



# 地球規模生物多様性情報機構 (GBIF) 日本のデータポータルサイトは何を提供しているか

重元 康昌<sup>1,2)</sup>, 桑名 良和<sup>2)</sup>, 菅原 秀明<sup>2)</sup>  
<sup>1</sup>富士通株式会社, <sup>2</sup>国立遺伝学研究所

Contact us: gbif-d@genes.nig.ac.jp

## ◆GBIFの概要

- 地球規模生物多様性情報機構【Global Biodiversity Information Facility (GBIF)】は2011年に180万種10億件の生物多様性情報(標本・観察データ)の提供を目標としている。
- 各国の拠出金で運営され、各国ならびにNGOはそれぞれの大学などの機関(ノード)を介してGBIF標準のデータ形式と通信方式で生物多様性データを公開している。
- 2009年6月現在の統計

データ提供機関(大学や博物館など)	289機関/39の国または地域
生物多様性情報(標本・観察データ)の件数	約1億7000万件

- 国別の標本・観察データの登録数(多い順)

国名	データ提供機関数	標本・観察データ数
アメリカ	78	6400万
スウェーデン	1	1800万
イギリス	8	1700万
フランス	13	1100万
ドイツ	19	660万
オランダ	3	550万
デンマーク	2	360万
ノルウェー	3	340万
オーストリア	7	330万
スペイン	5	290万
...		
<b>日本</b>	<b>4</b>	<b>140万</b>
...		

## ◆GBIFにおける遺伝研の役割

- 国内の標本・観察データを集成し、国際GBIFへ公開。
  - ✓ 東京大学、科学博物館、兵庫人と自然の博物館、理化学研究所などからデータを収集し、国際GBIFへ提供、公開。
  - ✓ マウスや昆虫、植物、乳酸菌の標本・観察データを約51万件を遺伝研から公開。
- 国際GBIFデータの国内での活用を促進。
  - ✓ 年1回のシンポジウム開催
  - ✓ GBIF日本ポータルサイトの公開
    - チュートリアルページの日本語訳
    - 和名による検索
    - バーコードオプライフのデータを用いた生物種同定システム

**GBIFデータ公開を支援します。GBIFフォーマット適合テストはこちら→**

**GBIFデータ公開の受付**

お手持ちの標本や観察結果のデータを国際プロジェクト GBIF を介して発信するお手伝いを致します。GBIFのデータ例やデータの標準や観察データがGBIFの規約に従っているかチェックを行うシステムも公開しておりますので、こちらからご利用ください。

**GBIF検索**

フリーキーワード検索(生物種名/国名/データセット名)  Search

和名によるGBIF検索: 例 ツバメ

和名によるGBIF検索は、ライフサイエンス統合データベースセンターの生物学名日本語一般名対照辞書を利用して実現しております。(S.Kawamoto et al. in DBCLS for LSDB of Mex)

生物種名で検索: 例 Puma concolor (Linnaeus, 1771)

国名で検索: 例 Japan

データセット名で検索: 例 Flora of Japan Specimen Database (東大植物標本データベース)

検索はコンピュータ上のGBIF本部サーバによって実行されます。

**GBIF検索チュートリアル(日本語)**

GBIFデータポータルサイトの使い方をご紹介します。 **GBIF検索のチュートリアルページ(日本語訳)**

<http://gbif.ddbj.nig.ac.jp/>

## ◆和名による標本・観察データの検索

- ライフサイエンス統合データベースセンターの「**生物学名日本語一般名対照辞書**」を利用して、和名から学名への変換機能を実現。
- 国際GBIFデータポータルでは和名を受け付けていない。
- GBIF日本ポータルで和名から学名へ変換後に標本・観察データを検索する。

和名「ツバメ」で検索した例

入力された「ツバメ」は、生物学名日本語一般名対照辞書を用いて以下の学名または英名に変換されました。下記のいずれかをクリックしてください。

**Hirundo rustica** **クリック**

Hirundo rustica gutturalis

**国際GBIFデータポータル**

**検索結果**

Search Results for: **Hirundo rustica**

Scientific names Common names Countries Datasets

**クリック**

Species **Hirundo rustica** (English: Barn Swallow)

Subspecies **Hirundo rustica erythrogaster**, Animalia

Subspecies **Hirundo rustica erythrogaster**, Animalia

Subspecies **Hirundo rustica erythrogaster**, Animalia

Subspecies **Hirundo rustica erythrogaster**, Animalia

Subspecies **Hirundo rustica gutturalis**, Animalia

Notospecies **Hirundo rustica gutturalis x tyleri**, Animalia

Notospecies **Hirundo rustica gutturalis x tyleri**, Animalia

Subspecies **Hirundo rustica gutturalis**, Animalia

Subspecies **Hirundo rustica mesoleuca**, Animalia

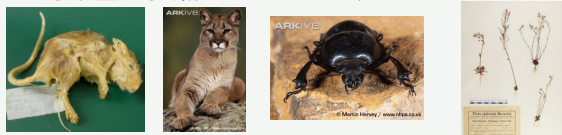
View all scientific names matching "Hirundo rustica".

**ツバメの観察データ分布図**

任意のDNA配列

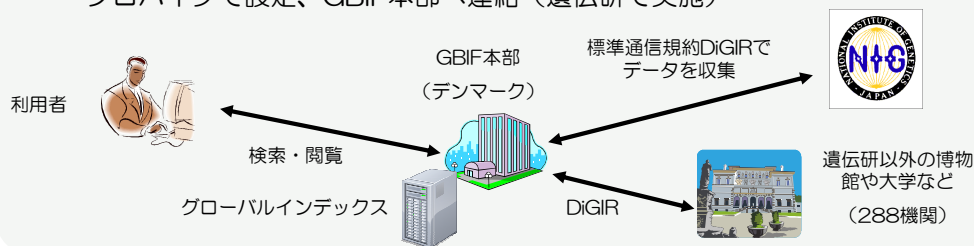
## ◆GBIFで公開している標本・観察データの例

- ✓ 学名 Mus musculus Linnaeus, 1758 (ハツカネズミ)
- ✓ 分類情報 Kingdom: Animalia, Phylum: Chordata, ...
- ✓ 一般名 House Mouse (英語) (各国の一般名が公開されています。)  
Souris Commune (フランス語)
- ✓ 標本の採集年月日、または観察年月日
- ✓ 標本の採集地、または観察地(緯度・経度、国、県、群や区)
- ✓ 標本・観察データの分布表示(標本データが多い地域を色付けして表示)  
(左下のツバメの観察データ分布図を参照してください)
- ✓ 標本・観察データの画像



## ◆GBIF公開手順

- お手持ちの標本・観察データについて、GBIFで採用しているデータ標準 Darwin Core との関連付けを行う。  
Darwin Core 48項目の例 (全ての項目が埋まっている必要はありません。)
- ✓ カタログ番号(例: C1-119467)
- ✓ 大学や博物館のコード(例: 兵庫県博の場合: HYO)
- ✓ 種別(観察/標本/遺伝資源、種)
- ✓ 学名や分類(界、門、綱、目、科、属、種、亜種)
- ✓ 収集もしくは採取した年月日
- ✓ 採集地(国、州・県、群・区、緯度・経度)
- Darwin Coreに沿った Excel もしくはタブ区切り形式ファイルのご提供 (ご相談ください。)
- 標本・観察データファイルのチェック
  - ✓ 必須項目に漏れが無い
  - ✓ 規定の型に合致しているか(緯度・経度は数値でなければならない)
  - ✓ カタログ番号がユニークであるか
- 標本・観察データ全体概要情報(Abstract, Keywords, 管理者)のご提供
- 遺伝研のサーバにデータを登録し、GBIF本部との通信ソフトウェア DiGIR プロバイダで設定、GBIF本部へ連絡(遺伝研で実施)



## ◆有用サイト集

- ✓ 収録サイトにタグ付与
- ✓ 収録サイトをカテゴリーに分類
- ✓ 収録サイトにコメント付与
- ✓ 収録サイトのサムネイル自動収集
- ✓ 収録サイトに対するコメント受付
- ✓ 収録サイトのランキング

トップ GBIF公開支援 国内活動 チュートリアル Barcode同定 有用サイト 国際GBIFサイト

有用サイト集

1. 生物多様性情報システム Global Earth Observation System of Systems (GEOSOS)

2. 環境生物多様性センター (EBC)

3. 生物多様性情報システム (Bioinformatics)

4. Surface Ocean - Lower Atmosphere Study (SOLAS)

5. Site CUI12000

## ◆Barcode of Life (BoL) データを用いた生物種同定システム

- BoLプロジェクトでは、Cytochrome Oxidase Subunit 1 (COXI) や、Internal Transcribed Spacer Region (ITS) のような特定遺伝子の配列の一部を生物種識別のバーコードとしてデータベースに蓄積している。
- New Yorkの女子高生2人がレストランやスーパーの魚の切り身を集め、大学に届けてBoL DBで解析したところ、56点のうち14点が「偽」魚と判明し、New York Timesなどに大きく報道された。その後、学術誌にも論文発表された。
- このBoLのデータやDBJの16S rRNAのデータを用いた生物種同定システムもGBIF日本ノードから公開している。

バーコードオプライフデータを用いた生物種同定システム

指定した参照データベースバーコードオプライフデータベース

比較対象データ

生物種同定結果のサマリ

taxonomic level (axonomic assignment) Probability (%)	
phylum	Arthropoda 100.00
class	Insecta 100.00
order	Lepidoptera 100.00
family	Nymphalidae 100.00
genus	Agraulis 100.00
species	Agraulis vanillae 100.00

種同定結果

任意のDNA配列

系統樹による確認