

平成13年度
生物多様性情報データベース構築
フィジビリティ・スタディ(F S)報告書

報告日：2002年3月4日

グループ名：特殊な生理機能を有する海産動物の
データベースの開発

グループ代表者氏名：道端 齊

1. 構築検討したデータベースについて

<p>データベースの概要及び特徴</p>	<p>陸上生物には例を見ない特殊な生理・生命維持機構を有するいくつかの海産動物種に焦点を当て、それらの分類や生態をより明らかにするとともに、その生命現象の発現に必須な蛋白や糖鎖などの分子生物学的知見をも加えた包括的データベースを構築し、将来のバイオインフォマティクスへとつなげるデータベース化をめざす。特殊な生理現象を生じさせた機構こそが多様性を生み出した源であると考え、特殊性の解明によってより普遍性の高い生命観を導き出すことを目的の一つとする。</p> <p>本年度のFSでは、重金属のバナジウムを濃縮するホヤ類、特異な受精様式を示すヒトデ類、共生する化学合成細菌からエネルギーを得るヒゲムシ類を対象として、分類学から生理学・分子生物学に至るデータの一部を収集した。</p>
<p>類似のデータベース</p>	<p>国外では、ホヤ類の全分野の研究論文に関するデータベースとして Lambert 教授夫妻による http://nsm.fullerton.edu/~lamberts/ascidian/、ヒトデの包括的なデータベースとしてスミソニアン博物館の http://www.nmnh.si.edu/gophermenus/TheEchinodermNewsletter.html、多様な海産動物種に関する Alvin 号のホームページ http://www.ocean.udel.edu/deepsea/home/home.html、無脊椎動物の酸素輸送蛋白質に関するデータベース http://www.sb-roscoff.fr/I02BiP/がある。</p> <p>国内では、日本海洋科学技術センター（JAMSTIC）http://www.jamstec.go.jp/が深海生物の総合的データベースを有する。JAMSTICではヒゲムシに近縁なハオリムシを初めとして深海生物の包括的データベースを作成している。</p>
<p>生物分野</p>	<p>現代生物学は主に陸上生物をモデルとして得られた知識を普遍化して成り立っている。しかし、陸上動物はわずか5動物門に属するのに対し、多種多様な海産動物種は化石種を含めれば35動物門のすべてに出現し、それらの中には特殊な生理・生命維持機能を具有する動物群が数多く存在する。従って、これらの特殊な生理現象を有する海産動物の生理現象を明らかにすることは、陸上生物の研究から得られた生物学的常識を革新的にパラダイムシフトする可能性があり、本データベースの開発は大きな意義を有する。</p> <p>本提案により構築されたデータベースは、海洋生物学、系統進化学、形態学、生理学、生化学、分子生物学、微生物学、構造生物学等の全生物学分野に利用されるのみならず、天然物化学、結晶構造化学、錯体化学、生物無機化学等にも波及的效果がある。</p>

2. データ源およびデータの現在の状況について

<p>データ源</p>	<p>1) H13年度データ取得可能なデータ源(件数など): 多様な生理機能に関する自機関の論文 6 件および他機関の論文 17 件の合計 23 編の論文にもとづくデータを 551 件入力した。3 月中に他機関の論文をもとに数十件のデータを追加する予定である。</p> <p>2) H14年度データ取得が可能なデータ源の見込み(件数など): ホヤ類の分類については、時岡の「相模湾産海鞘類図譜」と西川の「日本海のホヤ類」のモノグラフをあわせて入力した日本周辺産約 150 種の形態と分布の情報に富むデータベース http://etij.c.shikoku-u.ac.jp があり、多数の標本が名古屋大学に所蔵されている。これらに基づいてホヤのデータを整備する。</p> <p>3) H15年度以降のデータ取得が可能なデータ源の見込み(件数など): ヒトデの分類に関しては、「相模湾産海星類」が生物学御研究所から、また日本資源保護協会から「日本陸棚周辺の棘皮動物」が刊行されている。これらの標本の多くは国立科学博物館と生物学御研究所に所蔵されている。 ヒゲムシ類に関しては、約 100 種が A.V. Ivanov (1963) のモノグラフに記載されており、その本は金沢大学理学部附属臨海実験所が所蔵する。また、標本は国立科学博物館と同臨海実験所が所蔵する。 これらを順次入力していきたい。</p>
<p>データの発生・収集場所</p>	<p>自機関の論文 6 件および他機関の論文 17 件の合計 23 編の論文にもとづくデータを入力した。生物種数として 79 種、データ数として 551 件入力した。3 月中に他機関の論文をもとに数十件のデータを追加する予定である。</p>
<p>他機関の場合の連携協力について</p>	<p>日本動物学会では、地球上に現存する一億種を越えるとも言われる多様な生命体のすべてを包括的に記載し、ゲノム DNA、細胞、配偶子を保存解析する「ガイアリスト 21」プロジェクトを提案している。 本データベースは、特異な生理機能を有する海産動物のホヤ類、ヒトデ類、ヒゲムシ類とそれぞれの類縁の動物群に絞り込んで、「ガイアリスト 21」プロジェクトを意識し、分類・生態学のみならずタンパク質・糖質とそれらの遺伝子のデータを記載し、それらの相関を調べるバイオインフォマティクスの基礎的データベースを目指す。</p>

<p>データフォーマット</p>	<p>データベース構成としては、多様な生理機能を統一したフォーマットで平坦に記述した「付加情報(仮称)」データベースと、関連する種の分類・生態情報を記述した「種(仮称)」データベースを軸とし、それらに付随する「文献」「シノニム」「ホモニム」データベースを整備した。また、データの中で写真や図版は一括して「写真」データベースに保存した。ただしデータベースにはプレビュー用の 500x500 ドットのデータ(JPEG)のみとし、高解像度のデータ(TIFF または JPEG)は別に保管することとした。これらのデータベースのデータは全て独立の ID を付与して管理することとした。</p> <p>データ作成・管理には Windows および Macintosh 上で動作するファイルメーカー Pro を用いた。</p>
<p>デジタル化されたデータについて</p>	<p>1) 現在保有するデータ総件数と保存媒体 (H13 年度末見込み)</p> <p>特異な生理機能のうち「生体元素」に関するデータを 393 件、「受精」に関するデータを 87 件、「酸素運搬」に関するデータを 3 件を入力した。またそれらに付随する種の情報を 79 件、文献を 25 件、写真・図版を 43 件入力した。最新版は広島大学大学院理学研究科附属臨海実験所のコンピュータ内に保存しており、メンバーに対しては随時 CD-R に記録したものを配布している。</p> <p>2) 平成 14 年度の見込み</p> <p>ホヤ類の分類について日本周辺産約 150 種の分類学的データおよび標本データを入力する予定である。</p> <p>3) 平成 15 年度以降の見込み</p> <p>ヒトデの分類に関して「相模湾産海星類」および「日本陸棚周辺の棘皮動物」のデータ、ヒゲムシ類に関しては約 100 種のデータを入力する予定である。</p>
<p>データ・ベースの実現方式とデータのクオリティ</p>	<p>データベースはパソコン上で動作するファイルメーカー Pro で作成・管理している。このシステムの利点は特に専用のハードを必要としないこと、専門家でなくともある程度のデータベースが作成できること、維持管理が容易であることが挙げられる。しかしその反面、データ保護・履歴機能は貧弱であり、入力者の注意努力に依存するところが大きい。</p> <p>データ自体の信頼性は、元となる文献・書籍を電子化した後にも紙媒体として保存しておき、随時チェックする必要がある。</p>

3 . FS で得たデータベース化する際の知見について

データ・フォーマットやデータの加工内容などについて	<p>データにはすべて ID を付け、ID に基づいてデータベース相互間で参照することとした。特に図版は独立したデータベースに保存し、各データベースにはその ID のみを記録することとした。</p> <p>問題点として、日本の利用者の利便を考えると日本語と英語の両方を入力の方が好ましいが、データベースが複雑化し入力コストも上昇することが挙げられる。現時点では主として日本語で入力している。将来的に可能であれば、相互に自動翻訳するシステムを内蔵することが望ましい。</p>
---------------------------	--

4 . その他

データベースを公開する上での問題点の解決について	<p>一般に投稿した論文の著作権はその科学雑誌に、また特許はその権者に帰属するので、公開に当たってはその都度許可が必要となる。特に図版を利用する場合、問題になる。この問題は、GBIF プロジェクト全体で国際的に解決すべき問題であろう。</p> <p>またデータの公開に当たっては、外部からの不正アクセスや内容の書き換えが出来ないようにする必要がある。</p> <p>現在用いているファイルメーカーPro は簡易公開機能を備えているため常に最新の情報を公開することが可能だが、データの保護という点を考慮すると、公開時には静的なデータに変換した後、http サーバーを用いて公開するのが良いと考える。当面は、非公開で運用する。</p>
その他	<p>分類・生態のみならず、生理学から分子生物学・生化学に至る包括的データベースへの道筋を見いだした。同時に、多くの問題点を洗い出すことが出来た。</p>

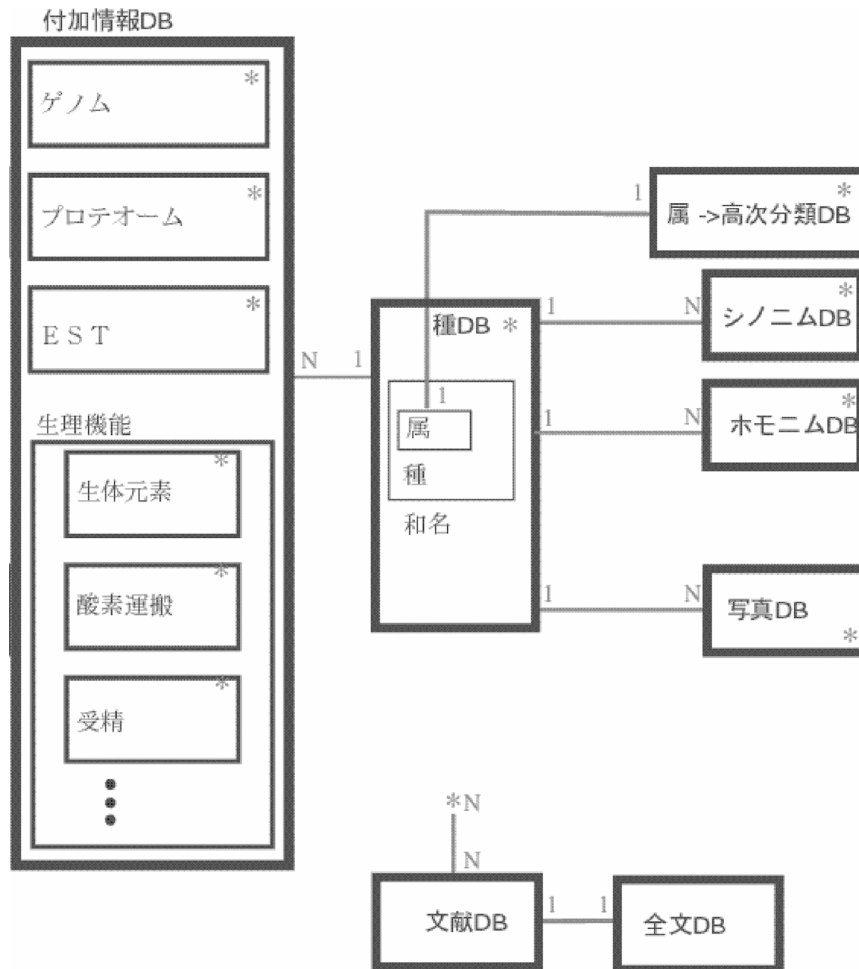
5 . 確立できた推進体制

開発責任者	道端 齊 広島大学大学院理学研究科附属臨海実験所・教授
研究協力者	笹山雄一 金沢大学理学部附属臨海実験所・教授 福森義宏 金沢大学理学部・教授 松野あきら 島根大学生物資源科学部・教授 西川輝昭 名古屋大学博物館・教授 宇山太郎 広島大学大学院理学研究科附属臨海実験所・講師 植木龍也 広島大学大学院理学研究科附属臨海実験所・助手 松本 緑 慶應大学理工学部・講師

6 . FS 後の推進スケジュール（案）

	H 14	H 15	H 16	H 17
1.データベース基本設計	■			
2.データベース詳細設計		■		
3.情報機器の導入		■	■	■
4.データベースプログラミング			■	
5.データ作成・入力	■	■	■	■
6.運用試験				■
7.試験公開				■
8.公開				■

(別紙) 本グループで作成したデータベースの概念図



- ・種および高次分類データベース
- ・高次分類データベース
- ・シノニムデータベース
- ・ホモニムデータベース

以上の4つのデータベースにより個々の生物種の分類、形態、生態、地理的分布といった情報をまとめる。

- ・付加情報データベース(仮称)

特殊な生理機能およびゲノムサイエンスの情報をまとめる。

- ・写真データベース
- ・文献データベース

上記の情報に関連する写真と文献をまとめる。