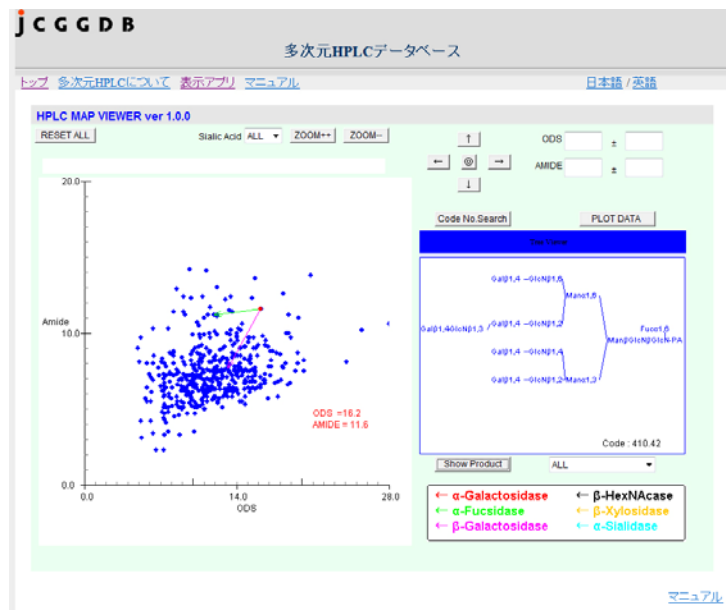
(1-4) データベース関係

①DB 管理者数	1 名
②キュレータ・アナレータ数	2 名
③データ構造	RDBMS, XMLDB
④DB 管理ソフト	MySQL
⑤サーバの OS	Linux サーバ
⑥サーバ規模	1 台
⑦DB へのアクセス数	不明

⑧独立 IP 数

⑨その他、特記事項

下記、画面コピー



をクリックし赤点になったら構造の情報が表示される。構造の下にあるプルダウンメニューから「All」を選択し「show product」ボタンを押すと消化酵素反応後の構造を2種類のカラムで溶出した場合の実験値への移動を矢印で表示される。

詳細

Code No.	410.42
Data access	to Viewer to XML to TEXT
Unit No.	1-2-4-7-9-12-13-14-17-21-24-28-39-53
Structure Linear	-
ods	16.2
amide	11.6

点をダブルクリックすると次のページにある画像・詳細画面が開く。

詳細画面とそれらのデータをダウンロードできる。

(2) データ (又はDB) の連結、統合化整備 (※試験的、限定的公開済みのものも含む。)

通番	データ (又はDB) の	公開 /	概要 (データの種類 (生物種) ・数量 (kB)
----	--------------	------	---------------------------

	名称 ※URLがあれば記述	非公開	等)、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述) ※ 公開している場合は、開始年月、利用状況(平均利用者数、アクセス数、ダウンロード数等の数値的指標で記述) ※ 必要に応じて画面コピー等の図表添付可
1	http://jcgdb.jp/idb		構造による横断検索から複数のDBを検索できるようになった。そのDBの中に下記のDBを検索できるようにした。
2	http://hplc.glycoanalysis.info/	公開	統合DB用にカスタマイズ

(3) DB基盤システム、ツール等開発成果物の整備(※試験的、限定的公開済みのものも含む。)(該当なし)

(4) その他の成果物((2)、(3)に該当しないもの)(該当なし)

DB名: PACDB

(1) 保有データ情報

※ 貴機関グループ内で保有するデータに関して、以下の内容を記述して下さい。

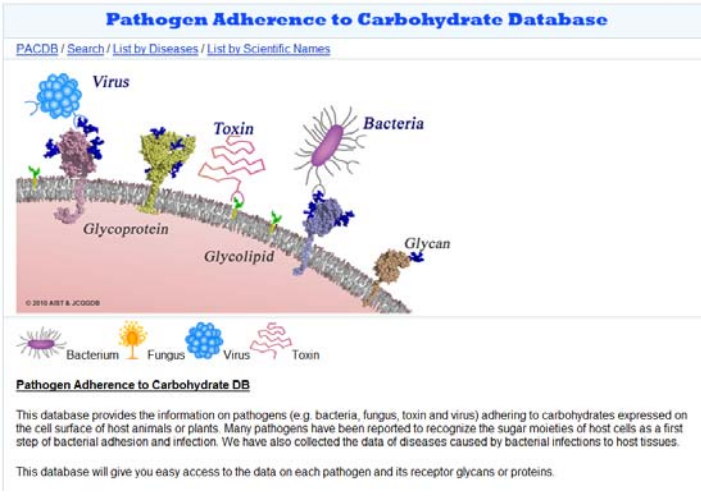
(1-1) データの種類

①生物種	病原体(細菌、真菌、ウイルス、トキシン:129種類) Actinobacillus pleuropneumoniae, Actinomyces naeslundii, Actinomyces odontolyticus, Actinomyces viscosus, Adeno-associated virus, Adenovirus, Anaplasma marginale, Anaplasma phagocytophilum, Bacteriophage, Bordetella parapertussis, Bordetella pertussis, Borrelia afzelii, Borrelia burgdorferi, Borrelia garinii, Borrelia hermsii, Burkholderia cepacia, Burkholderia pseudomallei, Campylobacter jejuni, Campylobacter upsaliensis, Candida albicans, Chlamydia pneumoniae, Chlamydia trachomatis, Clostridium botulinum, Clostridium difficile, Clostridium perfringens, Clostridium tetani, Coronavirus, Cowpox virus, Coxsackievirus, Cryptococcus neoformans, Cytomegalovirus, Diphyllbothrium hottai, D
------	--

	<p>iphylobothrium nihonkaiense, Entamoeba histolytica, Enterococcus faecalis, Equine rhinitis A virus, Escherichia coli, Foot-and-mouth disease virus, Fusobacterium nucleatum, Haemophilus influenzae, Haemophilus parainfluenzae, Hantaan virus, Helicobacter bilis, Helicobacter canis, Helicobacter felis, Helicobacter hepaticus, Helicobacter mustelae, Helicobacter pylori, Hepatitis virus, Herpes simplex virus, Herpes virus, Histoplasma capsulatum, Human immunodeficiency virus, Influenza virus, Klebsiella pneumoniae, Lactobacillus casei, Lactobacillus johnsonii, Lassa virus, Legionella pneumophila, Leptospira interrogans, Listeria innocua, Listeria monocytogenes, Lymphocytic choriomeningitis virus, Moraxella catarrhalis, Mycobacterium avium, Mycoplasma gallisepticum, Mycoplasma pneumoniae, Myxoma virus, Neisseria gonorrhoeae, Neisseria meningitidis, New world arenavirus, Newcastle disease virus, Norovirus, Paracoccidioides brasiliensis, Paramyxovirus, Parvovirus, Pasteurella haemolytica, Picornaviruses, Plasmodium falciparum, Poliovirus, Polyoma virus, Propionibacterium, Propionibacterium freudenreichii, Propionibacterium granulatum, Pseudomonas aeruginosa, Pseudomonas cepacia, Pseudomonas maltophilia, Rabbit fibroma virus, Rabbitpox virus, Rabies virus, Ralstonia solanacearum, Reovirus, Rhinovirus, Rotavirus, Rubulavirus, Saccharomyces cerevisiae, Salmonella enteritidis, Salmonella spp., Salmonella typhimurium, Semliki forest virus, Sendai virus, Serratia marcescens, Shigella dysenteriae, Shigella flexneri, Spirometra erinaceieuropaei, Sporothrix schenckii, Staphylococcus aureus, Staphylococcus epidermidis, Staphylococcus haemolyticus, Staphylococcus saprophyticus, Streptococcus agalactiae, Streptococcus crista, Streptococcus gordonii, Streptococcus mitis, Streptococcus mutans, Streptococcus parvaanguinis, Streptococcus pneumoniae, Streptococcus pyogenes, Streptococcus sanguis, Streptococcus sobrinus, Streptococcus suis, Theiler's murine encephalomyelitis virus, Treponema pallidum, Trypanosoma cruzi, Vaccinia virus, Vibrio alginolyticus, Vibrio cholerae, Vibrio harveyi, Yersinia pestis</p>
②試料・ライブラリー等の種類、数	<p>病原体（細菌、真菌、ウイルス、トキシン：129種類、亜種、亜株により細分化すると351種類） 全体が把握できていない。恐らく数十パーセント程度と推定。</p>
③測定方法	<p>病原体が糖鎖を認識する論文を検索し、文章中や表の中からマニュアルで収集。</p>
④データの内容	<p>論文の Pubmed ID、Taxonomy ID、生物種名、JCGGDB糖鎖番号、結合する・しないの情報、感染によって引き起こされる</p>

	病気の名称、糖鎖を認識する分子名（アドヘシン、レクチンなどの名称）
⑤その他、特記事項	構造のDBとリンク済み。

(1-2) データソース

①現在のデータ量	446 KB (Oracle)
②データ区分	■自前 ■文献データ
③将来の増加の見込み	平成22年度、更に収集する予定。
④権利関係	所有者（産業技術総合研究所） 公開（ <input checked="" type="checkbox"/> 可 <input type="checkbox"/> 否 <input type="checkbox"/> その他 [])
⑤その他、特記事項	<p>http://jcggdb.jp/search/PACDB.cgi</p>  <p>The screenshot shows the PACDB website interface. At the top, it says 'Pathogen Adherence to Carbohydrate Database'. Below that are navigation links: 'PACDB / Search / List by Diseases / List by Scientific Names'. The main content is a diagram illustrating pathogen adhesion. It shows a cross-section of a cell membrane with glycoproteins and glycolipids. A virus is shown binding to a glycoprotein, a bacterium to a glycolipid, and a toxin to a glycan. A legend at the bottom identifies the pathogens: Bacterium (rod-shaped), Fungus (spore-like), Virus (spherical), and Toxin (chain-like). Below the legend, there is a short paragraph explaining that the database provides information on pathogens (bacteria, fungus, toxin, and virus) adhering to carbohydrates on host cell surfaces.</p>

(1-3) データの管理状況

①更新頻度等の管理状況、体制	年に2度程度更新
②その他、特記事項	

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

①DB 管理者数	3 名
②キュレータ・アノテータ数	3 名
③データ構造	糖鎖構造：XML ファイル、その他、RDBMS
④DB 管理ソフト	Oracle 10g(10.2.0)、XML データベース eXist
⑤サーバの OS	Solaris
⑥サーバ規模	1 台
⑦DB へのアクセス数	12698 アクセス[2010-01-19(公開) ~ 2010-03-31]
⑧独立 IP 数	426
⑨その他、特記事項	

(2) データ (又は DB) の連結、統合化整備 (※試験的、限定的公開済みのものも含む。)

通番	データ (又は DB) の名称 ※URL があれば記述	公開 / 未公開	概要 (データの種類 (生物種)・数量 (kB 等)、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述)
1	糖鎖構造による統合検索	公開	統合検索から PACDB を検索できる。 公開日：2010 年 1 月 19 日 平成 21 年度：12698 アクセス

(3) DB 基盤システム、ツール等開発成果物の整備 (※試験的、限定的公開済みのものも含む。)(該当なし)

(4) その他の成果物 ((2)、(3) に該当しないもの) (該当なし)

DB 名：糖鎖関連疾患とその原因遺伝子のデータベース

(1) 保有データ情報

※ 貴機関グループ内で保有するデータに関して、以下の内容を記述して下さい。

(1-1) データの種類

①生物種	ヒト～植物まで
②試料・ライ	ヒトの疾患 (80 種類) を収集。

②その他、特 記事項	
---------------	--

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

①DB 管理者数	2 名
②キュレータ・ アノテータ数	2 名
③データ構造	
④DB 管理ソフト	Oracle 10g (10.2.0)
⑤サーバの OS	Linux
⑥サーバ規模	1 台
⑦DB へのアクセ ス数	
⑧独立 IP 数	
⑨その他、特記 事項	

(2) データ (又はDB) の連結、統合化整備 (※試験的、限定的公開済みのものも含む。)

通 番	データ (又はDB) の名称 ※URL があれば記述	公 開 / 未公開	概要 (データの種類 (生物種)・数量 (kB 等)、 本プロジェクトで実施した特徴点、進捗状況、 今後の計画・課題などを簡潔にわかりやすく 記述)
1	GGDB やグリコシダー ゼのDBと連携	未公開	平成22年度実現の予定

(3) DB 基盤システム、ツール等開発成果物の整備 (※試験的、限定的公開済みのものも含む。) (該当なし)

(4) その他の成果物 ((2)、(3) に該当しないもの) (該当なし)

DB 名 : JCGGDB Reports

(1) 保有データ情報

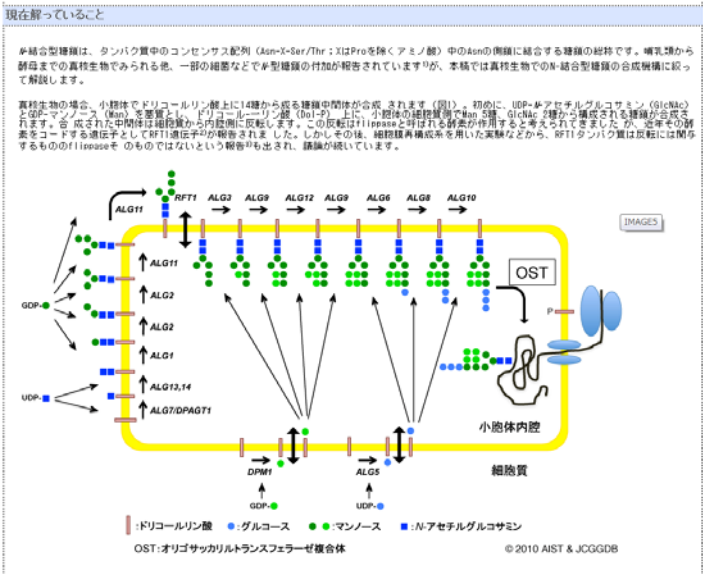
※ 貴機関グループ内で保有するデータに関して、以下の内容を記述して下さい。

(1-1) データの種類

①生物種	パスウェイは哺乳類を中心に。ツールとしての技術は全生物。
②試料・ライ	糖鎖の生合成や分解に関する記事を中心に3題 (英語1題、和文

ブラリー等の種類、数	3題)
③測定方法	原稿依頼し執筆。仕上がった原稿はDBに登録して公開。
④データの内容	記事、パスウェイの経路、遺伝子の情報。
⑤その他、特記事項	記事に出現する遺伝子名や糖鎖構造名と JCGGDB の情報がリンクする

(1-2) データソース

①現在のデータ量	27 KB (Oracle)、1 KB (eXist)
②データ区分	■自前 □第三者 ■文献データ
③将来の増加の見込み	論文で報告リリース後に公開する。平成22年度末日までに20~30報用意する予定。
④権利関係	所有者(産業技術総合研究所) 公開(■可 □否 □その他[])
⑤その他、特記事項	<p>http://jcgddb.jp/doc/</p>  <p>現在解っていること</p> <p>N-結合型糖鎖は、タンパク質中のコンセンサス配列 (Asn-X-Ser/Thr; XはProを除くアミノ酸) 中のAsnの側鎖に結合する糖鎖の総称です。哺乳類から酵母までの真核生物でみられる他、一部の細菌などでN-糖鎖の付加が報告されていますが、本稿では真核生物でのN-結合型糖鎖の合成機構に絞って解説します。</p> <p>真核生物の場合、小胞体でドリコルリン酸上に14糖から成る糖鎖中間体が合成されます(図)。最初に、UDP-N-アセチルグルコサミン (GlcNAc) とGDP-マンノース (Man) を基質とし、ドリコルリン酸 (Dol-P) 上に、小胞体の細胞質側でMan 5糖、GlcNAc 2糖から構成される糖鎖が合成されます。合成された中間体は細胞質から内腔側に反応します。この反応はflippaseと呼ばれる酵素が利用すると考えられてきました。近年その酵素をコードする遺伝子としてRFT1遺伝子が報告されました。しかしその後、細胞質再構成系を用いた実験などから、RFT1タンパク質は反応には関与するもののflippaseそのものではないという報告も出され、議論が続いています。</p> <p>LEGEND: ■ ドリコルリン酸 ● グルコース ● マンノース ■ N-アセチルグルコサミン OST:オリゴサッカリルトランスフェラーゼ複合体 © 2010 AIST & JCGGDB</p> <p>内腔側では、さらに糖転移酵素が作用し、Dol-P-ManやDol-P-グルコース (Glc) を基質としてGlc 3糖、Man 9糖、GlcNAc 2糖から構成される糖鎖中</p>

(1-3) データの管理状況

①更新頻度等の管理状況、体制	分野ごとに担当者を決め、その領域で新しい知見が論文で報告されたときにその都度更新する。平成22年度は20題を目標。
----------------	---

②その他、特記事項	
-----------	--

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

①DB 管理者数	2 名
②キュレータ・アノテータ数	1 名
③データ構造	RDBMS
④DB 管理ソフト	Oracle 10g(10.2.0)
⑤サーバの OS	Solaris 10
⑥サーバ規模	1 台
⑦DB へのアクセス数	不明
⑧独立 IP 数	
⑨その他、特記事項	

(2) データ (又は DB) の連結、統合化整備 (※試験的、限定的公開済みのものも含む。)

通番	データ (又は DB) の名称 ※URL があれば記述	公開 / 未公開	概要 (データの種類 (生物種)・数量 (kB 等)、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述)
	なし		

(3) DB 基盤システム、ツール等開発成果物の整備 (※試験的、限定的公開済みのものも含む。)(該当なし)

(4) その他の成果物 ((2)、(3) に該当しないもの) (該当なし)

DB 名 : JCGGDB GlycoScience Protocol Online Database

(1) 保有データ情報

※ 貴機関グループ内で保有するデータに関して、以下の内容を記述して下さい。

(1-1) データの種類

①生物種	こだわらない
②試料・ライ	平成 21 年度 : 44 プロトコルを整備。

ブラリー等の種類、数	
③測定方法	原稿依頼し執筆。仕上がった原稿はDBに登録して公開。
④データの 内容	器具、機器、手順、ノウハウ、Q&A、注意事項やコメント
⑤その他、特 記事項	プロトコルを評価できる機能を追加。

(1-2) データソース

①現在の データ量	登録の途中のため、データ量は不明										
②データ 区分	■自前										
③将来の 増加の 見込み	平成22年度末日までに100プロトコルを用意する予定。										
④権利関 係	所有者（産業技術総合研究所・立命館大学） 公開（ <input type="checkbox"/> 可 <input checked="" type="checkbox"/> 否 <input checked="" type="checkbox"/> その他〔著作権を産総研に移行する予定・平成22年度8月までに公開予定〕）										
⑤その他、 特記事項	 <p>The screenshot shows the JCGGDB Protocol Database interface. It features a search bar for 'Glycan Analysis by MSⁿ Spectral Matching' and a dropdown menu for the registrant, currently set to 'Dr. JCGGDB ADMIN (Ver 1.0)'. Below the search bar are tabs for 'Introduction', 'Protocol', and 'References'. The 'Protocol' tab is active, displaying a table with the following details:</p> <table border="1"> <tr> <td>Category</td> <td>Isolation & structural analysis of glycan</td> </tr> <tr> <td>Protocol name</td> <td>Glycan Analysis by MSⁿ Spectral Matching</td> </tr> <tr> <td>Keywords</td> <td>tandem mass spectrometry MSⁿ spectral database collision induced dissociation MALDI</td> </tr> <tr> <td>Reagents</td> <td> <ul style="list-style-type: none"> 2,5-Dihydroxybenzoic Acid, proteomics grade (Wako Pure Chemical, Osaka, Japan) Water, Ultrapure for LC/MS (Wako Pure Chemical, Osaka, Japan) Ethanol, for HPLC (Wako Pure Chemical, Osaka, Japan) </td> </tr> <tr> <td>Instruments</td> <td> <ul style="list-style-type: none"> MALDI-QIT-TOF mass spectrometer equipped with an Ultra Cooling Kit (AXIMA-QIT, Shimadzu, Kyoto, Japan) GMDB (URL: http://riodb.dv.tbase.aist.go.jp/rcmg/glycodb/MS_ResultSearch) </td> </tr> </table> <p>The 'Protocol' section includes a numbered list of steps for MSⁿ Spectra Acquisition, with links to 'Comment' for each step.</p>	Category	Isolation & structural analysis of glycan	Protocol name	Glycan Analysis by MS ⁿ Spectral Matching	Keywords	tandem mass spectrometry MS ⁿ spectral database collision induced dissociation MALDI	Reagents	<ul style="list-style-type: none"> 2,5-Dihydroxybenzoic Acid, proteomics grade (Wako Pure Chemical, Osaka, Japan) Water, Ultrapure for LC/MS (Wako Pure Chemical, Osaka, Japan) Ethanol, for HPLC (Wako Pure Chemical, Osaka, Japan) 	Instruments	<ul style="list-style-type: none"> MALDI-QIT-TOF mass spectrometer equipped with an Ultra Cooling Kit (AXIMA-QIT, Shimadzu, Kyoto, Japan) GMDB (URL: http://riodb.dv.tbase.aist.go.jp/rcmg/glycodb/MS_ResultSearch)
Category	Isolation & structural analysis of glycan										
Protocol name	Glycan Analysis by MS ⁿ Spectral Matching										
Keywords	tandem mass spectrometry MS ⁿ spectral database collision induced dissociation MALDI										
Reagents	<ul style="list-style-type: none"> 2,5-Dihydroxybenzoic Acid, proteomics grade (Wako Pure Chemical, Osaka, Japan) Water, Ultrapure for LC/MS (Wako Pure Chemical, Osaka, Japan) Ethanol, for HPLC (Wako Pure Chemical, Osaka, Japan) 										
Instruments	<ul style="list-style-type: none"> MALDI-QIT-TOF mass spectrometer equipped with an Ultra Cooling Kit (AXIMA-QIT, Shimadzu, Kyoto, Japan) GMDB (URL: http://riodb.dv.tbase.aist.go.jp/rcmg/glycodb/MS_ResultSearch) 										

(1-3) データの管理状況

①更新頻度 等の管理 状況、体制	ユーザ登録した人は誰でもプロトコルを登録できるようにする。
------------------------	-------------------------------

②その他、特記事項	
-----------	--

※ 既にデータベースを保有している場合は、以下についても記述して下さい。

(1-4) データベース関係

①DB 管理者数	2 名
②キュレータ・アノテータ数	1 名
③データ構造	RDBMS
④DB 管理ソフト	Oracle 10g(10.2.0)
⑤サーバの OS	Solaris 10
⑥サーバ規模	1 台
⑦DB へのアクセス数	
⑧独立 IP 数	
⑨その他、特記事項	

(2) データ (又は DB) の連結、統合化整備 (※試験的、限定的公開済みのものも含む。)

通番	データ (又は DB) の名称 ※URL があれば記述	公開 / 未公開	概要 (データの種類 (生物種)・数量 (kB 等)、本プロジェクトで実施した特徴点、進捗状況、今後の計画・課題などを簡潔にわかりやすく記述) ※ 公開している場合は、開始年月、利用状況 (平均利用者数、アクセス数、ダウンロード数等の数値的指標で記述) ※ 必要に応じて画面コピー等の図表添付可
	なし		

(3) DB 基盤システム、ツール等開発成果物の整備 (※試験的、限定的公開済みのものも含む。) (該当なし)

(4) その他の成果物 ((2)、(3)に該当しないもの) (該当なし)