

平成21年度目標と進捗報告（植物オミックス情報および蛋白質構造情報）

	H20年度最終目標	H21年度目標
1. シロイヌナズナオミックス注釈（豊田）	シロイヌナズナの発現、表現型、リソースに関する計6DBを統合化して公開。	前年度に公開したデータについて、アノテーション作業を行い、アノテーションデータを公開する。また、それ以外のデータについても国際的な合意をえつつ、データ公開を進める。
2. 高等動植物由来蛋白質構造のアノテーション（横山）	高等動植物等由来の蛋白質構造データを付随する実験データを含めて30件公開。	平成21年度は、タンパク3000プロジェクトで解明された高等動植物等由来の蛋白質構造データのアノテーション、およびそれに付随する実験データのキュレーションとアノテーションを行う。
3. 微生物由来蛋白質構造のアノテーション（国島）	微生物由来蛋白質に関わる試料調整（発現プラスミド構築実験 1万、培養実験 5千、精製実験 3千）、結晶化実験データ（結晶化条件：90万件、観察1000万件）と200件の回折実験データ公開。 変異導入蛋白質に関わる実験データ150件、重原子導入蛋白質に関わるデータ500件を公開。	平成21年度は、前年度に公開したデータについてアノテーション作業を行い、アノテーションデータを公開する。またそれ以外のデータについても国際的な合意を得つつデータ公開を進める。
4. アノテーションシステムの開発運用とデータ変換（豊田）	アノテーションシステムの開発運用と変換データの中核機関への提供（契約締結後）	前年度に公開したデータについて、アノテーション作業を行うためのシステムを運用・改善し、作成されたアノテーションを公開用データに変換する。

平成20年度進捗状況(豊田担当①)植物オミックス情報注釈)

年度目標

①10件のうち6件のDBを統合化して公開

12月時点進捗

①**発現に関する4件**はゲノムブラウザに統合して公開した。
表現型とリソースはセマンティックウェブ化が終了し下記外部公開サーバから公開を予定。

DB番号	データベース名	オーム空間
1	RARGE 完全長cDNAクローン (RAFL)	transcriptome
2	タイリングアレイ	transcriptome
3	454 small RNA library	genome
4	代謝質量分析プロファイル	metabolome
5	RARGE トランスポゾンタグライ ン(RARGE)	genome
6	RIKEN Arabidopsis Phenome Information Database (RAPID)	phenome
7	Activationタグライン	phenome
8	BRCリソース	resource
9	ストラクチュローム(高等動植 物等由来)	proteome
10	ストラクチュローム(微生物由 来)	proteome

セマンティクス化データ

The screenshot shows the RIKEN Hub Database interface. The search results for 'シロイヌナズナ' are displayed in a grid of tiles. The tiles are numbered 1 through 10, corresponding to the table entries. The tiles are: 1. RIKEN RARGE Database (RAFL cDNA, RARGE Alternative splicing, RARGE Promoter, RARGE Transposon); 2. RIKEN OmicBrowse - Arabidopsis (Tiling Array); 3. Plant Small RNAs (miRNA, siRNA, tasiRNA); 4. RIKEN Activation tagging line (ACT); 5. RIKEN Ds transposon line (DS); 6. RIKEN Arabidopsis Phenoty... (RAPID); 7. RIKEN Arabidopsis Protein...; 8. TAIR PO Annotation (Growth, Struct); 9. RIKEN Arabidopsis Phenome Information Database (RAPID); 10. RIKEN Arabidopsis Phenome Information Database (RAPID). The tiles are highlighted with red and green boxes.

平成20年度進捗状況(豊田担当④)理研のDB統合化のためのモデルケース構築)

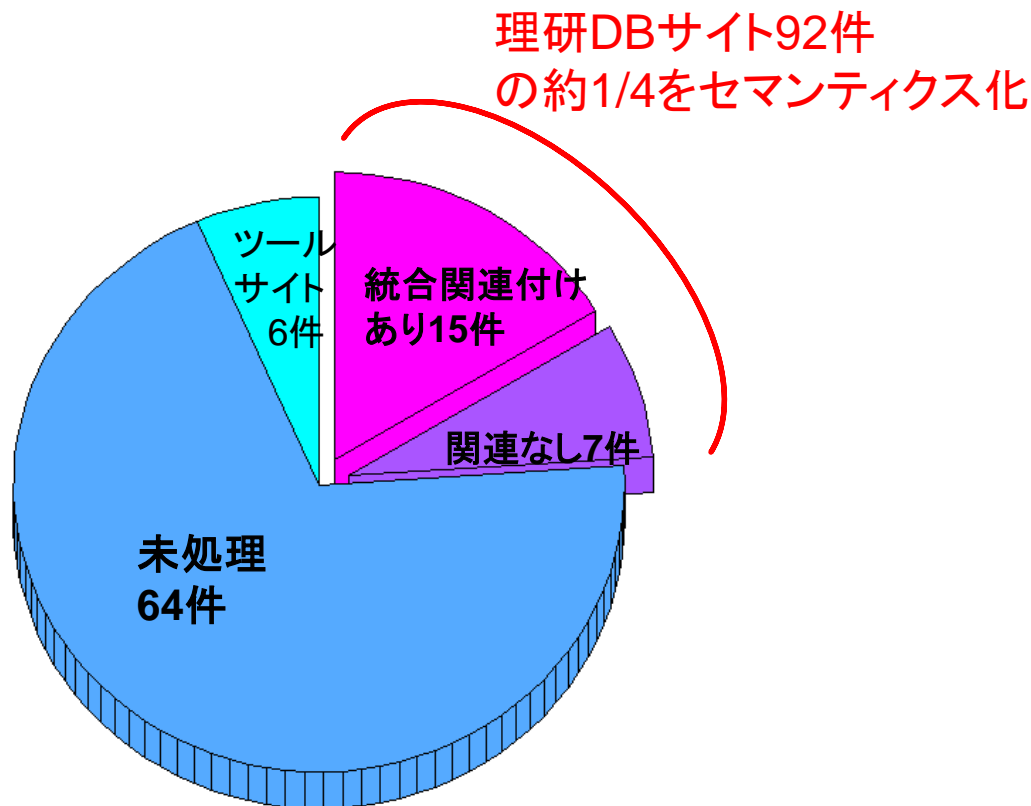
年度目標

④アノテーションシステムの開発運用と
変換データの中核機関への提供(契約締結後)

12月時点進捗

④アノテーションシステムを使って上記以外の理研データベースの統合化を推進している。一方、外部公開用サーバについては、本委託事業費に計上していなかったため、理研内の仮サーバを臨時に用いて、年度内には公開化を完了する。

◆植物以外の理研DBの統合化準備



◆メタデータ登録の準備

④投稿するデータの作成に関するヘルプ

データ名を英語で入力して下さい *入力必須
bacpedia

データ名を日本語で入力して下さい *入力必須
bacpedia

外部URL(英語)を入力して下さい *入力必須

1.1 責任者を選択して下さい *入力必須

ムーア エイリアン
中尾 愛子
新海 暁男
高島 明彦
後藤 彰
中野 明彦

ヘルプ

1.2 連絡先のメールアドレスを入力してください。
kunisima@spring8.or.jp

1.3 担当機関を選択してください(複数選択可)。 *入力必須
(独)理化学研究所 ヘルプ

1.4 連絡先住所を入力してください。
〒679-5148 兵庫県佐用郡佐用町光部1-1-1

2. 該当するデータベース分類を選択してください。複数選択が可能です。「その他(自由記述)」を選択した場合は、具体的に記入してください。

ヘルプ
右画面からデータを選択するか検索して下さい。

その他の分子生物学データベース ヘルプ

3. データベースの要約を入力してください。

キュレーション進捗

■ 公開ターゲット(暫定)

- 結晶構造系
- 高等動植物由来タンパク質

■ データ収集状況

- ターゲット配列: 52件
- **発現データ: 30%**
- **精製データ: 30%**
- **結晶化データ: 42%**
- **X線解析データ: 55%**
- **PDB登録数: 61件**

- Locus Tag
 - Genome ProteinID
 - Organism
 - Product Name
 - Related Locus Tag
- Plasmid Exp. Record
 - Host Cell
 - Vector
- Culture Exp. Record
 - Culture Type
- Purification Exp. Record
 - Purification Method Record
- Sample Exp. Record
 - Sample Reagent Formulati on
- Crystallization Exp. Record
 - Plate Observation
 - Well Condition
- Diffraction Exp. Record
 - Diffraction Image
- Bacpedia Protein Structure Analysis Record

Copyright © 2009 RIKEN, Japan. All rights reserved.

詳細画面

ダウンロード

播磨微生物関連実験情報トップページ

微生物由来タンパク質のX線結晶構造解析実験情報および
変異導入タンパク質の結晶構造の均一で詳細な情報

平成21年度業務計画 (国島)

微生物由来蛋白質構造のアノテーション

平成20年度に公開したデータについてアノテーション作業を行い、アノテーションデータを公開する。またそれ以外のデータについても国際的な合意を得つつデータ公開を進める。

1. データの追加及び更新
2. 公開したデータについてアノテーション作業

A)データの回収
[データの拡散を防ぐ]

B)データの整理
[データ間の紐づけや情報の修正]
[情報の品質を高める]

C)アノテーション
[情報の価値を高める]

D)データの公開
[理研のデータ公開
基盤から公開]

①微生物由来タンパク質のX線結晶構造解析実験情報と立体構造データに関する体系的データの提供

平成20年度公開 → 発現プラスミド構築実験:1万件、培養実験:5千件、精製実験:3千件、結晶化実験データ(条件:約90万件、観察約1千万)、回折実験データ:200件

A)完了 B)ほぼ完了 C)今年度公開予定分のデータは登録済み。今後、今年度最終版のデータと差し替える(テキストデータ)。画像類は今年度最終版のデータ登録を既に行っている。

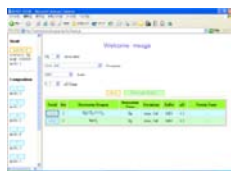
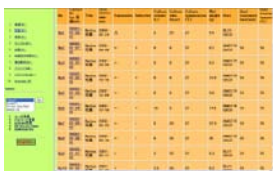
②微生物由来変異導入タンパク質の結晶構造に関する均一で詳細なデータの提供 平成20年度公開 → 150件
進捗は①と同じ

③重原子実験情報の提供 平成20年度公開 → 500件

A)完了 B)完了 C)平成20年度公開最終データ登録は完了し、公開に向けて微調整を行っている。

タンパク質結晶構造解析研究グループ(国島)

①微生物関連実験情報 ②変異体構造情報 ③重原子実験情報



分子機能情報

ローカル環境にあるデータ

ローカル環境にあるデータ

①微生物実験情報
放射光システム生物学研究グループ(倉光)

クレンジング等が完了したデータ

提供

[Omixml形式]

各データアイテムに
統一的なID



Riken Semantic Web 基盤(豊田)

