Licensed under a Creative Commons 表示 2.1 日本 license ©2010 ライフサイエンス統合データベースセンター

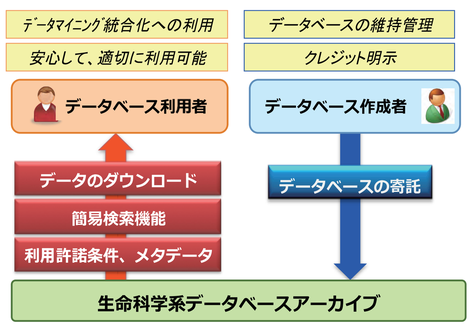
文部科学省委託研究開発事業 「統合データベースプロジェクト（Integrated Database Project）」

「生命科学系データベースアーカイブの成果と今後」

○坂東 明日佳，熊谷 禎洋，八塚 茂，西川 哲夫，畠中 秀樹，大久保 公策，高木 利久
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
bando@dbcls.rois.ac.jp

「生命科学系データベースアーカイブ」とは？

あのデータベースがまるごとダウンロード可能に



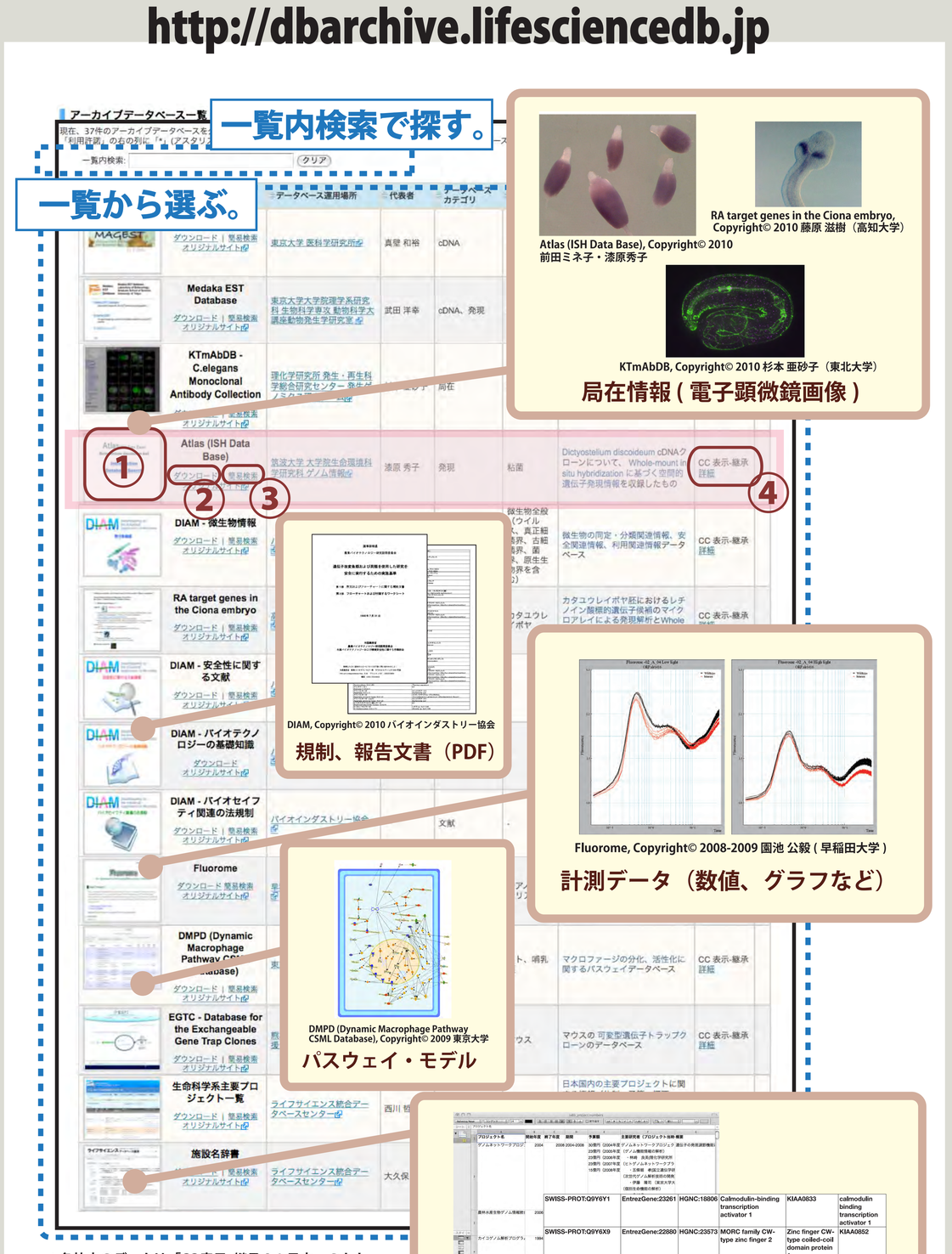
国内で生産された生命科学系 DB を誰もが利用できるよう、ダウンロード可能なかたちで長期にわたり維持していくサービスである。

【要旨】

「生命科学系データベースアーカイブ」では、公的資金により産生されたデータを将来にわたり誰もが容易に活用できるよう、データベースを丸ごとダウンロード可能なかたちで維持している。2010 年 10 月 5 日現在までに 37 件のデータベースの受け入れ・公開を完了した。今後は文部科学省所管プロジェクト公募要領でのデータ提供の明文化や省庁の枠を超えたデータ共有の動きなどにも対応しながら受け入れ体制を強化していく。各データベースの公開形式についてはデータの一括ダウンロードページに加えて、データの説明（メタデータ）や利用許諾条件の提示、さらにデータの一覧や部分ダウンロードも可能な簡易検索ページを立ち上げてきた。今後はデータ生産者によるアーカイブ作成の支援に向けて、これまでのアーカイブ化の試行錯誤をガイドラインとしてまとめ、データ生産者の支援に活用することを検討している。

【成果】37 件の多様な DB の受け入れ・公開が完了

http://dbarchive.lifesciencedb.jp



＊各枠内のデータは「CC 表示-継承 2.1 日本」のもと配布されています。

一覧や検索から、目的のデータベースを探す。

【成果】DB のアーカイブ化の手順の構築

① データベースの説明、データ項目の説明（メタデータ）

名称	Atlas (ISH Data Base)
別名	
作成者（所属）	前田ミネ子（作成時: 大阪大学）、*漆原秀子(筑波大学)
データベース運用場所	筑波大学 大学院生命環境科学研究科 ゲノム情報
連絡先	〒305-8572 茨城県つくば市天王台1-1-1・・・
オリジナルサイト	http://dictycdb.biol.tsukuba.ac.jp/~tools/bin/ISH/index.html
一覧表示	無し
クエリ検索	有り
Webサービス	無し
ユーザ登録	無し
説明	Dictyostelium discoideum cDNAクローンについて、 Whole-mount in situ hybridization に基づく空間的遺伝子発現情報を収録したもの。
生物種	Dictyostelium discoideum(Taxonomy ID: 44689)
運用開始年月日	2003年10月1日
最終更新年月日	2004年4月2日
背景・プロジェクト	文部科学省特定領域研究　ゲノム研究
データベースの特長・有用性・活用方法	シンプルな発生・分化をすることでモデル生物として利用されているDictyostelium discoideumの発生期に発現する遺伝子について局在データを示すことにより、それらの遺伝子機能について推察することができる。 また、部位特異的な遺伝子についてDicty_cdBへのリンクから遺伝子の構造とアノテーション情報を得ることができる。
論文等	Changing patterns of gene expression in・・・

＊一部項目抜粋

データベースの説明

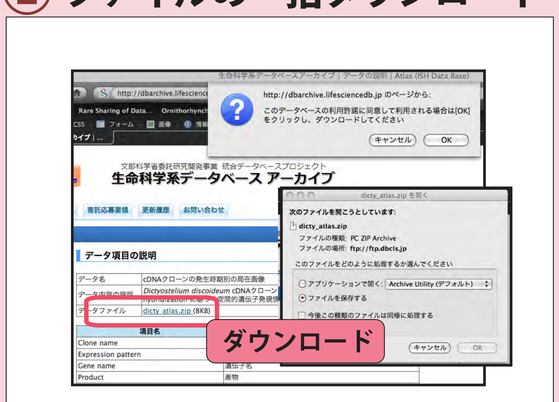
項目名	項目の説明
Clone name	クローン名
Expression pattern	発現パターン（prespore(予定胞子) / prestalk(予定柄細胞)）
Gene name	遺伝子名
Product	産物
Accession number	アクセッションナンバー
mound	発生時期「mound」の局在画像
tipped mound	発生時期「tipped mound」の局在画像
standing slug	発生時期「standing slug」の局在画像
young slug	発生時期「young slug」の局在画像
migrating slug	発生時期「migrating slug」の局在画像
slug	発生時期「slug」の局在画像
mexican hat	発生時期「mexican hat」の局在画像
culminant	発生時期「culminant」の局在画像
mid culminant	発生時期「mid culminant」の局在画像
late culminant	発生時期「late culminant」の局在画像
very late culminant	発生時期「very late culminant」の局在画像
fruiting body	発生時期「fruiting body」の局在画像

データの項目の説明

フォーマットの整った説明つきで、DB 全体及びダウンロードデータの内容を把握できる。

DB の説明では、MEDALS（経産省）と連携して項目の決定・改訂を行っている。

② ファイルの一括ダウンロード



・一括でデータをダウンロード。

③ 簡易検索サイト（TogoDB）

・データの閲覧。
・検索して目的のデータのみダウンロード。

寄託されたデータをもとに、ダウンロードファイルと、TogoDB（データベーステンプレート）を用いた簡易検索サイトを作成。

Atlas (ISH Data Base), Copyright© 2010 前田ミネ子・漆原秀子

④ 利用許諾

Atlas (ISH Data Base), Copyright© 2010 前田ミネ子・漆原秀子　licensed by CC 表示ー継承 2.1 日本

データベース作成者の氏名や所属をクレジットとして掲載する。

「CC 表示 - 継承」を標準利用許諾として採用。

【今後】プロジェクトの進行と連携した公開作業へ

● 公募要領でのデータ提供の明文化との連動

統合データベースへのデータ提供

本プログラムはライフサイエンス分野のデータ産生を伴うプロジェクトであるため、本プログラムの実施者は、ライフサイエンス分野における研究に資するよう、本プログラムの実施の結果産生したデータ又はデータベースを、一般に利用可能な形で公開するべく努めるものとします。そのため、本プログラムの実施者は、毎事業年度にまとめる成果報告書の提出に合わせて、論文発表等で公表された成果に関わる生データの複製物、又は本プログラムで構築した公開用データベースの複製物を、ライフサイエンス統合データベースセンター※又は文部科学省の指定する機関に提供するものとします。提供された複製物については、ライフサイエンス分野のデータベース整備（統合データベース構築）に利用するため、これら複製物に関わる知的財産権を非独占的に複製・改変その他必要な形で利用できるものとします。（後略）

平成 21 年度
・NBRP
・脳科学研究戦略推進
・橋渡し研究支援推進
・ターゲットタンパク
・セルイノベーション
の公募要領での記載

● 省庁の枠を超えたデータベースのアーカイブ作成への動き

【今後】データ生産者自身によるアーカイブ作成の支援へ

● ①～④ をデータ生産者自身が実行できるよう、我々のアーカイブ化の経験をもとにしたガイドライン作りを検討中。

● ガイドラインを利用したアーカイブ作成支援の実践。

DBCLS Database Center for Life Science
大学共同利用法人　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
〒113-0032　東京都文京区弥生 2-11-16　東京大学工学部 12 号館　TEL:03-5841-6754　FAX:03-5841-8090