本件の取り扱いについては、資料配布と同時解禁
（2009 年 3 月 31 日（火）午後 2 時）でお願いします。

2009 年 3 月 31 日
独立行政法人理化学研究所

理研のデータベース構築基盤の公開基準をセマンティックウェブに統一

－ ライフサイエンスネットワーキングシステム（理研サイネス）の運用を開始 －

本研究成果のポイント

〇各分野で活躍する日本の研究者が中核となる国際連携を推進する情報基盤へ
〇大量データを扱うライフサイエンス分野の統合データベース事業でも有用性を発揮
〇個々のデータベースを丸ごと研究成果物として発表できる学術メディアとして期待

独立行政法人理化学研究所（野依良治理事長）は、ライフサイエンスを主体にしたデータベースの構築基盤システムを理研内で一元化し、国際標準規格「セマンティックウェブ形式※1」に準拠したデータ公開を大規模に実施するための共通基盤「理研サイネス」を構築しました。これは、理研生命情報基盤研究部門（理研 BASE、豊田哲郎部門長）による研究成果です。

近年のライフサイエンスが、大量のデータを扱う科学に発展したことなどから、研究成果を論文形式にして発表するだけでなく、ウェブ上でアクセス可能なデータベースとして独自に発信していく機会が増えました。しかし、研究論文の発表では学術雑誌などの専用のメディアが発達しているのに対し、データベースの発信・発表では、個々の研究者が自らウェブサイトを立ち上げてサービスしなければならず、特に、発信・発表後も継続的にサービスを維持するための運用コストが、研究者にとって大きな負担となっていました。さらに、個々の研究者が立ち上げた独自サイトが、多数乱立するにつれ、公開方法がまちまちで国際基準規格に準拠していないサイトも多くなり、利用者から見ても分かりづらく、統合的な活用が妨げられていました。

そこで理研 BASE では、研究者がウェブサーバーを維持する必要がなく、個々のデータベースを丸ごと研究成果物として発信・発表することができる共通基盤「理研サイネス」を開発しました。理研サイネスは、研究者自身が、サイバースペース上でバーチャルな研究プロジェクトを組織することを支援し、数万個以上の研究プロジェクト群の収容を想定しています。この新たに開発したデータベースの構築基盤システムは、各研究プロジェクトを機密性高く区切り、未公開情報の管理や大規模なデータを介した研究業務フローを、プロジェクトごとに柔軟に設定することができることから、研究プロジェクト内のデータガバナンス※2 を向上させ、新しいタイプの学術メディアとしても期待できます。この基盤で構築したデータベース群は、国際標準規格を採用し、公開が容易なため、「理研総合データベース（RIKEN Hub-Database）」として理研が集約的に維持管理を継続し、各分野の日本の研究者が国際連携研究の核として主導権を発揮するチャンスを提供します。

本研究は、理研内の戦略的裁量研究費所内連携推進事業として行ったものです。理研サイネスは、閲覧機能に限り、Firefox というウェブブラウザに対応した試用版で、理研総合データベースとともに、3 月 31 日から公開しました（http://database.riken.jp/）。

1. 背 景

近年のインターネットの発達によって、研究成果データを発信するとともに学術的な発表を行うメディアとして、ウェブが活用されるようになりました。最近では、個々のデータベースをウェブ上で公開すること自体が、成果発表の行為として頻繁に行われています。しかし、研究論文の発表では、学術雑誌という専用のメディアが発達しているのに対し、データベースの発表では、個々の研究者が自らウェブサイトを立ち上げてサービスしなければならず、発表後も継続的にサービスを維持するための運用コストが研究者にとって大きな負担となっていました。また、個々の研究者が立ち上げた独自サイトが、多数乱立するにつれ、公開方法が国際基準規格に準拠していないサイトも多くなり、利用者側から見てどのサイトで何が公開されているか分かりづらく、統合的な活用が妨げられていました。

従来、ウェブ上で大勢が共同でコンテンツを作成する場合は、Wiki※3 が一般的に用いられていました。Wiki は、百科事典のように人間が直接読んで解釈するコンテンツを大勢で書き込む場合には優れています。しかし、ライフサイエンス研究分野では、そのコンテンツを膨大な実験データとコンピュータ上で自動照合して、実験データの解釈に利用する必要があり、そのようなコンテンツのデータベースを Wiki だけで構築することは困難な状況となっていました。このため、各研究者が、データベース構築から公開化までの活動を、長期にわたりセキュリティの高い状態で一括的に管理・運営することができる「データベース構築のためのインキュベーション基盤」が強く求められるようになってきました。また、世界的にバイオインフォマティクス研究者が不足している状況から、多種類のコンテンツをコンピュータが自動的に統合化処理できるように、形式や語彙などの基準をそろえてコンテンツ作成と品質管理が可能となる機能を持った新しいデータベース構築基盤システムも必要とされていました。

これまでの日本では、優れた情報基盤を国際的に提供することの意義が十分認識されておらず、日本の研究者が大きく貢献をした研究成果データベースであっても、欧米の研究者がいち早く提供した情報基盤に吸収される傾向にありました。このため、国際的なデータ連携型研究プロジェクトで、日本の研究者が核となって主導権を発揮できる優れたデータベース構築基盤システムを整備することが急務となっていました。

2. 研究手法と成果

理研 BASE は、研究者がウェブサーバーを維持する必要がなく、個々のデータベースを丸ごと研究成果物として発信・発表できる共通基盤「理研サイネス」を開発しました（図 1）。理研サイネスは、以下の特徴を備えています。

①数万個以上の個別データベース構築活動を、大勢の研究者がインターネット経由で並行して実施する。
②大規模なデータを介した業務フローを柔軟に設定でき、人的連携や自動処理を容易化する。
③各活動群をセキュリティの高い状態で区切り、未公開の状態でデータベース構築ができる。
④構築したデータベースをその基盤から直接公開できる。
⑤公開後も、研究者がシステムの維持コストを負担することなく、その基盤でコンテンツを継続的に更新することができる。
⑥複数の世界標準形式に準拠したデータ配信が容易。

理研サイネスは、研究者自身がサイバースペース上でバーチャルな研究プロジェクトを組織

できるよう支援し、数万個以上の研究プロジェクト群を収容することを想定して開発したデータベース構築基盤システムです（図 2）。各プロジェクトをセキュリティの高い状態で区切り、未公開情報の管理や大規模なデータを介した研究業務フローを研究プロジェクトごとに柔軟に設定できるように、国際標準規格（セマンティックウェブ形式）にセキュリティ管理を施した独自の技術を開発しました。これによって、従来の書き込み型の Wiki 機能に加え、形式や語彙などの基準をそろえてコンテンツを作成し、コンピュータが自動的にデータを統合化処理させることも可能にしました。理研サイネスでは、分かりやすいインタフェースでセマンティックウェブ形式のデータ作成が行えるため、大勢の研究者が連携して効率的に多目的な活用が可能な研究用データを作成することできます。セマンティックウェブ形式で統一的に作成されたデータは、各技術分野に特有のデータフォーマットへも容易に自動変換できるため、多様な世界標準形式に準拠したデータ配信が可能です。そのため、個々の研究者がデータベースを発表するための、新しいタイプの学術メディアとしても利用できます。この基盤で公開したデータベース群は、「理研総合データベース」として理研 BASE が集約し、維持管理していきます。

これまで日本は、国際連携において情報基盤を提供することに遅れ、欧米の研究者がいち早く提供した情報基盤に日本のデータが吸い込まれていく傾向となっていました。しかし、理研サイネスは、国際的なデータ連携型プロジェクトにおける情報連携基盤として多目的に利用できるため、日本の研究者に国際連携研究の核として主導権を発揮できるチャンスをもたらします（図 3）。

3. 今後の期待

理研では、文部科学省の委託研究開発事業である統合データベースプロジェクト※4 を理研サイネスの基盤で実施しており、植物オミックスデータベースおよびタンパク質結晶化実験データベース（約 1,000 万以上のデータ）を、国際標準規格であるセマンティックウェブ形式で理研サイネスから公開します。さらに今後、データダウンロードを可能にし、研究成果データの共有化を進める予定です。

セマンティックウェブでは、各データの意味や関係性を、コンピュータが自動解釈可能な形式で表現するため、これらの膨大なデータベースの単なる閲覧だけでなく、コンピュータでさまざまな角度から自動解析を試みることができるようになると期待されます。これにより、ウェブサイエンス※5 の成果をライフサイエンスに活かしていくことができます。

また、理研サイネスは、遠隔地に分散したライフサイエンス研究者たちを結びつける情報基盤として、今後はさらに、トランスレーショナルリサーチの分野における情報連携基盤としての利用に期待が高まります。現在、理研免疫・アレルギー科学総合研究センターと連携して、アジアにおける原発性免疫不全症の専門医と臨床医をつなぐネットワーク構築に理研サイネスの基盤を活用する計画で、今後も国際的な研究プロジェクトを支援していく予定です。

＜報道担当・問い合わせ先＞
（問い合わせ先）
独立行政法人理化学研究所
生命情報基盤研究部門
部門長　豊田　哲郎（とよだ　てつろう）

＜補足説明＞

※1 セマンティックウェブ形式

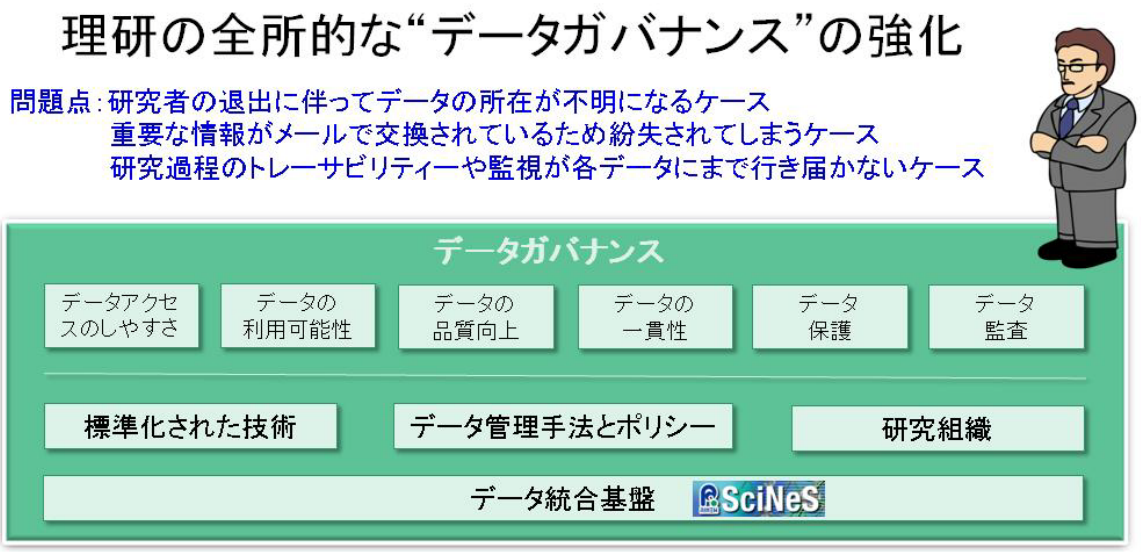
ワールドワイドウェブ（WWW）の発展形として、英国の計算機科学者であるティム・バーナーズ・リーによって提唱されているウェブ技術。ワールドワイドウェブは、ネットワーク上に置かれた文書などのリソース間をハイパーリンクでつなぐもので、インターネット上の標準的なインフラとして爆発的な成功を収めた。しかし、ハイパーリンクは、人間がそのリンクをたどりながら読み進めていくのには適しているものの、単純に 2 つのリソースを結びつけているだけなので、そのリンクがどのような関係づけを意味しているかは表現していない。リンクの意味については、テキストに書かれた内容を人間が読んで解釈するしかなく、コンピュータが意味を認識し、高度な知識処理を行うための情報をほとんど含んでいないことが、ワールドワイドウェブの問題点として指摘された。この反省から、ウェブにセマンティクス（意味論）を与えることが求められるようになった。つまり、情報を持つ文書を機械可読な形で提供できるようにし、また、リンクにその関係を示す値を付け加えられるようにすることで、私たちが自ら読む以上の情報を、コンピュータの力を借りて取り出すことを目指すのがセマンティックウェブである。

参考資料：セマンティックウェブのための RDF/OWL 入門（神崎正英著、森北出版株式会社、2005 年）
「コンピュータの中の脳　-情報基盤の進化論-」（豊田哲郎）生体の科学 59(1):20-32, 2008
（http://omicspace.riken.jp/publications/evolution/page7.html）

※2 データガバナンス

データの生成、管理、活用にかかわる個々のプロジェクトを研究所規模で統制することにより、一貫性のある適切なデータ管理を実現させるもの。データガバナンスはこの目的を実現させるための人、研究プロセス、情報技術を包含する枠組みで、以下に掲げる課題が含まれる。

- データへのアクセスのしやすさ、利用可能性、データ品質の向上
- データ処理結果の一貫性と信頼性の向上
- 個々の研究室におけるデータ損失のリスクを軽減
- 高度なデータ保護の実現
- データが持つ研究活動を加速させる能力を最大限に引き出す手法の確立
- データ品質向上ための説明責任の明示



※3 Wiki

サーバ上のウェブ文書を書き換えるシステムの一種。ネットワークにつながっていれば、どこからでも、修正・更新が行えるため、多人数の共同作業で書き換えをすることに向いているのが特徴。

※4 統合データベースプロジェクト

文部科学省が推進するライフサイエンスデータベースの統合化事業。日本のライフサイエンス関係データベース整備戦略の立案・評価支援、データベース統合化の基盤技術開発、ポータルサイトの整備などを行い、統合化を推進している。第 3 期「科学技術基本計画」（平成 18 年 3 月 28 日閣議決定）に基づき、総合科学技術会議が策定したライフサイエンス分野の推進戦略である、統合データベース整備事業である。情報・システム研究機構ライフサイエンス統合データベースセンターを中心に、8 つの研究参画機関と 4 つの補完課題実施機関、3 つの分担機関から構成される。理研は平成 19 年度から 4 年計画で行われている「統合データベースプロジェクト補完課題」に参画し、植物オミックス情報およびタンパク質構造情報を同プロジェクトに提供する。

参考資料：文科省統合データベースプロジェクト
ライフサイエンス分野の統合データベース整備事業

※5 ウェブサイエンス

ウェブ世界を探求する新たな学問領域として、2006 年 11 月に英国の計算機科学者のティム・バーナーズ・リーらが立ち上げた。ウェブの驚異的な成長を可能にしている構成原理を解明すること、インターネット上の人間関係がどのように進展し、社会習慣をどう変えるかを明らかにすることを目指している。ウェブ世界はライフサイエンスの研究スタイルだけでなく、遺伝子進化にも大きな影響を与えると予想されている。

参考資料：「ウェブサイエンスの誕生」　日経サイエンス 2009 年 1 月号
「コンピュータの中の脳　-情報基盤の進化論-」生体の科学　59(1):20-32, 2008
http://database.riken.jp/item/cria42s2-ria42s2
「ゲノム解読から生命戦略の解明へ」Bionics 26-30, Feb., 2007
http://database.riken.jp/item/cria42s2-ria42s1

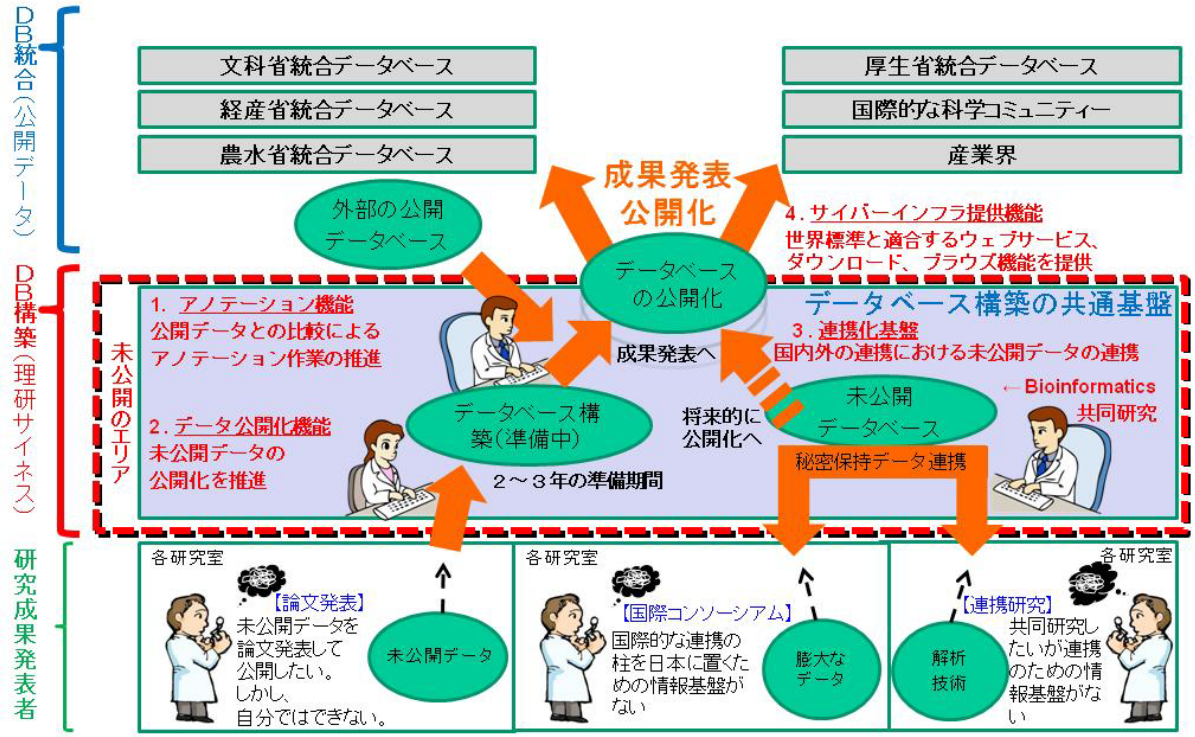


図 1　理研 BASE が提供する大規模データベース構築基盤（理研サイネス）

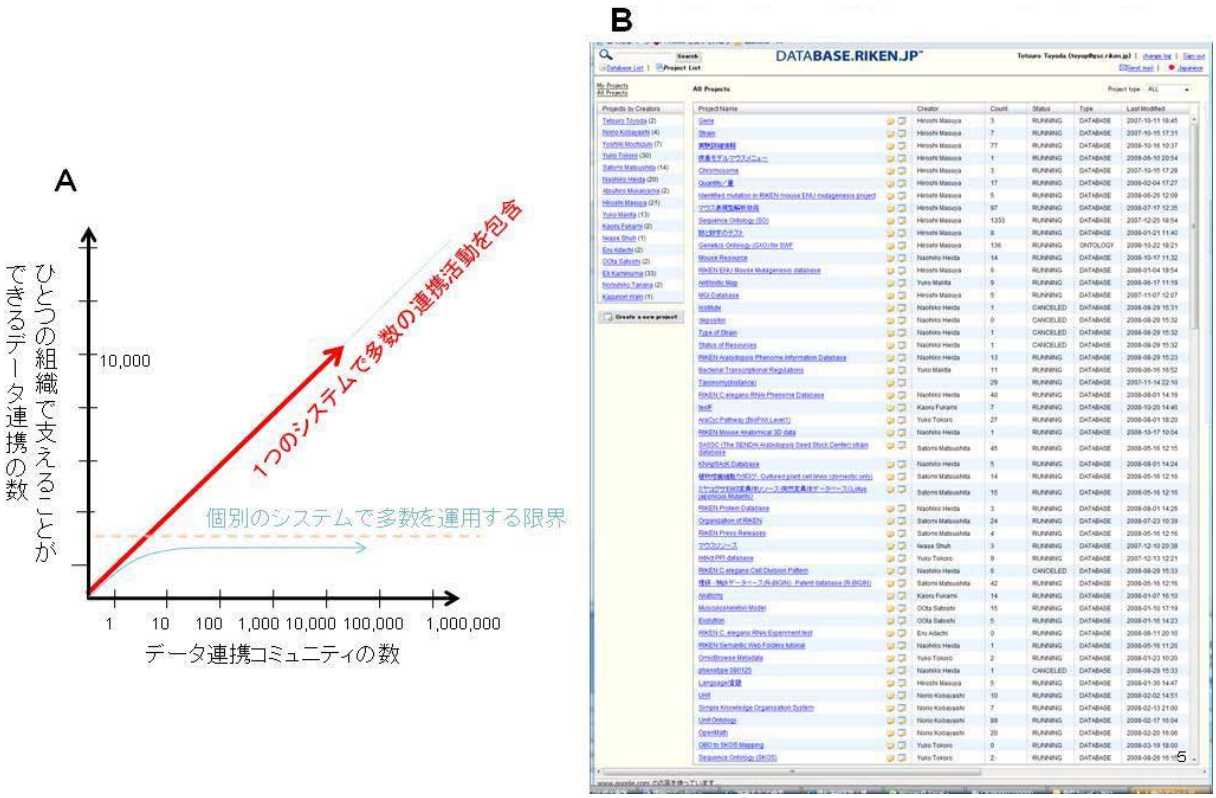


図 2　理研サイネスの中では、多数のデータベース構築連携が進行中

A. 理研サイネス（赤矢印）では、セマンティックウェブにセキュリティ管理を施した独自の技術を採用したため、一つのシステムで数千の異なるデータベースを、それぞれの研究者に対してセキュリティの高い状態で提供できる。

B. 理研サイネスで進行中の連携プロジェクト一覧

図 3　世界の研究者の英知と最新テクノロジーを連携させるイメージ図

（参考）**理研サイネスの詳細**

● 大規模データを介した連携研究のための情報基盤

・大規模データを介した共同研究のための理研サイネス

“ライフサイエンスネットワーキングシステム”【3】

・連携化支援機能【4】

・コラボレーション・レビュー機能　【5】

・メッセージ機能【6】

・電子ラボノート機能【7】

・コミュニティ機能【8】

・プロジェクト運営機能【9】

・オントロジープロジェクト【10】

・オントロジー構築機能【11】

・異種オントロジー間対応付け支援機能【12】

・データベースプロジェクト【13】

・データベース構築機能【14】

・データベース公開機能【15】

・文書タグ付けルール設定機能【16】

・自動文書タグ付け機能【17】

・Wiki 機能【18】

・データベーステンプレートプロジェクト【19】

・データベースウィザード機能【20】

・リポジトリ機能【21】

・ライブデータベースリポジトリ機能【22】

・解析ツール拡張機能【23】

・データ自動解析機能【24】

・データ可視化機能【25】

・一括ダウンロード機能【26】

・データマート機能【27】

・データフロー制御機能【28】

・ラボラトリーオートメーション【29】

・自動更新機能【30】

・自動セマンティックウェブ変換機能【31】

・データベース自動統合更新機能【32】

・推論検索機能【33】

・推論検索エンジン GRASE（PosMed, etc.）【34】

・階層構造一括検索機能【35】

・塩基配列または化合物構造のようなキーワード以外の検索【36】

・安定した運用【37】

・スーパーコンピュータとの連携【38】

・さまざまなサービスの機能【39】

・災害対策【40】

・システムメンテナンスとサポートスタッフ【41】

近年のライフサイエンスが大量のデータを扱う科学へと急速に進展したことから、データ解析を専門に行うバイオインフォマティクスの人材が世界的に不足しています。理化学研究所生命情報基盤研究部門（理研 BASE）では、バイオインフォマティクスの専門家を集めた研究組織として、幅広い分野の研究者と直接連携しながらデータ解析研究【1】を行ってきましたが、さらに多くの研究者どうしが効果的に連携しあえる情報基盤を提供することで、間接的にもバイオインフォマティクスの連携研究を強化する必要があると考え、大小さまざまな多数の研究連携群が大規模なデータベースを介して連携研究を行える情報システム「理研サイネス」を開発しました。

理研サイネスが提供する連携化支援機能【4】では、セキュアに区切られたサイバースペース内で未公開レベルのデータを扱える環境を個々の研究者に提供できるだけでなく、そこで編纂したデータの一部、または全部をデータベースとしてそのままインターネットへ公開化し、研究成果のパブリケーションメディアとしての役割も担うことができます。この理研サイネスを使うことで、理研のさまざまなデータベース公開を共通して行える窓口である「理研総合データベース」の運営も開始しました。この共通窓口からデータベースを公開化していくことで、データベース群の統合化や横断検索が必然的に実現されるため、文部科学省が推進するライフサイエンス分野の統合データベース委託研究開発事業でも、理研サイネスを使って理研の公開データベースの統合化が効率的に推進されています。

一方、疾患研究ではオミックスによるデータ駆動型のアプローチへの期待が高まっています。このアプローチでは、解析に必要となる網羅的分子から臨床データを戦略的に蓄積し、それら多様かつ膨大なデータを介して、ゲノミクス研究者、バイオインフォマティクス研究者、基礎研究者、臨床医、専門医など幅広い分野の研究者が効果的に連携しあえる情報基盤が必須となっています。この情報基盤は多階層・双方向的であるとともに、それぞれの階層が多様な研究分野の個々の現実的ニーズに応えられる独立性も求められるため、単なるデータベース構築機能だけでは不十分です。そこで、理研 BASE では、さまざまな研究プロジェクトに携わる大勢の研究者間でのデータフロー制御や自動データ処理を実現する機能を付加することによって、ライフサイエンス研究全般にわたってさまざまな研究プロジェクトに対応できる理想的な情報基盤システムとしての理研サイネスを開発しました。

理想的な生命情報基盤に求められる沢山の機能を総称して、「ライフサイエンスネットワーキングシステム【3】」と呼んでいることから、理研が運営するシステムには「理研サイネス（RIKEN SciNeS：RIKEN Life Science Networking System）」という名称をつけました。理研サイネスには「プロジェクト運営機能【9】」があり、ユーザが新規プロジェクトの立ち上げ申請をすると、理研サイネス内にそのプロジェクト用の仮想的なワーキングエリアを作成、申請者にはそのオーナー権限が与えられてプロジェクトの運営を開始できるようになります。オーナーは、複数の参画者を自分のプロジェクト内に招き入れてネットワーク経由で研究データを介したさまざまな連携活動を行うことができ、コンテンツの相互閲覧や共同編集を時系列で行いつつ、各コンテンツに対してペーパーレビューのような高度な査読手続きを適宜設定して品質管理を行える「コラボレーション・レビュー機能【5】」や、コンテンツ内の各データアイテムにコメントを添付して時系列でやりとりできる「メッセージ機能【6】」が提供され、実験計画から結果までの研究ログが時系列でデータベースに記録される「電子ラボノート機能【7】」も活用ができます。

理研サイネスのプロジェクト運営機能には、一般的なソーシャルネットワーキングサービス（SNS: Social Networking Service）にある「コミュニティ機能【8】」のほかに、ライフサイエンス研究を強力にサポートするための特別な機能として、「オントロジープロジェクト【10】」、「データベースプロジェクト【13】」、「データベーステンプレートプロジェクト【19】」の 3 種類の「プロジェクト運営機能【9】」が提供されています。

「オントロジープロジェクト【10】」では、参画者の研究者らが、理研サイネス上でオントロジーの個々のターム項目についてコメントを交換し合いながら、共同でオントロジーを定義していくことができる「オ

ントロジー構築機能【11】」や「異種オントロジー間対応付け支援機能【12】」も提供されています。また、理研サイネスでは、国際的な連携にも対応できるように英語と日本語の併記でデータ作成することが可能になっています。

また「データベースプロジェクト【13】」では、「データベース構築機能【14】」を提供します。このプロジェクトでは、オーナーが許可した者以外では各データにアクセスできないため、理研サイネスを共同研究の初期段階からデータベース構築に利用することができ、後からアクセス権の設定を変更するだけで、データの一部または全部をデータベースとして容易に外部公開する「データベース公開機能【15】」を整備しています。また、理研サイネスでは、オントロジーの各概念クラスがデータベースとしての機能を備えています。そして、オントロジーの継承関係で複数のデータベースの統合関係を指定することで、各データベースは、オントロジーのクラス関係で自動的に階層化しており、あるデータベースで検索をかけると、その階層以下の全データベースを対象に一斉検索を実行する「階層的構造一括検索機能【35】」も連動します。サイネスを利用することで、研究者達は、データを扱うためのシステムを個別に準備する必要がなく、データ作成に専念できるようになりました。しかし、一般のユーザが、データベースを構築するために複雑なオントロジー概念やさまざまな条件を自分で指定することは難しいことです。そこでデータベースの立ち上げを容易化するための仕組みが「データベーステンプレートプロジェクト【19】」です。これは、データベース構築にあたり使用すべきオントロジーやデータクラスに関する諸条件を専門家があらかじめデータベースの雛型として登録しておいたもので、理研サイネスが提供する「データベースウィザード機能【20】」によって、ユーザはテンプレートから派生させた新しいデータベースプロジェクトを容易に立ち上げることができます。この際、同一テンプレートから派生したデータベースは、共通の親クラスを継承しているため、上述の階層化の仕組みによりそれらの派生データベースは互いに理研サイネスの中で仮想的に統合化されます。

また、データベース統合化は、外部の利用者にとっても大きな利便性をもたらします。例えば、理研には、300 を超える研究室があり、それぞれが自発的に研究成果の発表を行っています。データ公開もさまざまな形式で実施、現在約 90 のデータ提供サイトが、ワールドワイドウェブに向けて公開している状況です。このため、どこにどのようなデータを公開しているのかが、専門分野外の人には分かりづらく、公開データの再利用や、コンピュータを使った統合利用が容易ではありませんでした。また、ライフサイエンスで扱われるデータのほとんどは、形式が定まっていない集積困難なデータで、これらについては、塩基配列や DNA チップデータのような集積リポジトリが存在しておらず、データ公開を難しくしていました。そこで理研サイネスでは、各研究室が公開するデータの永続性を研究所全体の責任として担保するために、各データアイテムに対してグローバルにユニークな ID 付与して外部公開し、外部者はその ID から、該当するデータアイテムを容易に参照できる「リポジトリ機能【21】」を提供します。この ID は、国際標準として提案されている URI や Handle ID、LSID といったほかの ID 表記法にも相互変換可能なように配慮して設計されており、ほかのリポジトリシステムと連携させることが可能です。また、理研サイネスでは各データアイテムにこの ID を最初から割り振り、公開・未公開に限らずすべてのデータをシームレスに関係づけています。このため、膨大な実験データやそれに関連するさまざまなデータを、データの種類に応じて別々の外部データベースに振り分けて登録した場合であっても、後日それら全体の関係性を ID によって復元することが可能なように配慮されています。また、理研サイネスのリポジトリ機能は、従来のリポジトリのようにファイルや画像など個々のデジタルコンテンツを格納する機能だけでなく、データベース自体を利用可能な状態のまま丸ごとひとつのコンテンツとして格納する「ライブデータベースリポジトリ機能【22】」という新しい特徴があります。

理研 BASE では、上記の理想的なサイネスの構築を進めるにあたり、セマンティックウェブ技術を導入いたしました。セマンティックウェブは今世紀に入って急速に国際標準としての仕様が確立した新しい情報技術です。現在、大型研究所全体のセキュアなデータ連携と大規模なデータ公開の両方が行える情報基盤を、

国際標準であるセマンティックウェブ技術によって構築しているのは理研以外にありません。このシステムはセマンティックウェブとファイルシステムの両方の長所を併せ持つ“Semantic Web Folders（SWF)”という理研 BASE が独自に開発した情報技術を使って実装されており、数千個の異なるデータベース群を一つのシステム内に包含して同時に運用することを想定し、データベースの数が増えても安定した動作が可能なシステムとして設計しています。

セマンティックウェブの特徴は、データの意味をコンピュータが自動解釈可能な形式で保持する点にあり、これにより、コンピュータが各データの型を自動判断して、それぞれに適切なデータ処理を自動的に行う「データ自動解析機能【24】」の実現を可能にすることができます。セマンティックウェブに追加されたデータを、システム内の巡回ロボット（クローラー）が、プロジェクトごとに設定されたセキュリティ制限に配慮しながら定期的に抽出し、各種データごとに必要な処理を自動実行して、人間にわかりやすい形式のレポートを自動作成し、データの種類に応じて適切なフォーマットでダウンロード用のファイルを自動作成しています。特にオミックスデータや画像データの自動解析では、定量化処理などによりさまざまな派生データが自動生成されるため、これらデータ間の関連性もセマンティックウェブ形式で理研サイネス内に記録保持することは、データのトレーサビリティーとデータガバナンスの強化に必須です。また、プロジェクトのオーナーは各データへのアクセス権やライセンス条項を細かく設定できるため、データの「一括ダウンロード機能【26】」や、条件に合う一部のデータのみを絞り込んでダウンロードする「データマート機能【27】」を理研サイネスから提供されています。さらに、既に別のサイトから公開されているデータベースをクローリングによってサイネスに取り込む「自動セマンティックウェブ変換機能【31】」も提供されており、プロジェクトのオーナーなど限られたユーザが利用可能です。

理研サイネスのデータフロー制御機能では、コンピュータによる自動処理のフローだけでなく、その途中で、人間が介在するデータ操作が必要になるケースにも対応しています。たとえば、後述する原発性免疫不全症の専門医と臨床医をつなぐ医療ネットワークの構築では、医師の診断や専門家同士のやり取りの順序制御も含めた「データフロー制御機能【28】」が使われています。また、ラボ内でデータと人的活動の連携を制御する「ラボラトリーオートメーション【29】」の実現にも理研サイネスは有効です。さらに、理研サイネスでは「解析ツール拡張機能【23】」により、各種データの解析に必要な処理プログラムを追加することで、システム拡張を柔軟に行うことができます。例えば、上述の巡回ロボットは、理研サイネスの拡張用インタフェースを介してデータにアクセスする拡張プログラムとして実装しています。また、オミックスデータでは統計処理プログラムによる自動解析やゲノムブラウザなど、データの種類に応じた拡張プログラムに対応付けることも可能です。これにより、データの種類に応じた「データ可視化機能【25】」を理研サイネスが提供します。

最近では、パブリケーションに伴ってデータベースを公開するケースが増えていますが、一般的にそれらすべてのデータベースにその後も継続的なメンテナンスの予算が投入されるとは限りません、従来のように各データベースを個別のシステムで運用する方式だと、データベースのメンテナンスが途絶えて運用が止まり、データの利用ができなくなるケースが増えて問題となっていました。これに対し、理研サイネスでは多数のデータベースを一つのシステム内に包含して全体として運用するため、各データベースへのアクセスが不能になる事態を避けることが可能で、コストパフォーマンスの面からも非常に効率的です。また、上述の巡回ロボットがコンテンツの自動更新を行うため、人的な更新が途絶えてしまったデータベース X に対しても自動更新が継続できます。例えば、新しいデータベース N が追加された際に、そのデータベース N の新しいコンテンツから上記データベース X のコンテンツにリンクを追加すれれば、「自動更新機能【30】」によってデータベース X のコンテンツの中にもデータベース N の情報が自動かつ統合的に更新するため、「データベース自動統合更新機能【32】」が実現できます。

最近では Wiki を使った国際的なコンテンツづくりが流行していますが、Wiki は人間が読むための文書を

書き込みやすい半面、コンピュータによるデータの再利用が難しい欠点があります。セマンティックウェブはこの欠点を補うものではある反面、人間には理解しにくくなるという欠点があります。そこで理研サイネスでは、セマンティックウェブ上に Wiki で編集可能なファイルを位置づけることで、「Wiki 機能【18】」の長所とセマンティックウェブの長所を両立させています。

理研サイネスは研究者間で国際的な共同作業を行いながら、多様なデータベースを共同構築し、セマンティックウェブおよびワールドワイドウェブの両方の形式でデータ公開できる情報基盤です。これを使って理研では、マウスやシロイヌナズナなどモデル動植物の遺伝子に医学文献（Medline）を対応付ける作業を国際的な連携で人手により進めてきました。遺伝子名はシンボル名などの略語があり、文献との対応付けは単なる文字列比較だとうまくいかないため、理研サイネスが提供する「文書タグ付けルール設定機能【16】」を使って各遺伝子のタグ付けルールをデータ編集の専門家が人手で丹念に作成してきたことで、今後は、新しい文献が追加されても、そのタグ付けルールが自動適用されるため信頼性の高いタグ付けデータを「自動文書タグ付け機能【17】」で更新していくことができるようになっています。また、これらのデータを有効に活用するためにサイネスでは「推論検索機能【33】」を提供することで、セマンティックウェブの長所を最大限に生かすことができます。理研 BASE では上記の対応付けデータを対象にした「推論検索エンジン GRASE【34】」を既に開発し、Positional Medline（PosMed）という名称の検索サイトをインターネット上に公開しています（http://omicspace.riken.jp/PosMed/）。

多くの利用が見込まれるにつれ、データに対する安全対策も必要になっており、理研では、横浜研究所（神奈川県）と和光研究所（埼玉県）の間でデータストレージを二重化させるなどの「災害対策【40】」を行っています。

現在、文部科学省の委託研究開発事業である統合データベースプロジェクトを、理研ではサイネスの情報基盤で実施しており、植物オミックスデータベースおよびタンパク質結晶化実験データベースを、国際標準であるセマンティックウェブ形式で理研サイネスから公開する予定です。セマンティックウェブでは、各データの意味や関係性をコンピュータが自動解釈可能な形式で表現するものであるため、これらの膨大なデータベースを人間が単に閲覧するためでなく、コンピュータでさまざまな角度から自動解析を試みることができるようになると期待できます。

また、サイネスは遠隔地に分散したライフサイエンス研究者たちを結びつける情報基盤として、今後は個別化医療の分野における情報連携基盤としての利用に期待を寄せています。例えば、理研免疫・アレルギー科学総合研究センターの小原収グループディレクターらは、日本における原発性免疫不全症の専門医と臨床医をつなぐネットワーク構築にこれまで成功してきており、今後はこのネットワークをさらにアジア全体に広げ、さまざまな免疫疾患研究にオミックス的アプローチを応用していくための日英両言語対応版情報インフラを理研サイネスで推進する計画です。また、理研が構築を進めている「スーパーコンピュータとの連携【38】」や、インターネットを介した「さまざまなウェブサービス機能【39】」としてのアクセス、さらに「塩基配列または化合物構造のようなキーワード以外の検索【36】」にも対応できるよう今後拡張を行っていきます。

ライフサイエンスでは、計測手法の技術革新が目覚ましく、新しいタイプの測定データが次々と得られるため、多様化するデータタイプを効率的に扱うために、理研サイネスのような情報基盤に多くの研究者が依存するようになると予想され、新しいバイオインフォマティクス手法の研究と開発【2】を推進する一方で、その基盤を安定的に提供するための安定した運用【37】や、そのためのサポートスタッフ【41】も強化していく予定です。

理研サイネスのトップ画面

文科省委託事業（植物統合データベース）も理研サイネスで推進中

文科省委託事業シロイヌナズナオミックスデータの統合データベース

サイネス上で共同研究を行うための作業デスクトップが各研究者に割り当てられる

データベース構築・連携におけるやり取りを記録するためのセマンティックウェブメール機能

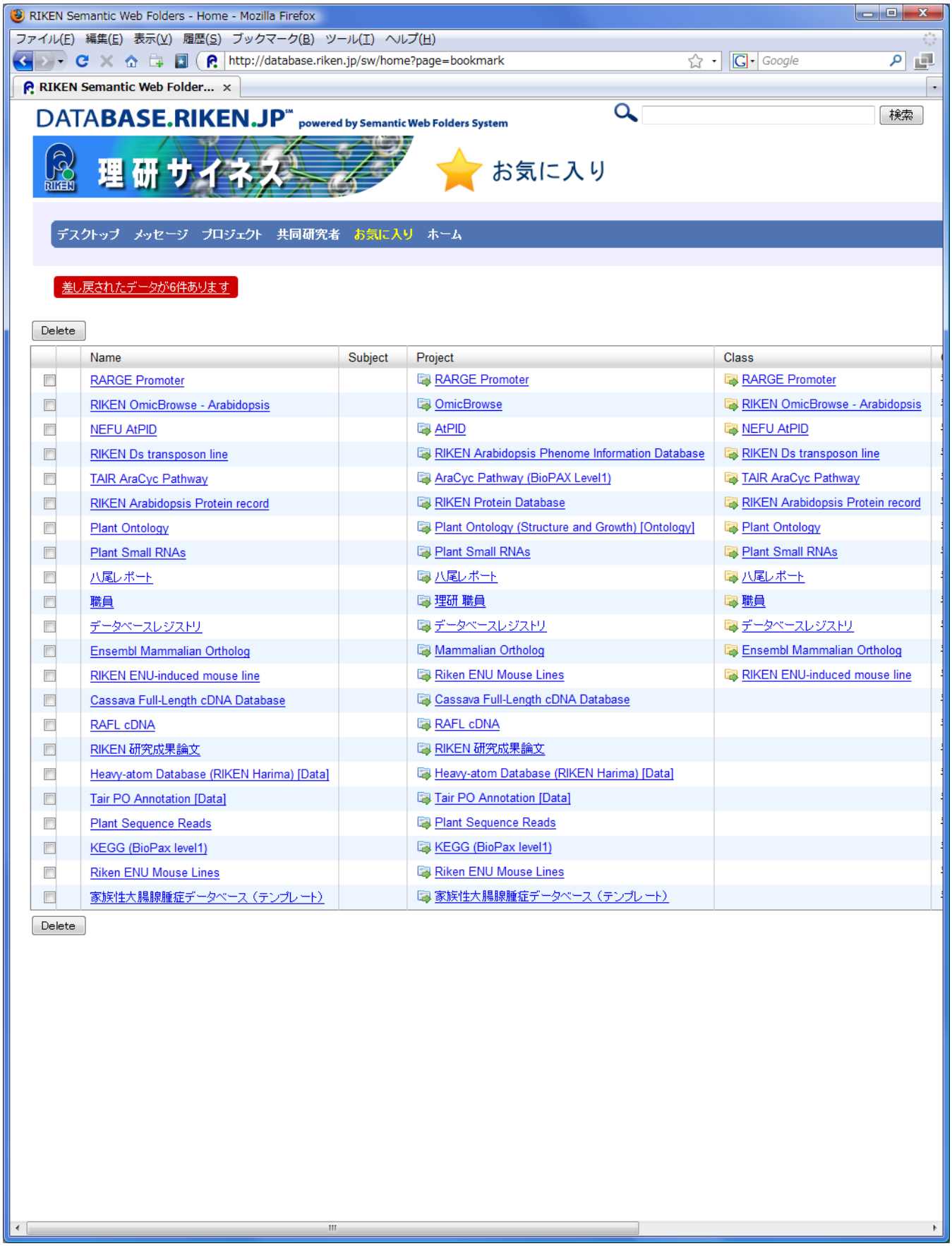
各研究者の共同研究プロジェクト一覧表示

プロジェクト名		オーナー	件数	進行状況	タイプ	更新日時
Plant Sequence Reads	view	Kaminuma Eli	7	RUNNING	DATABASE	2009-01-21 14:17:20
mRNA Targets of Plant Small RNAs	view	Kaminuma Eli	7	RUNNING	DATABASE	2009-01-21 14:17:01
ダウンロードファイル保存場所	view	小林 紀郎	6	RUNNING	DATABASE	2009-02-18 11:17:28
Strain	view	榊屋 啓志	7	RUNNING	DATABASE	2007-10-15 17:31:26
データベースレジストリ	view	小林 紀郎	1	RUNNING	DATABASE	2009-01-21 15:37:09
実験テンプレート	view	小林 紀郎	17	RUNNING	TEMPLATE	2009-02-16 20:51:00
Bacpedia beta (RIKEN Harima)	view	所 有子	25	RUNNING	DATABASE	2008-11-20 15:12:20
Plant Small RNAs	view	Kaminuma Eli	11	RUNNING	DATABASE	2009-01-21 14:16:39
Tair PO Annotation [Data]	view	Kaminuma Eli	5	RUNNING	DATABASE	2009-01-21 14:16:15
Bacpedia (Riken Harima SPring8-Center)	view	Kaminuma Eli	29	RUNNING	DATABASE	2009-02-03 11:23:32
理研総合データベース	view	所 有子	33	RUNNING	DATABASE	2009-01-23 09:50:37
RIKEN Arabidopsis Activation tagging line	view	Kaminuma Eli	12	RUNNING	DATABASE	2009-02-16 16:57:47
RIKEN Rice FOX Arabidopsis mutant line	view	Kaminuma Eli	0	RUNNING	DATABASE	2009-02-10 15:18:40
EC table	view	Kaminuma Eli	1	RUNNING	DATABASE	2009-01-21 14:24:39
文献レポジトリ	view	蒔田 由布子	1	RUNNING	DATABASE	2009-02-10 15:28:49
Arabidopsis KEGG PATHWAY Annotation	view	Kaminuma Eli	1	RUNNING	DATABASE	2009-01-21 14:20:36
理研データベースアンケート	view	所 有子	8	RUNNING	DATABASE	2009-01-28 12:08:25
家族性大腸腺腫症データベース（テンプレート）	view	平田 直彦	1	SUSPENDED	TEMPLATE	2009-02-18 09:59:19
Bacpedia beta version (Riken Harima SPring8-Center)	view	Kaminuma Eli	0	CANCELED	DATABASE	2009-02-17 16:54:13
Domain	view	望月 芳樹	1	RUNNING	DATABASE	2008-11-26 13:52:13
八尾レポート	view	所 有子	1	RUNNING	DATABASE	2009-02-17 17:28:40
Tair GO Annotation [Data]	view	Kaminuma Eli	5	RUNNING	DATABASE	2009-01-21 14:17:51
GO Slim Generic	view	所 有子	2	RUNNING	DATABASE	2008-10-15 13:48:58
GO Slim Plant	view	所 有子	2	RUNNING	DATABASE	2008-10-15 13:48:12
ファイルレジストリテンプレート	view	蒔田 由布子	1	RUNNING	TEMPLATE	2008-12-16 19:48:08
OmicBrowse	view	平田 直彦	3	RUNNING	DATABASE	2008-12-16 18:57:24
RIKEN 研究成果論文	view	Kaminuma Eli	3	RUNNING	DATABASE	2009-01-21 14:10:04
SWF 編集Q&A	view	Kaminuma Eli	2	RUNNING	DATABASE	2009-01-21 14:11:34
ファイルレジストリ	view	蒔田 由布子	1	RUNNING	DATABASE	2009-02-18 11:18:14
ファイルレジストリ	view	蒔田 由布子	0	RUNNING	DATABASE	2009-02-17 13:35:32
家族性大腸腺腫症患者データ（テンプレート）	view	平田 直彦	5	SUSPENDED	TEMPLATE	2009-02-18 10:04:07
KEGG (BioPax level1)	view	蒔田 由布子	27	RUNNING	DATABASE	2008-08-01 15:49:01
家族性大腸腺腫症遺伝子情報（テンプレート）	view	平田 直彦	1	SUSPENDED	TEMPLATE	2009-02-18 10:04:46
FAP（家族性大腸腺腫症）データベース	view	平田 直彦	4	SUSPENDED	DATABASE	2009-02-18 09:58:40
理研 職員	view	所 有子	1	RUNNING	DATABASE	2009-02-12 14:42:27
家族性大腸腺腫症変異データ（テンプレート）	view	平田 直彦	7	SUSPENDED	TEMPLATE	2009-02-18 10:03:38
Plant Ontology (Structure and Growth) [Ontology]	view	Kaminuma Eli	2	FINISHED	ONTOLOGY	2009-01-21 14:13:45
RIKEN Arabidopsis Phenome Information Database	view	平田 直彦	13	RUNNING	DATABASE	2009-01-22 12:30:41

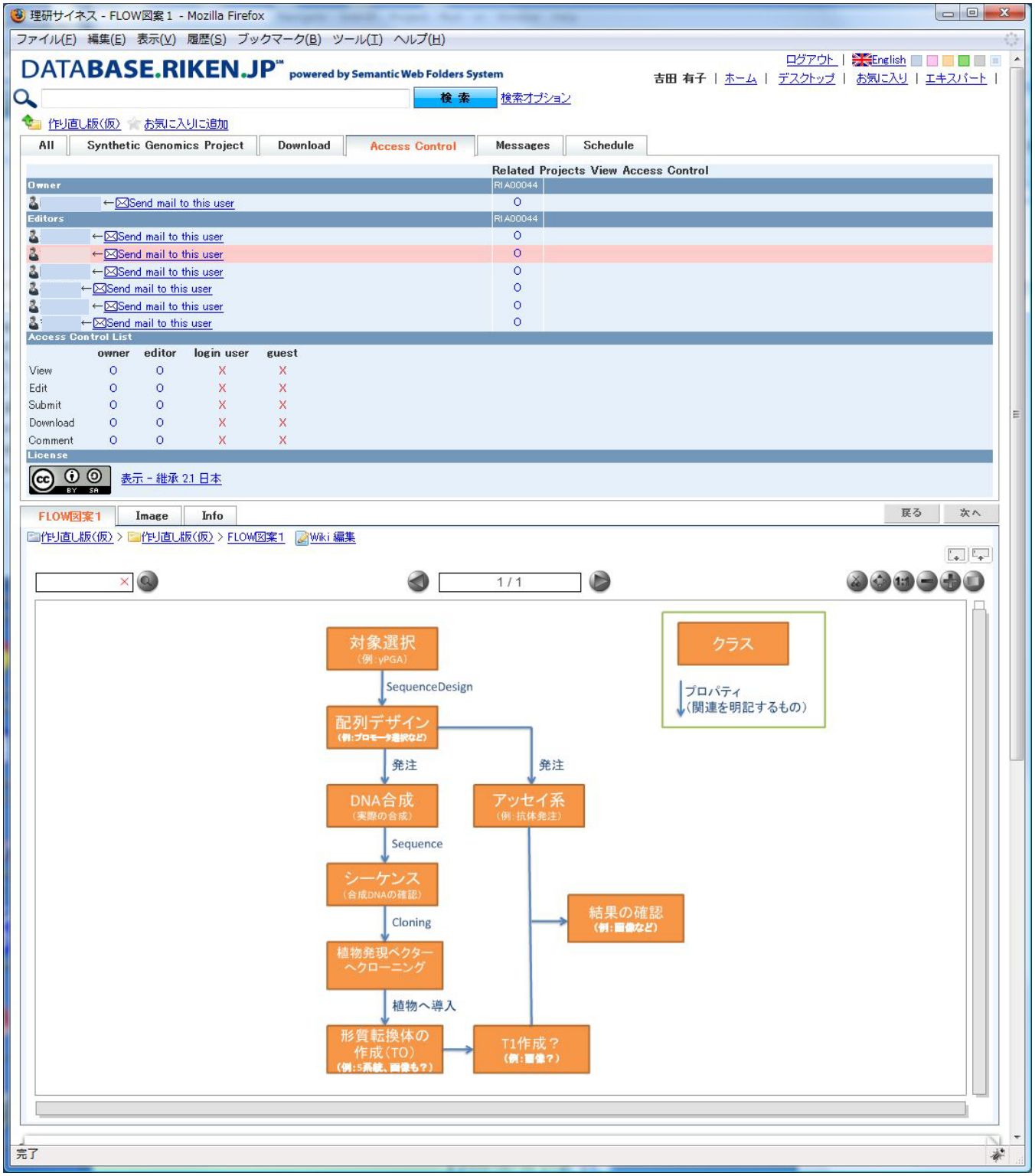
共同研究者一覧表示


RIKEN Semantic Web Folders - Home - Mozilla Firefox
ファイル(F) 編集(E) 表示(V) 履歴(S) ブックマーク(B) ツール(T) ヘルプ(H)
http://database.riken.jp/sw/home?page=friends
RIKEN Semantic Web Folder...
DATABASE.RIKEN.JP powered by Semantic Web Folders System
検索
理研サイネス
共同研究者
平田 直彦
デスクトップ メッセージ プロジェクト 共同研究者 お気に入り ホーム
設定 | ログアウト | English
差し戻されたデータが6件あります
共同研究者を招待
豊田 哲郎
男性
生命情報基盤研究部門
toyop@gsc.riken.jp
小林 紀郎
男性
生命情報基盤研究部門
nori@base.riken.jp
望月 芳樹
男性
オミック情報統合化研究チーム
fmoon@gsc.riken.jp
所 有子
女性
オミック情報統合化研究チーム
tokoro@gsc.riken.jp
浅野 (松下) さとみ
女性
生命情報基盤研究部門
ma31stm@gsc.riken.jp
8232
蒔田 由布子
女性
生命情報基盤研究部門
ymakita@gsc.riken.jp
8236
Kaminuma Eli
女性
eli@base.riken.jp
石井 学
男性
manabu@base.riken.jp
松嶋 明宏
男性
生命情報基盤研究部門
amatsus@base.riken.jp
8163
上田 浩美
女性
生命情報基盤研究部門
ueda@base.riken.jp
向山 敦洋
男性
オミック情報統合化研究チーム
mukai@gsc.riken.jp

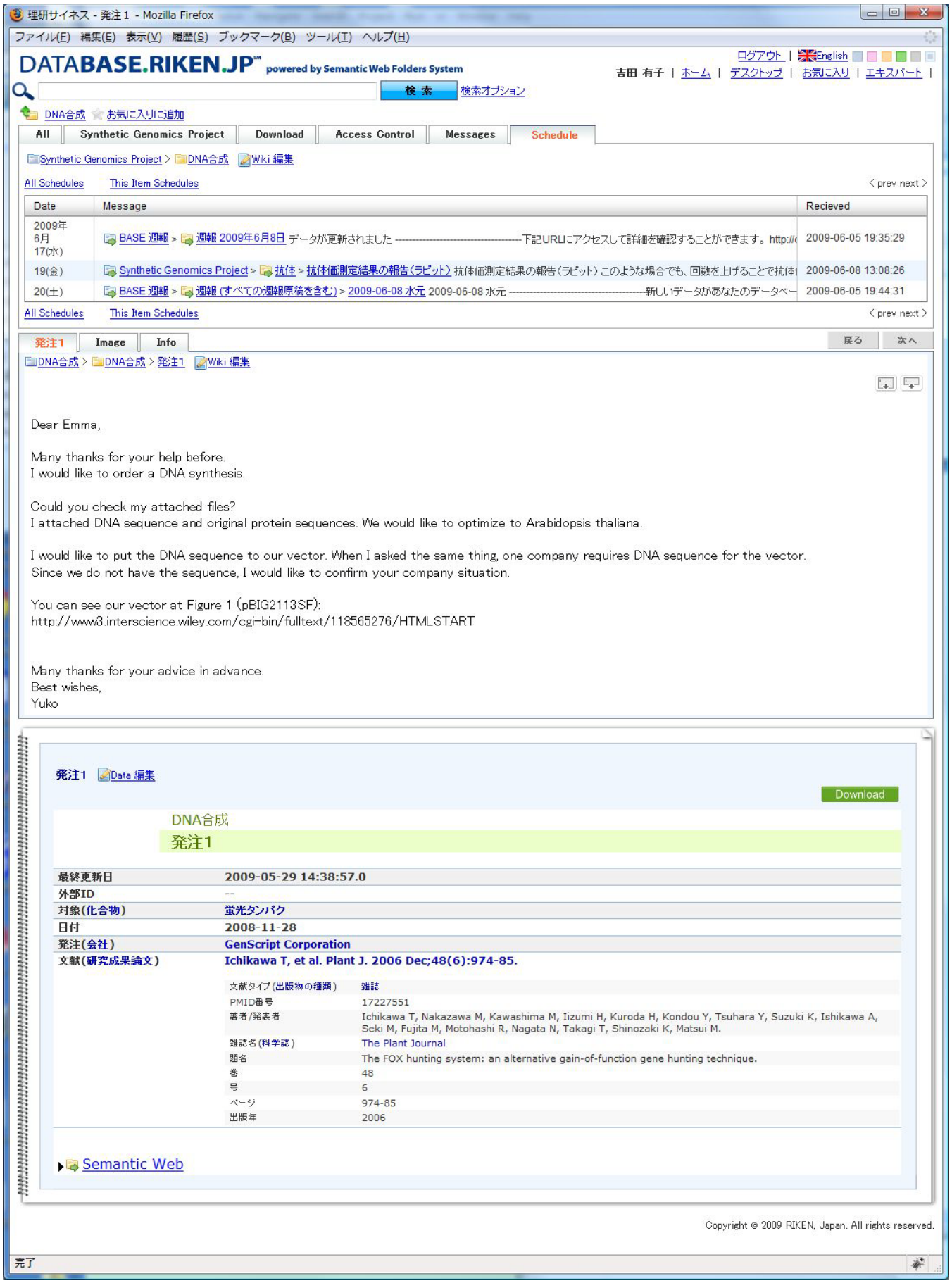
作業スペースやプロジェクト推進で使用するブックマーク機能



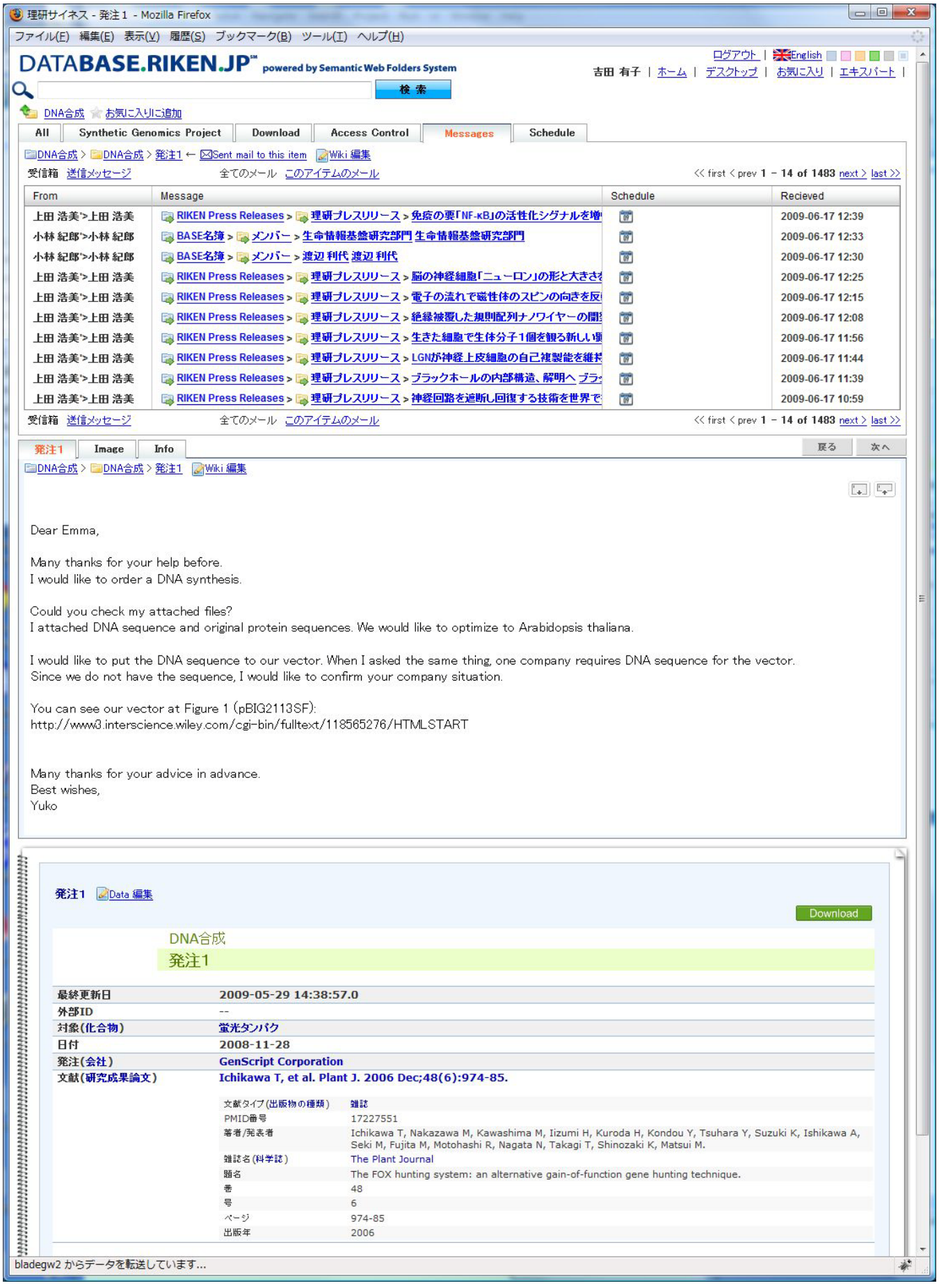
バーチャルラボ（電子ラボノート機能）



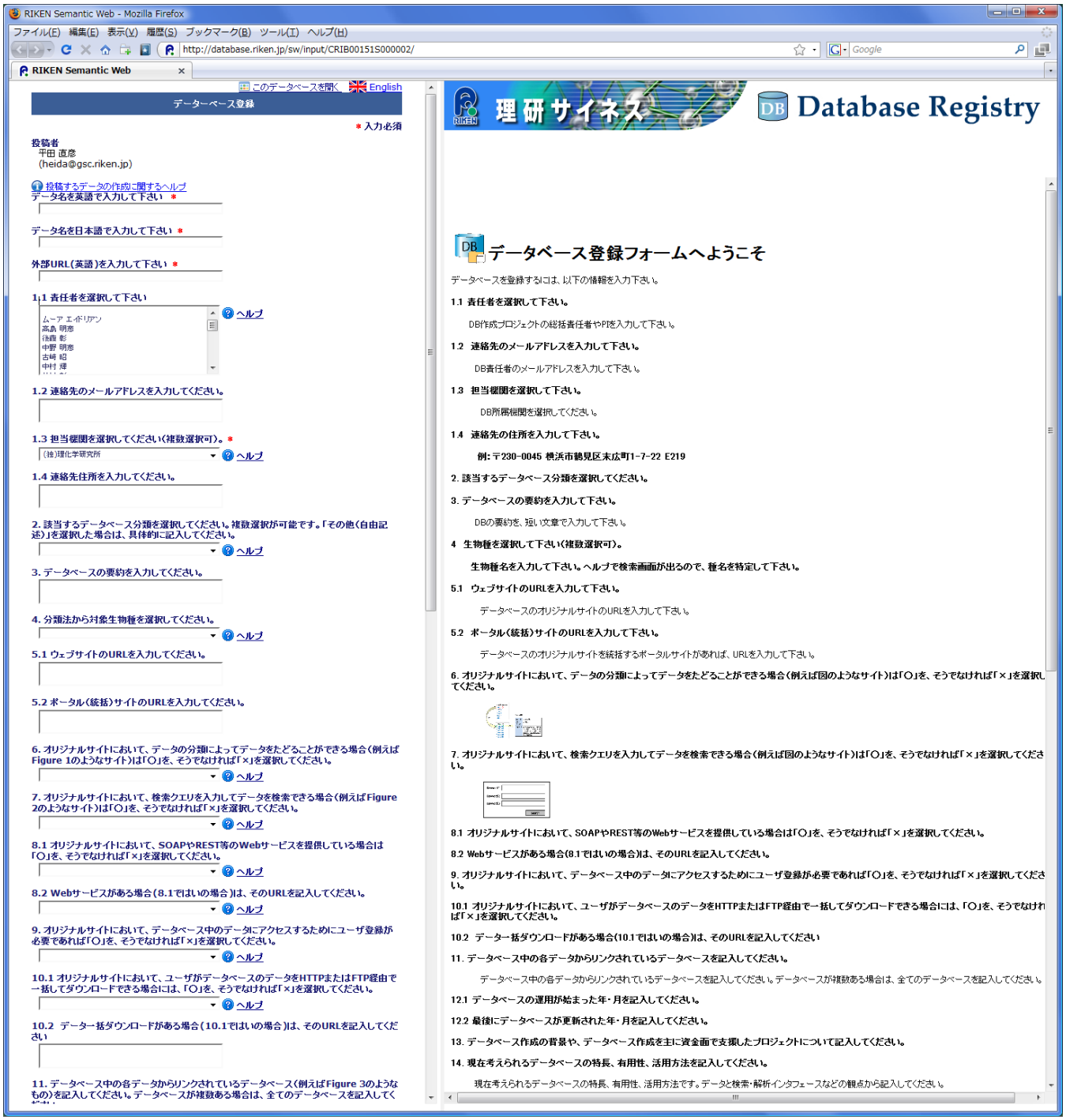
バーチャルラボ（スケジュール管理機能）



バーチャルラボ（メッセージ機能）



理研データベース登録フォーム

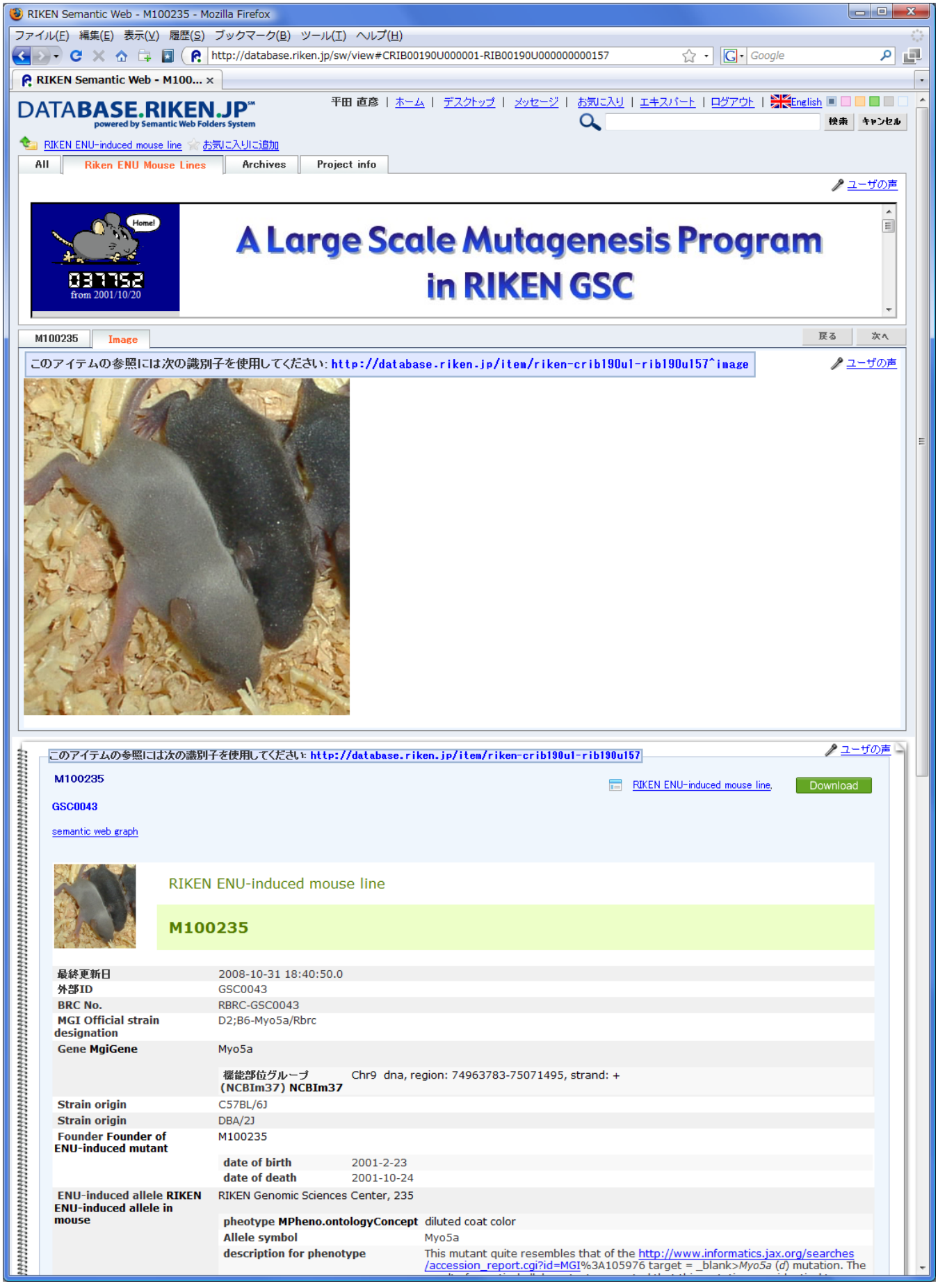


変異データ入力フォーム

患者データ入力フォーム

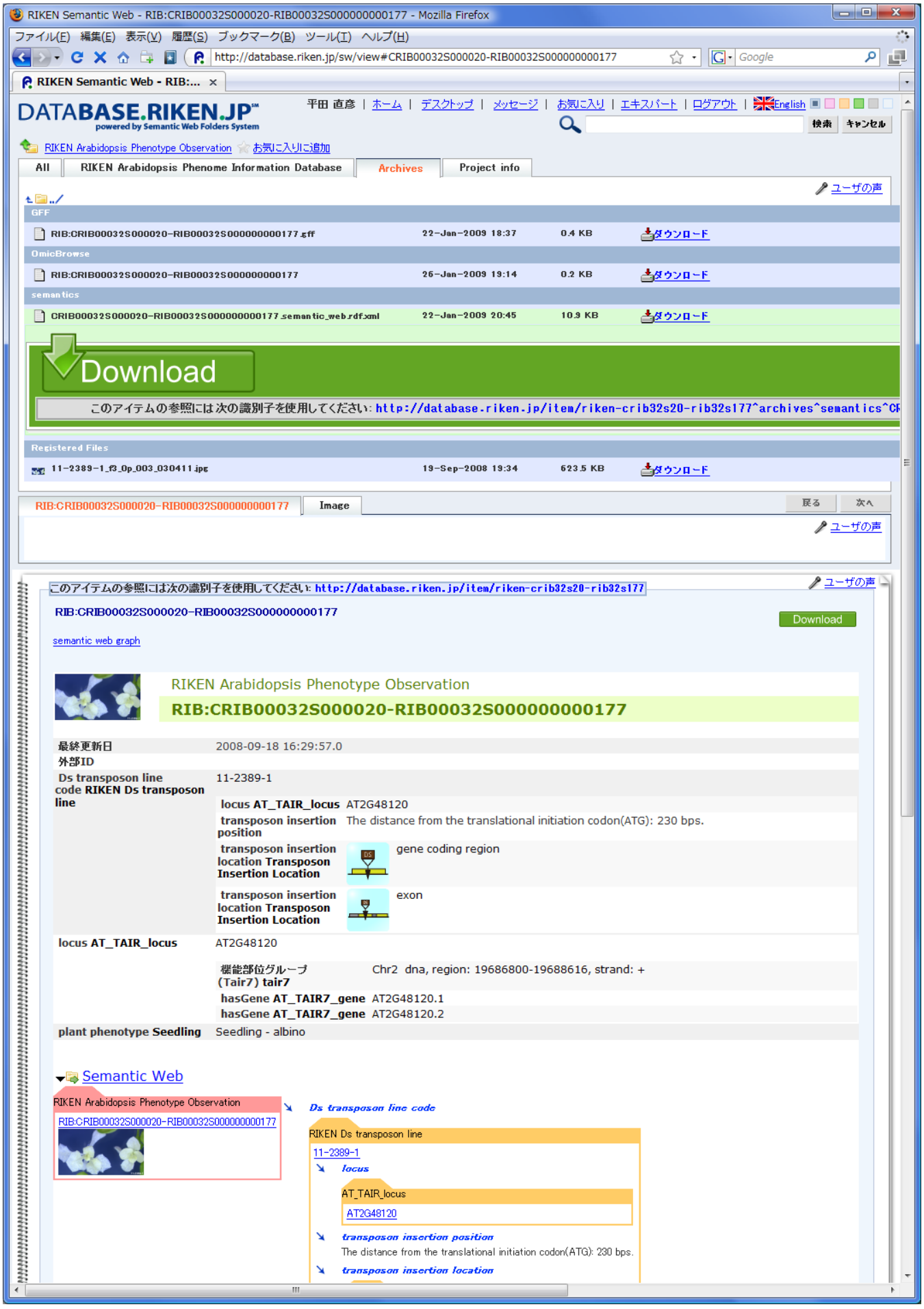
理研変異マウス系統データベース

理研変異マウス系統データベース

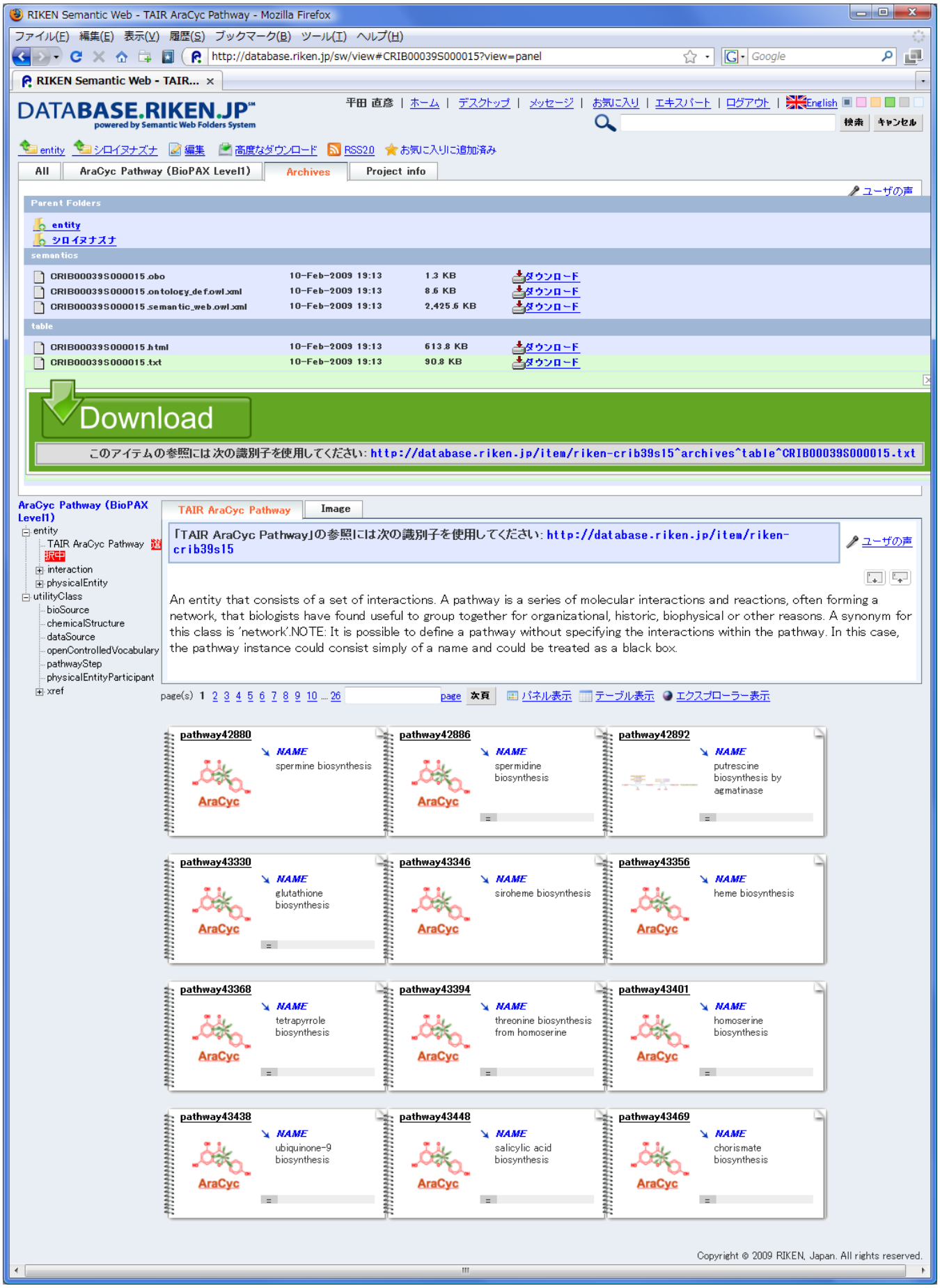


理研シロイヌナズナ変異体表現型データベース

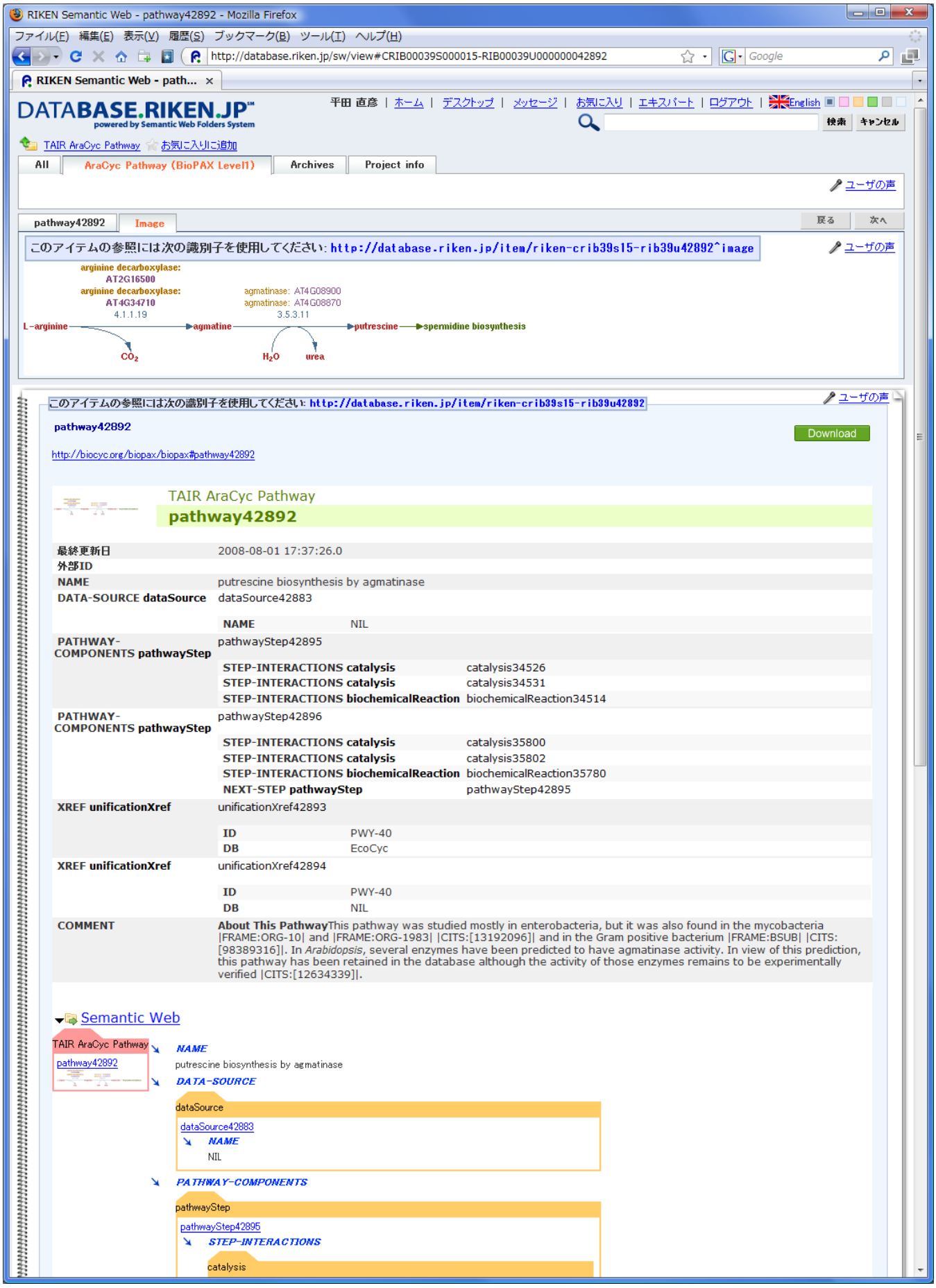
理研シロイヌナズナ変異体表現型データベース



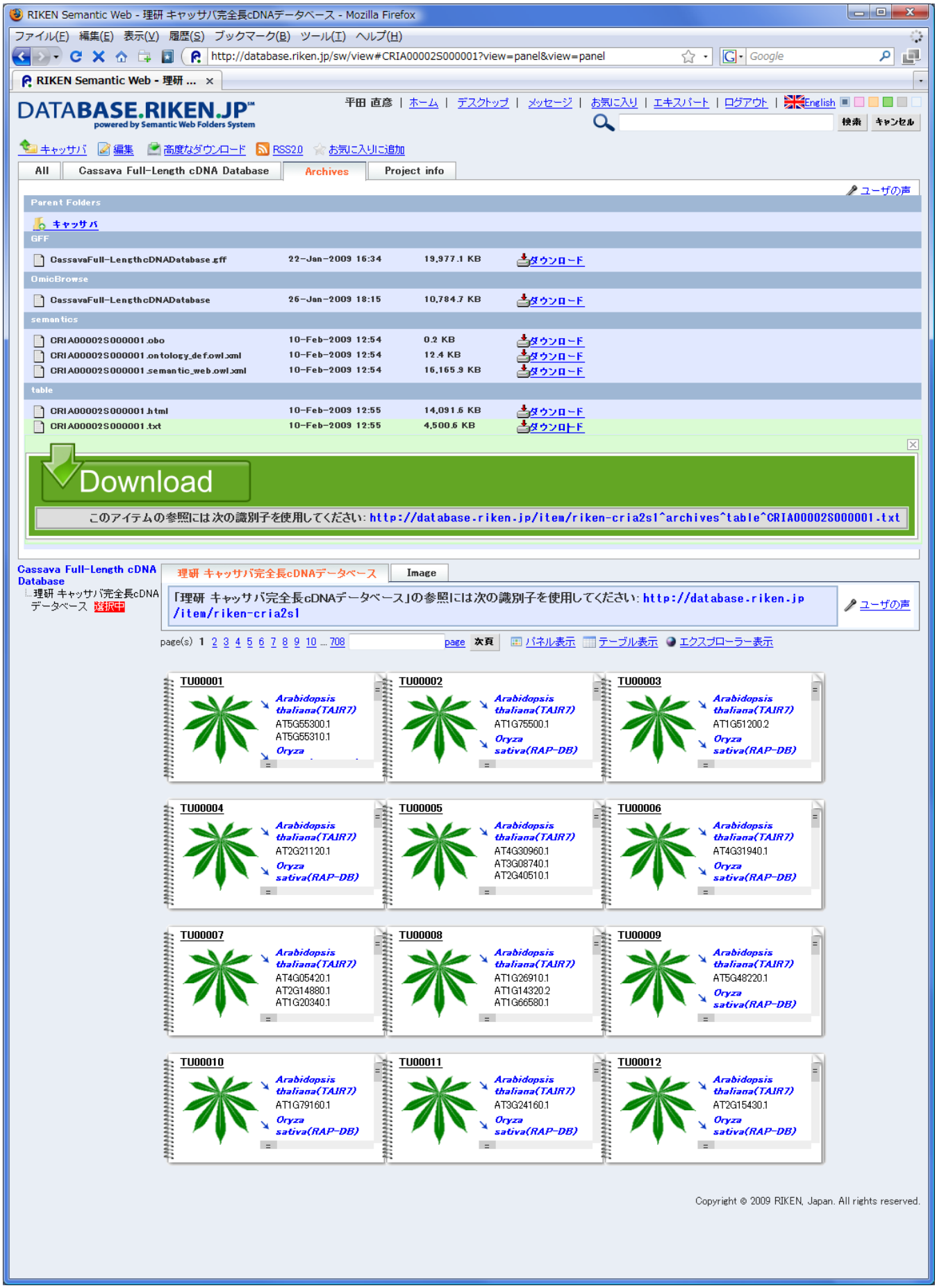
シロイヌナズナ生化学パスウェイデータベース



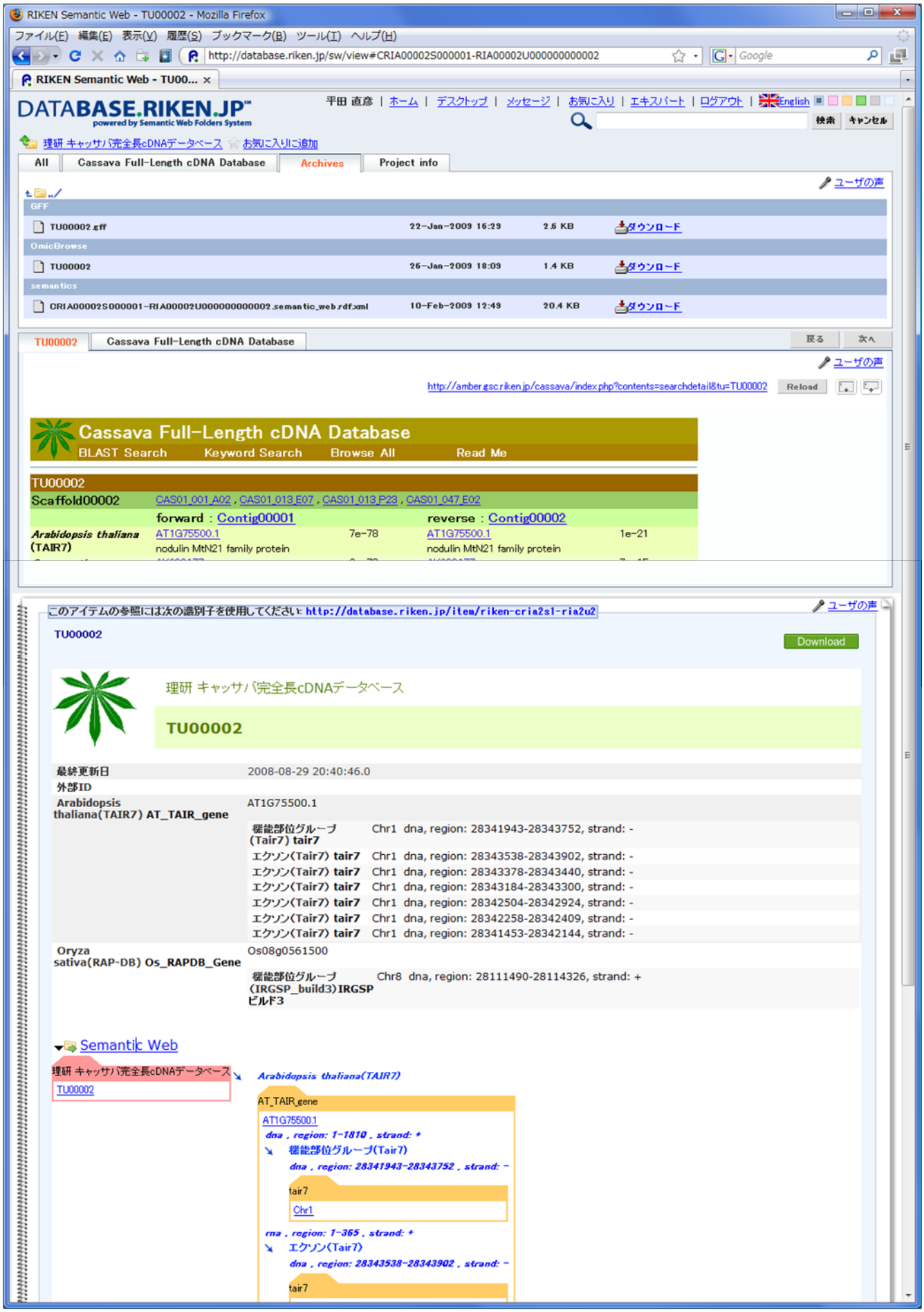
シロイヌナズナ生化学パスウェイデータベース



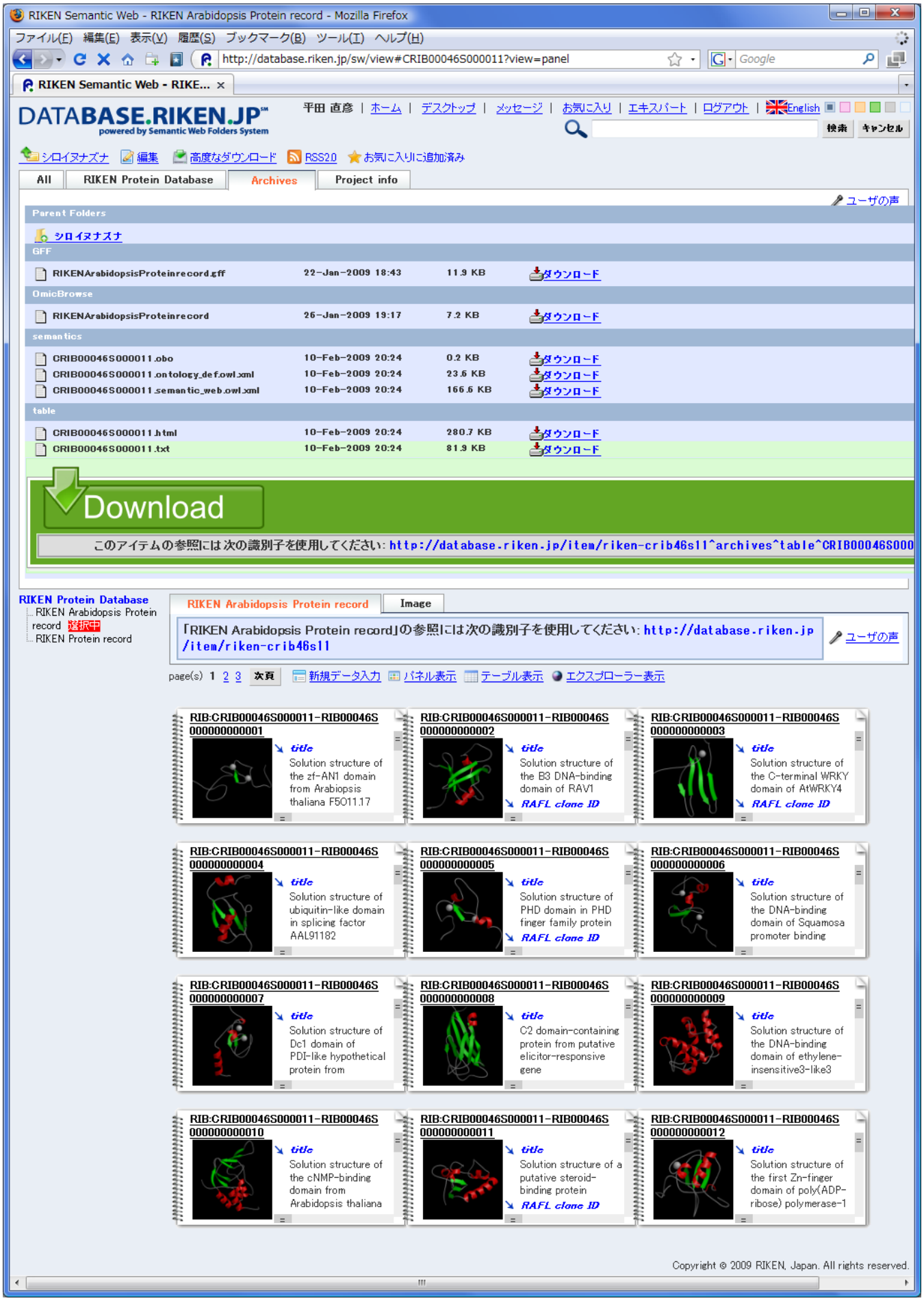
キャッサバ全長 cDNA データベース



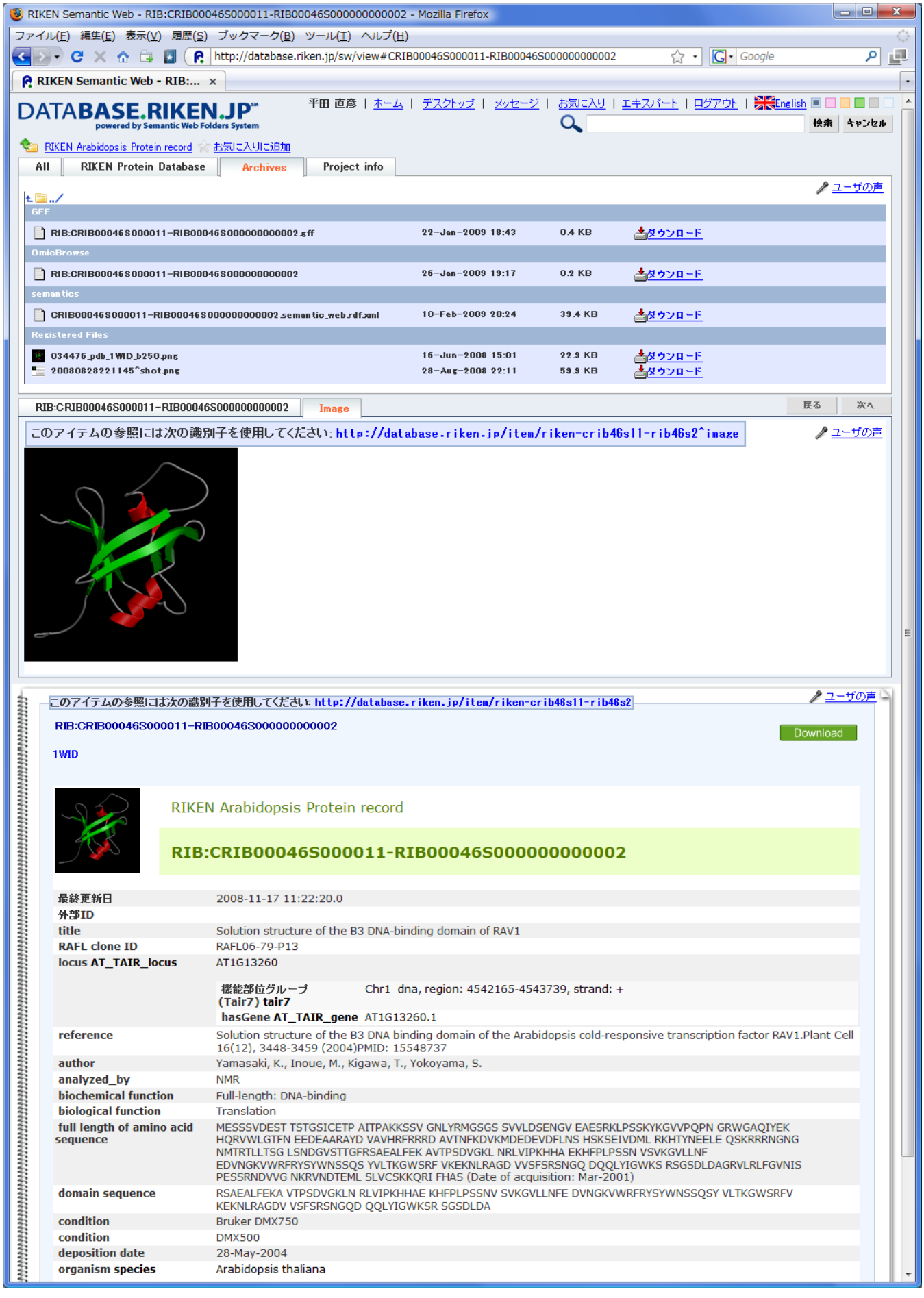
キャッサバ全長 cDNA データベース



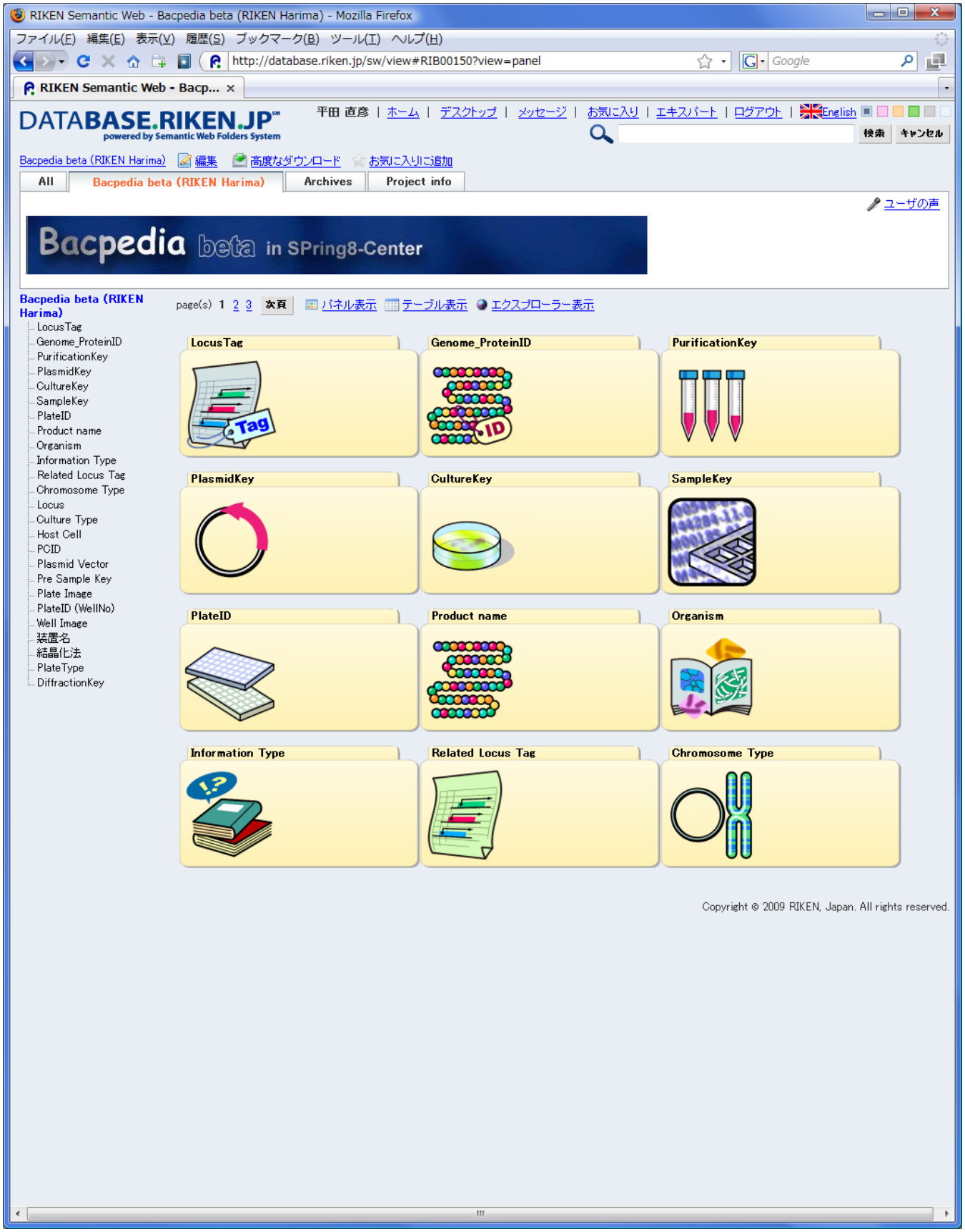
理研シロイヌナズナ蛋白データベース



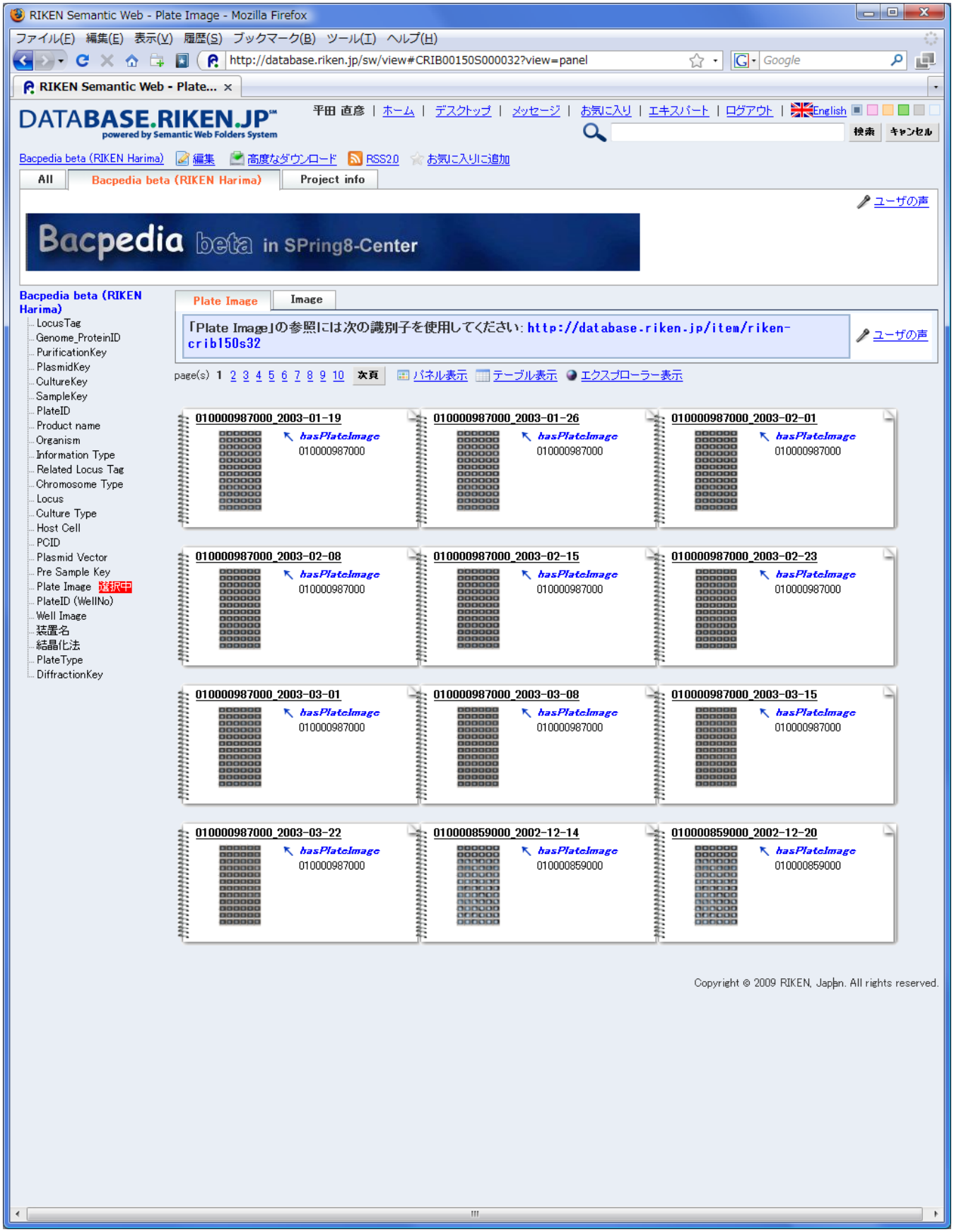
理研シロイヌナズナ蛋白データベース



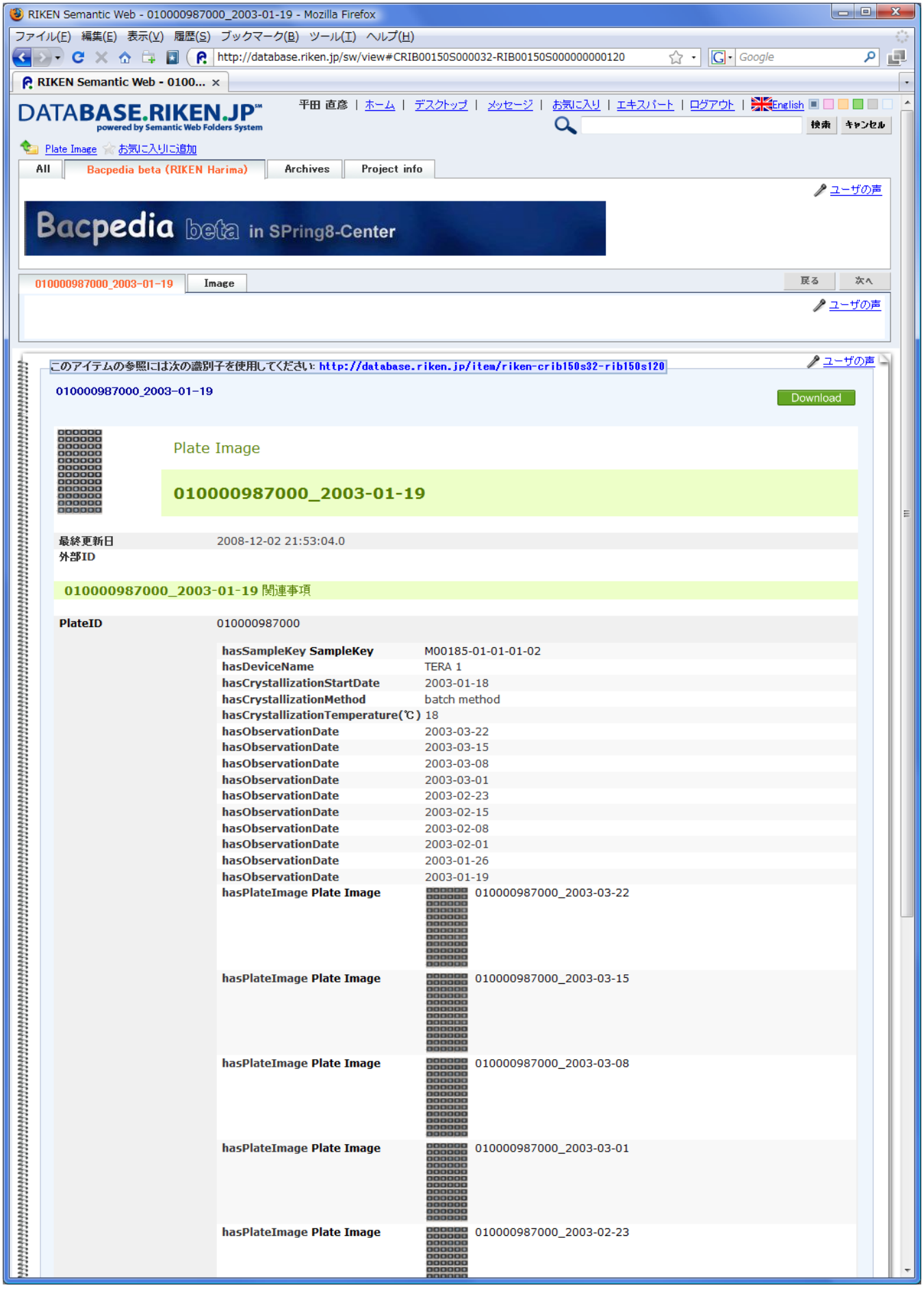
理研蛋白結晶回折データベース



理研蛋白結晶回折データベース



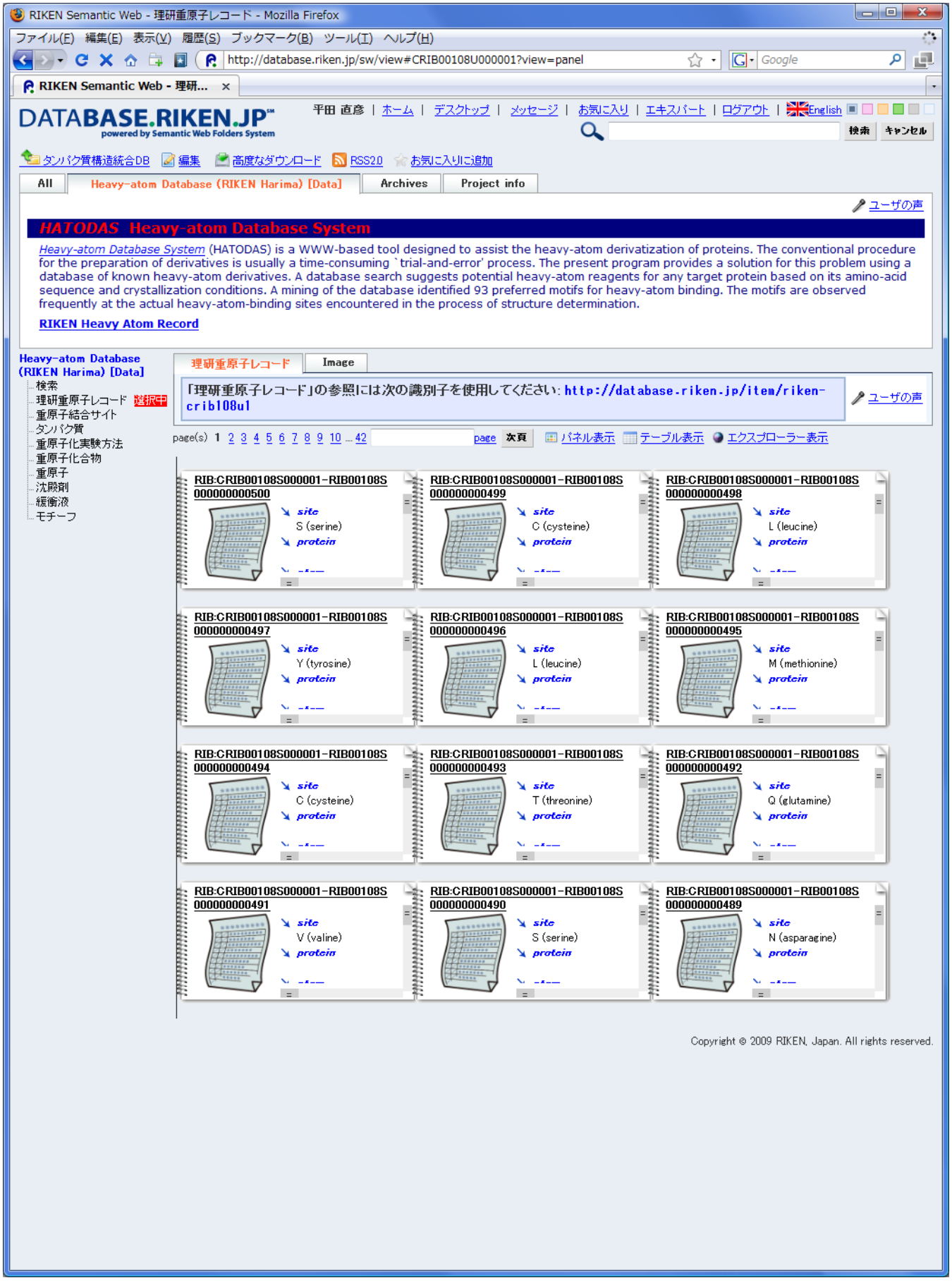
理研蛋白結晶回折データベース



理研蛋白結晶回折データベース

理研蛋白結晶回折データベース

理研重原子データベース

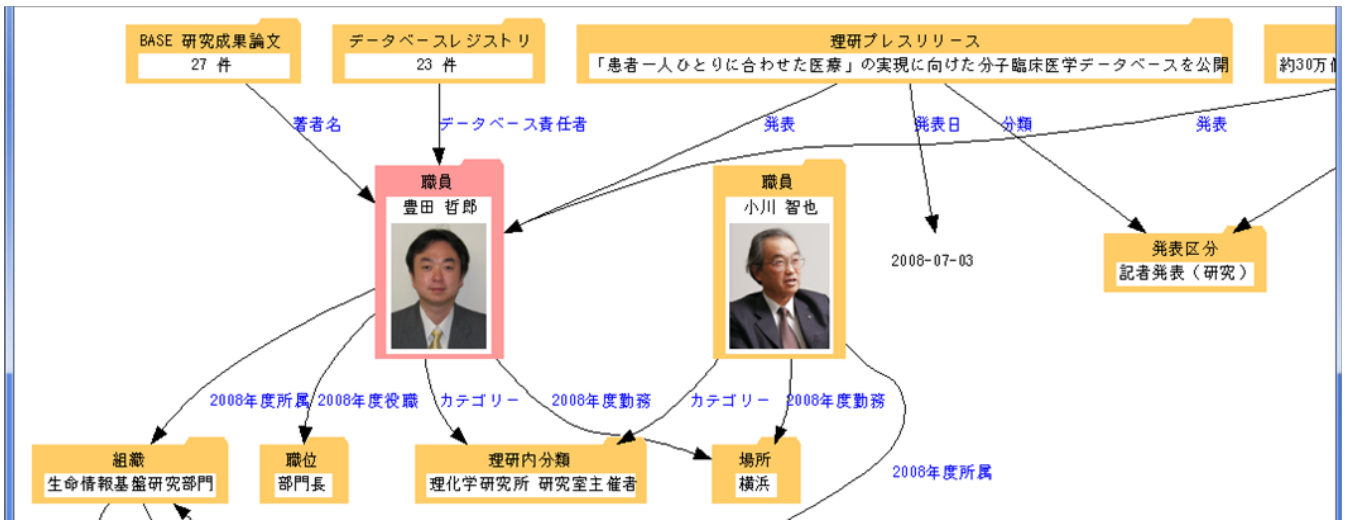


理研職員


RIKEN Semantic Web - 職員 - Mozilla Firefox
ファイル(F) 編集(E) 表示(V) 履歴(S) ブックマーク(B) ツール(T) ヘルプ(H)
http://database.riken.jp/sw/view#CRIB00139U000001
RIKEN Semantic Web - 職員
DATABASE.RIKEN.JP
powered by Semantic Web Folders System
平田 直彦 | ホーム | デスクトップ | メッセージ | お気に入り | エキスパート | ログアウト | English
検索 キャンセル
研究者 Agent 編集 高度なダウンロード RSS2.0 お気に入りに追加済み
All 理研 職員 Archives Project info
ユーザの声
理研サイネス
Researcher
RIKEN Persons
職員 選択中
職員 Image
「職員」の参照には次の識別子を使用してください: http://database.riken.jp/item/riken-crib139u1
ユーザの声
page(s) 1 2 3 4 5 6 7 8 9 10 ... 37 page 次頁
新規データ入力 パネル表示 テーブル表示 エクスプローラー表示
板谷 光泰
2008年度所属
生命情報基盤研究部門
2008年度勤務
横浜
八尾 徹
2008年度所属
(独)理化学研究所
2008年度勤務
横浜
カテゴリー
花田 耕介
2008年度所属
機能開発研究グループ
2008年度勤務
横浜
ムーア エイドリアン
2008年度役職
ユニットリーダー
2008年度所属
Moore研究ユニット
中尾 愛子
2008年度役職
チームリーダー
2008年度所属
連携支援チーム
2008年度勤務
新海 暁男
2008年度役職
チームリーダー
2008年度所属
機能解析第1研究チーム
中野 明彦
2008年度役職
グループディレクター
主任研究員
チームリーダー
2008年度所属
高島 明彦
2008年度役職
チームリーダー
2008年度所属
アルツハイマー病研究チーム
田仲 昭子
2008年度役職
チームリーダー
2008年度所属
応用生物学研究チーム
岡本 晃充
2008年度役職
独立主幹研究員
2008年度所属
岡本独立主幹研究ユニット
木寺 詔紀
2008年度役職
チームリーダー
2008年度所属
分子スケール研究開発チーム
海老原 章郎
2008年度役職
チームリーダー
2008年度所属
機能解析第2研究チーム
Copyright © 2009 RIKEN, Japan. All rights reserved.

理研職員


RIKEN Semantic Web - 豊田 哲郎 - Mozilla Firefox
DATABASE.RIKEN.JP
powered by Semantic Web Folders System
理研サイネス
Researcher
豊田 哲郎
Image
このアイテムの参照には次の識別子を使用してください: http://database.riken.jp/item/riken-crib139ul-rib139u312
Download
semantic web graph
職員
豊田 哲郎
最終更新日 2009-01-20 19:49:34.0
外部ID
2008年度役職職位 部門長
2008年度所属組織 生命情報基盤研究部門
上位組織 横浜研究所
略称 BASE
2008年度勤務場所 横浜
カテゴリー理研内分類 理化学研究所 研究室主催者
豊田 哲郎 関連事項
BASE 研究成果論文 27 件
理研プレスリリース 「患者一人ひとりに合わせた医療」の実現に向けた分子臨床医学データベースを公開
分類発表区分 記者発表(研究)
発表日 2008-07-03
理研プレスリリース 約30万個の分散データベース群を高速に統合結合する技術を開発
分類発表区分 記者発表(研究)
発表日 2008-03-19
参考文献 PMID: 18263641

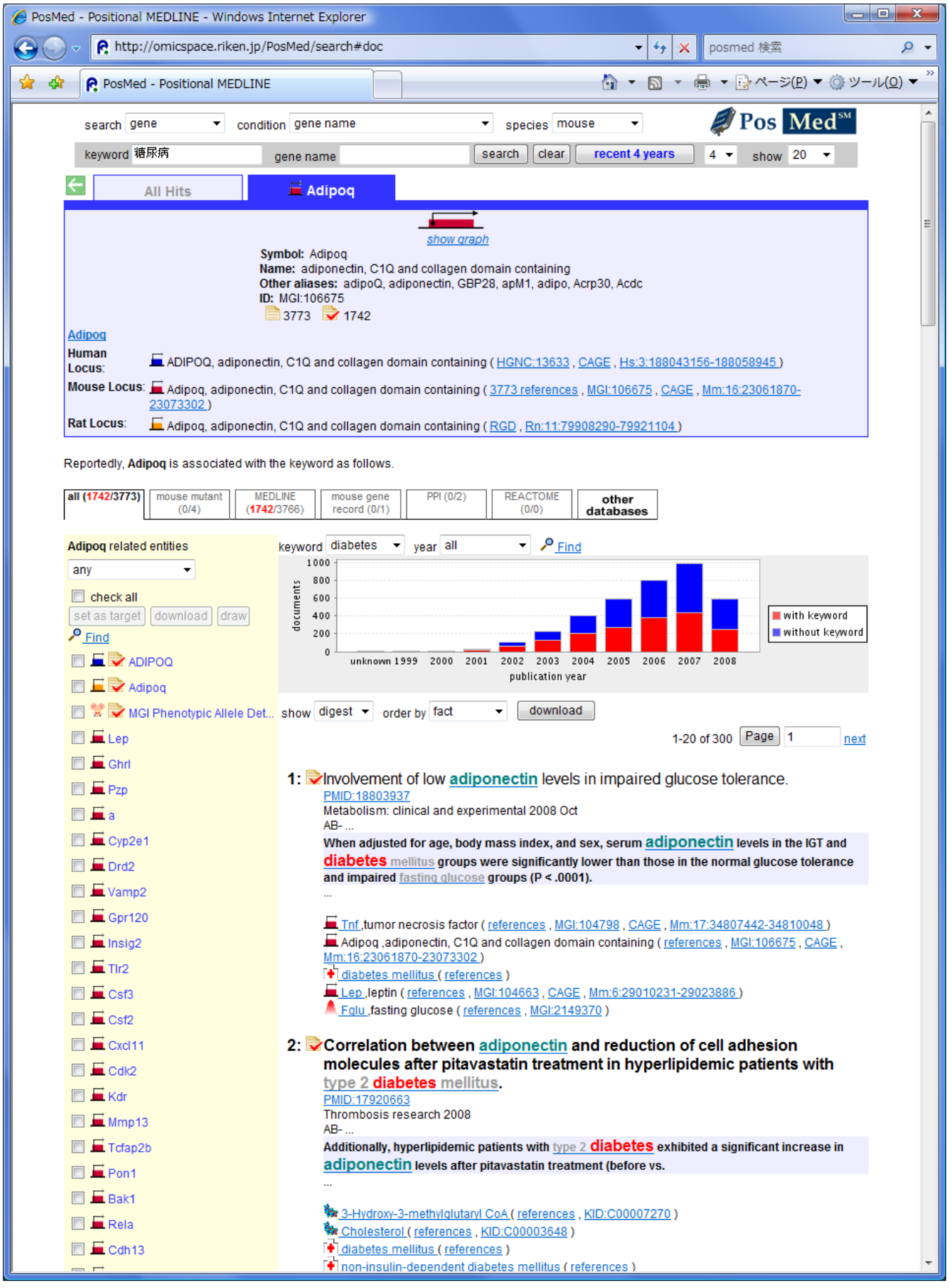

BASE 研究成果論文
27 件
データベースレジストリ
23 件
理研プレスリリース
「患者一人ひとりに合わせた医療」の実現に向けた分子臨床医学データベースを公開
著者名
データベース責任者
発表
発表日
分類
発表
職員
豊田 哲郎
職員
小川 智也
2008-07-03
発表区分
記者発表（研究）
2008年度所属
2008年度役職
カテゴリー
2008年度勤務
カテゴリー
2008年度勤務
組織
生命情報基盤研究部門
職位
部門長
理研内分類
理化学研究所 研究室主催者
場所
横浜
2008年度所属

理研独自のサーチエンジンを利用したデータベースの横断検索サイト

http://omicspace.riken.jp/db/index.html

理研推論検索サイト PosMed (*Positional MEDLINE*)

http://omicspace.riken.jp/PosMed/



理研推論検索サイト PosMed (*Positional MEDLINE*)

http://omicspace.riken.jp/PosMed/