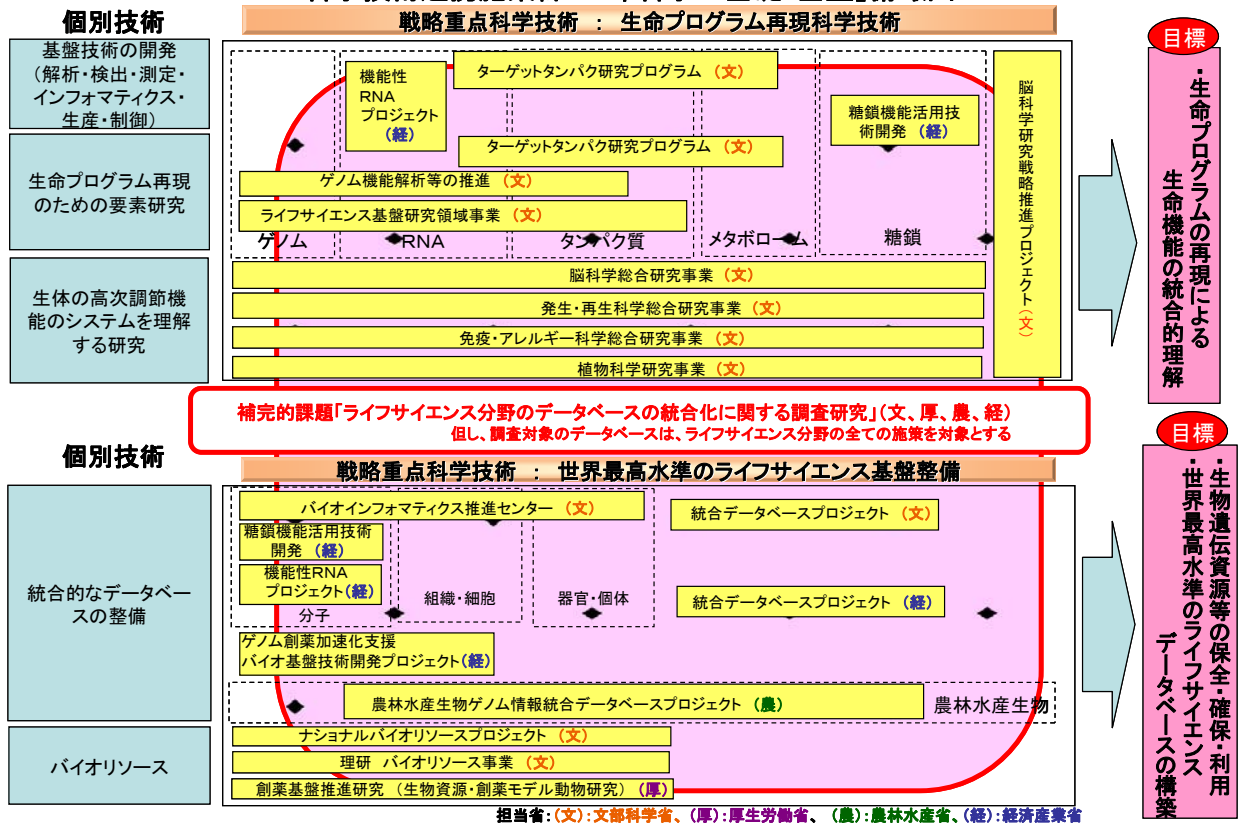


## 科学技術連携施策群「生命科学の基礎・基盤」俯瞰図



## 1. 生命科学の基礎・基盤

### (1) 目標

統合的なデータベース整備に向けた研究開発については、第1期、第2期科学技術基本計画の下に、ヒトを含めた動植物等のゲノム解読を行うプロジェクトや、3000種類のタンパク質基本構成要素の立体構造の決定を目指したタンパク3000プロジェクトに代表される大量のデータが産出される施策が多く行われてきたが、それぞれの研究成果についてプロジェクトごとにデータベースが構築されたためデータベース間の連携がとれておらず、また、利用者の利便性に充分配慮された設計となっていなかった。そのため、多くの資金や労力を使って産み出された研究資産であるデータが活用されにくい状態になっていた。これらのデータを、どのように効率良く研究者、産業界、さらに国民に還元し、新たな知見を得たり、新たな産業を創出していくのかが課題となっており、この問題点の解決に向けた制度的・技術的取組は急がれるところであった。

そのため、本連携施策群では、第3期科学技術基本計画における分野別推進戦略のライフサイエンス分野での戦略重点科学技術の中で、「世界最高水準のライフサイエンス基盤整備」及び「生命プログラム再現科学技術」

- ・ 「世界最高水準のライフサイエンス基盤整備」：産出・蓄積された生命情報データの利活用に必須である統合的なデータベース整備に向けた研究開発及び、質と量の科学的評価を踏まえた生物遺伝資源等の保全や確保を行う領域。
- ・ 「生命プログラムの再現科学技術」：生命体のDNA、RNA、タンパク質等の相互作用の解明及び脳などの生命の高次調節機構を理解し、細胞などの生命機能単位を、計算機を用いて仮想的にまたは部分機能を試験管内でシステムとして再構築し理解する研究領域。

に関する施策を対象として、関係省庁間の連携を強化して研究体制の構築を行うべく活動を行うこととし、特に情報の流通や蓄積の重要性が増大している生命科学の研究にとって急を要する研究基盤開発である“統合的なデータベース整備に向けた研究開発”を主に推進することを目標とした。

具体的には、我が国で取組まれてきたライフサイエンス分野の研究成果としてのデータベースの統合化を促進して利用者が使いやすいデータベースを構築することとし、それに向けて、各省横断的に現在存在するデータベースの基礎調査や活用状況等に関する調査研究を行う。そして、それらの調査結果を基に関係府省におけるデータベース統合化に向けた制度設計やロードマップ作成に向けた検討を行い、これらの大量のデータを産出した施策を新たな知見を得たり、新たな産業の創出に結びつけるための具体的計画を立てることを目標とした。

### (2) 活動

#### ① 府省間等の連携活動

統合的なデータベース整備に向けた研究開発に向けた連携施策群としての活動については、平成17年度及び18年度は「ポストゲノム連携施策群」として行ってきた。しかし、その連携施策群が対象とする領域は、基礎生物学から医学・臨床分野に及ぶ幅広い研究分

野であるため、活動をより専門化し円滑に行えるよう、この連携施策群を「生命科学の基礎・基盤」、「臨床研究・臨床への橋渡し研究」、「食料・生物生産研究」に分割することが適当であると判断し、平成19年度より3群に再編した。これらのうち、本連携施策群においては、“統合的なデータベース整備に向けた研究開発”の活動を引き継ぐとともに、補完的課題「生命科学データベース統合に関する調査研究」を行ってきた。

また、本連携施策群では、関係者の認識統一に向けて、関係府省間の情報交換の円滑化と連携強化を図り、真に利用し易い統合的なデータベース整備を目指し、打合せ会合〔有識者委員及び文部科学省、厚生労働省、農林水産省、経済産業省担当課長等を委員とする会合及び小打合せ会合〔対象とする研究分野を絞り、その分野の専門知識を有する有識者と各省担当課長を委員とする会合〕等を行い、議論を重ねた。さらに、平成19年度には、「生命科学の基礎・基盤連携群シンポジウム ライフサイエンスデータベース統合への取組」と題したシンポジウムを開催して、補完的課題の研究内容の報告、有識者講演及びパネルディスカッションを行い、国民に対しその情報発信を行った。

具体的には平成17年度においては、本連携施策群の登録施策の精査を行い、

- ・ 各施策は閉じた領域にとらわれず、これまでの研究成果を基礎から応用までのライフサイエンス分野全体と連携してゆくこと。
- ・ データベースを基礎から応用にわたる様々なポストゲノム領域の研究と連携させることが、ポストゲノム研究全体を活性化させるために重要であること。

等の課題を抽出した。これらの課題の抽出により、後述する科学技術振興調整費による「補完的課題」が“統合的なデータベース整備”に向けた調査研究と定められた。

第3期科学技術基本計画におけるライフサイエンス分野の推進戦略の検討に資することを視野に入れ、ライフサイエンスにおける重要な研究開発課題として考えられる課題候補群として、「生命情報統合化データベースの構築に関する研究開発」、「研究開発の基礎となる生物遺伝資源等の確保と維持」、「脳や免疫系等の高次複雑制御機構の解明など生命の統合的理解」等に区分・整理し、ライフサイエンス分野の方向性や研究目標に関する検討を行った。これらの検討結果は、分野別推進戦略（平成18年3月28日）の策定に大きく寄与した。

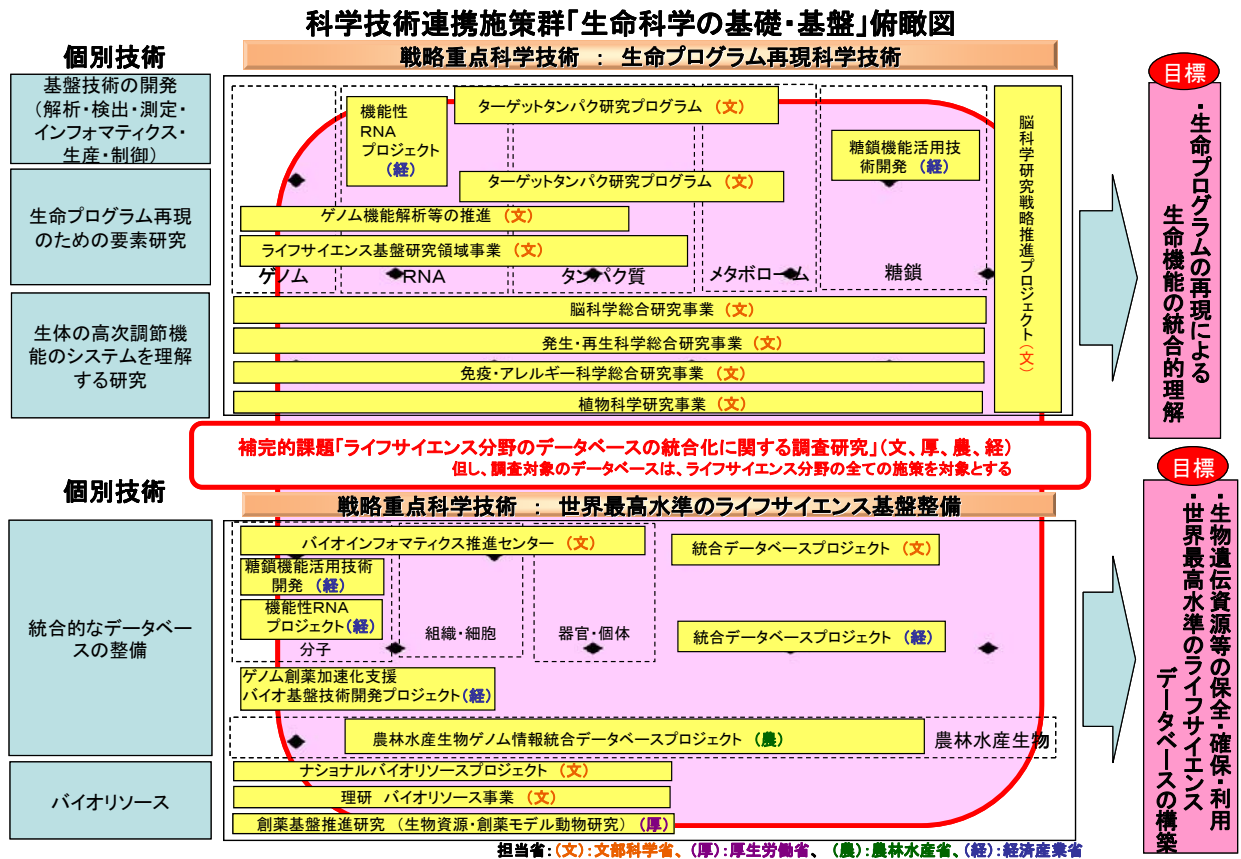
また、本連携施策群の主な目標である“統合的なデータベース整備”に向けた研究を行う補完的課題である「ライフサイエンス分野のデータベースの統合化に関する調査研究」として「生命科学データベース統合に関する調査研究」（研究代表者 大久保公策 教授 国立遺伝学研究所）を採択した。

平成18年度においては、分野別推進戦略におけるライフサイエンス分野の推進を府省連携のもとに実現するため、ライフサイエンス分野に該当する平成18年度施策を精査するとともに、平成19年度概算要求に向けての各府省の取組について情報収集及び意見交換を実施し、戦略重点科学技術の「生命プログラム再現科学技術」及び、「世界最高水準のライフサイエンス基盤整備」についての俯瞰図の素案を作成した。その結果、各戦略重点科学技術における各省施策の個々の位置付けを明確化でき、平成19年度の概算要求に反映させることができた。

また、補完的課題の調査研究の成果として、国内データベースの現状、データベースの構築法による型分類等、データベース統合化の方策のための基礎資料を取得した。その結果、データベースの構築法による型分類から統合に適するデータベースの型が分かり、各省の統合データベース施策の推進に寄与した。

平成19年度においては、改めて対象とする戦略重点科学技術に該当する平成19年度施策を精査、平成20年度概算要求に向けた各府省の取組について情報収集及び意見交換を実施した。また、今後取り組むべき連携強化が必要な課題としてバイオリソース整備を選定し、各省からの関連施策の現状説明及び意見交換を行った。

また、補完的課題については、取りまとめの最終年度であるため、その取りまとめを行った。



## ② 補完的課題の成果概要

### ・ 課題の概要

採択課題名：「生命科学データベース統合に関する調査研究」（平成17年度採択）

内容：ミレニアムプロジェクトや、ポストゲノムプロジェクト等で大規模なデータ産出型の施策が行われてきたが、その成果は広く利用されておらず、実用化研究に十分に活用できる状態にはなっていない。そのため、国内外の主要なデータベースについて内容、機能及び運営方法の調査を行い、データベースをデータおよび機能の両面から多面的に分類整理したうえで有効かつ可能なデータベース統合を検討し、

また、複数のデータベース機能を統合する際の技術的・制度的問題点を検討することを目的に試験的研究を行い、少数のデータベースの統合を試み、統合化の実現性を検証する。さらに、我が国の知的情報基盤の設計に資する客観的な資料並びに国内主要データベースを中心とする統合化に関するフィジビリティスタディーを通じた実現性検証結果の提示を行い、それらの成果を活かした制度設計やロードマップの作成に資する試案を提示する。

## ・ 成果の概要

### 目的)

ゲノムプロジェクトやタンパク3000プロジェクトに代表される大規模データ産出型プロジェクトは、その成果の一部としてデータベースが作成されるが、現在、それらは産学官の多くの研究者に必ずしも効率よく使われていない。また、広く生命科学の分野を見ると日本国内にデータベースが散在しておりその全貌は把握できておらず、それらも効率の良い使い方はなされていない。本研究では、これらの現状を鑑み、日本国内の生命科学データベースを網羅的に調査し、基礎データを得た上で、それらを有効に活用するために障害となる課題を抽出し、その成果を活かした制度設計やロードマップの作成に資する試案を提示することを目的とした。

具体的には、以下の3項目をミッションステートメントとした。

- ①関係府省におけるデータベース統合化に向けた取り組みの補完となる、関係府省の制度設計やロードマップ作成に資する試案の提示
- ②複数の国内主要データベースを統合化する際の技術的課題並びに解決策の提示及びフィジビリティ・スタディーを通じた実現性検証結果の提示
- ③国内外の医学分野・学術分野データベース、国内の産業分野データベースに関する技術的側面、制度的側面からの基礎調査結果の提示

### 成果)

それぞれのミッションステートメントに対して、以下の成果を得た。

- ① 関係府省におけるデータベース統合化に向けた取り組みの補完となる、関係府省の制度設計やロードマップ作成に資する試案の提示

データベース統合の課題は、各データベースに特有の問題を持っているため、今後、統合すべき各省のデータベースの整備に伴って生ずる特有の課題を抽出した上で、海外のよい点を見習いながら、我が国の実情に合わせた柔軟なルール作り（制度設計やロードマップの作成）を検討する必要がある。また、データベースの統合を進めていくにあたり、情報の流通に関して、国民を含めた理解推進に向けた方策も盛り込んでいく必要がある。

② 複数の国内主要データベースを統合化する際の技術的課題並びに解決策の提示及び  
フィージビリティ・スタディを通じた実現性検証結果の提示

フィージビリティ・スタディとして、i)解剖用語の自動分類機による全植物ESTの統合整理、ii)複数のDBにおける遺伝子ID系列の自動更新等を行った。その結果、これらの技術的問題は、情報工学で開発された手法を用いて解決できることが分かった。

③ 国内外の医学分野・学術分野データベース、国内の産業分野データベースに関する技術的側面、制度的側面からの基礎調査結果の提示

生命科学のデータベース全体を整理するために、その構築法からアプローチし分類した。方法としては、情報源の種類、処理方法等の点から分類した。その結果、データベースは、6つの型（バンク型、プロジェクト型、プログラム型、キュレーション型、知識モデル型、総説型）に分類できることが分かった。

(3) 成果と研究目標の進捗状況

補完的課題の実施により、各データベースが利用しづらいこと、使われている用語が不統一であること、データベースの数が多く（250個）、何が、どこにあるかが分かりづらいこと等の現状を把握した。また、データベースの統合が進んでいる米国の場合では、情報の流通が促進される環境が整備されていることが分かった。

これらの成果に基づき、平成18年度に文部科学省「統合データベースプロジェクト」、農林水産省「農林水産生物ゲノム情報統合データベースの構築」の施策が立ち上がり、また、平成20年度予算で経済産業省「統合データベースプロジェクト」が立ち上がることとなった。連携施策群では、これらの施策の推進及びフォローアップを行いつつ、施策間の連携を促進した。

具体的な連携施策群の活動における成果としては、平成17年12月14日にシステム改革専門調査会へ「科学技術連携施策群の活動状況報告ポストゲノム ー健康科学の推進ー 連携施策群ポストゲノムワーキンググループ会合におけるこれまでの検討結果の報告書」を報告した。この報告書の中において、検討対象施策の説明及び検討方法、検討の結果を報告した。その検討の結果の中での補完すべき点として指摘した事項と、その指摘に対する各省の対応は以下のとおりである。

- ① 文部科学省「ナショナルバイオリソースプロジェクト」、経済産業省「製品評価技術基盤機構」、厚生労働省「医薬基盤研究所/ヒューマンサイエンス振興財団」のリソースバンク間の情報共有について、関係研究機関の共同ウェブサイトを強化し相互の情報共有を図ることが重要であると指摘した。それに対し、現在、例えば、文部科学省「ナショナルバイオリソースプロジェクト」のホームページからは、“日本の遺伝資源情報サイト”にリンクが張られ、厚生労働省、農林水産省、経済産業省他の遺伝資源ウェブサイトのリストが閲覧可能となっている。

また、平成17年度、18年度には、補完的課題の中間報告が行われ、そこで以下の指摘等を行い、対応を求めた。

- ② 分かり易いポータルサイト（入り口となるウェブサイト）を作成するべきと指摘した。それに対し、文部科学省「統合データベースプロジェクト」では、情報を収集している物質、データベースの構築法などの複数の分類方法でデータベースを分かり易く分類し、かつ利用者の意見を集約する機能を持たせた辞書的ポータルサイトを作成し運営している。
- ③ 臨床データの統合化データベースの構築を促進するべきと指摘した。それに対し、例えば、文部科学省「統合データベースプロジェクト」では、医科学分野のデータベースを統合するためのデータベース構築方法の開発に着手している。

補完的課題の実施により、平成19年度には複数の国内主要データベースを統合化する際の技術的課題並びに解決策の提示、フィージビリティ・スタディを通じた実現性検証結果の提示、国内外の医学分野・学術分野データベース、国内の産業分野データベースに関する技術的側面、制度的側面からの基礎調査結果の提示をすることができ、また、関係府省のデータベースの統合化に向けた制度設計やロードマップの作成について問題点の抽出を行うことができ、当初の目標を概ね達成することができたが、統合すべき各省のデータベースの整備に伴い、それぞれに特有の課題を解決していく必要がある。そのため、今後、各省におけるデータベースの整備の進捗状況に合わせてデータベース統合化に向けた継続的な取組が必要である。

#### （4）今後の課題

「生命プログラムの再現科学技術」に関しては、今まで基礎データの収集を主眼とした施策が多く行われてきたが、これらのデータを活用し計算機内あるいは試験管内で生命の機能を再現するためには、複数の施策で取得された基礎データを有機的に活用できる統合的なデータベースが必要である。そのため、データベースの統合化の活動と連携させ、本研究領域を推進していくことが重要であると考えられる。

また、「世界最高水準のライフサイエンス基盤整備」の内、バイオリソース（生物資源）の領域について、バイオリソースの継続的維持を行うための方策の検討は急を要する課題である。また、新たな疾患モデル動物の開発を行い創薬の活動に寄与するなどの研究施策を継続して実施し、世界最高水準のバイオリソース基盤を整備していく必要がある。

もう一方の基盤であるデータベースに関しては、これまでの本連携施策群の取組により、文部科学省、農林水産省、経済産業省において統合データベース事業が新たに開始され、それぞれの事業において将来の統合的なデータベース化を目指した、これまでのライフサイエンス分野の研究成果のデータベース化が進められている。

しかしながら、こうしたデータベースの統合化に向けた取組は端緒についたばかりであり、今後、我が国のライフサイエンス分野の研究成果に誰もが容易にアクセスし活用でき、我が国の研究開発力を更に強化させる基盤としての統合データベースを整備していく必要がある。そのためには、有用なデータやデータベースの散逸を防ぎ、新しい情報を入力するなど恒常的に利用者の求める機能を提供できる環境整備や、国として人材が不足してい

るバイオインフォマティクス研究者や技術者の人材養成が重要であり、関係府省の役割分担を明確にしつつ連携強化を図りながら、早急にデータベースの統合化に向けた具体的な制度設計やロードマップを作成し、データベースの統合化を検討し、実施していく必要がある。