

報 告 書

ライフサイエンスデータベースの
統合・維持・運用の在り方

平成21年1月13日

科学技術・学術審議会

研究計画・評価分科会

ライフサイエンス委員会

ライフサイエンス情報基盤整備作業部会

目 次

1. DB整備の意義および必要性	2
2. 現状及び問題点	3
3. 諸外国の状況	4
4. DBの統合・維持・運用の在り方 ～機能～	4
5. DBの統合・維持・運用の在り方 ～体制～	6
6. DBの統合・維持・運用の在り方 ～移行時期～	8
(参 考)	
これまでの経緯	9
ライフサイエンス情報基盤整備作業部会における審議の過程.....	10
ライフサイエンス情報基盤整備作業部会委員名簿	11
ライフサイエンス情報基盤整備作業部会の設置について	12
(参考資料)	
(1) 我が国のデータベース整備.....	13
(2) 統合データベースプロジェクト.....	22
(3) 我が国のバイオインフォマティクスの現状	47
(4) 統合DBタスクフォース（仮称）の設置について	64
(5) 平成21年度概算要求における科学技術関係施策（ライフサイエンス分野）（継続案件）	65

1. DB整備の意義および必要性

- DNA塩基配列の読み取り速度の飛躍的向上に加えて、タンパク質の立体構造データや遺伝子の発現データも大量に産生する中で、データベース（DB）の活用なしにライフサイエンス研究を行うことは事実上困難になってきている。
- また、現代のライフサイエンス研究は、ゲノム科学のように大量データに基づく点が特徴であり、プロジェクトが完了して多数の論文が公表されるだけでは、生産された大量のデータが活かされないことになる。個々の研究室が論文の基となるデータをDBとして作成しているだけで、広範な公開と自由な利用が確保されないままでは、研究成果の再利用が達成されない。
- 我が国のライフサイエンス研究全体を推進、加速させていくためには、より網羅的かつ精確で付加価値の付いたライフサイエンスDBの整備（統合・維持・運用）が不可欠である。
- ライフサイエンスDBの整備（統合・維持・運用）により、ライフサイエンスという多岐に亘る学問体系を生命のシステムとして再統合・俯瞰することが可能となり、生命の理解や探求が深まり、新たな研究の糸口をもたらす。
- 整備された統合DBは、ライフサイエンス研究のためだけでなく、医療、民間企業、さらには一般社会に還元するための手段としても有効に機能する。学界のみならず産業界からもライフサイエンスDB整備（統合・維持・運用）の更なる強化・充実に対する強い要望が寄せられている。
- そのためには、今後、公開性・透明性・客観性・科学性を担保しつつ、国内でDBを戦略的に整備していくことが重要である。
- ライフサイエンスDB整備（統合・維持・運用）には長い年月を要するとともに、現在の技術進歩は二桁、三桁のデータ量の増加をもたらしており、また研究とは異なる側面を持つことから、その整備にあたっては恒常的な経費により、永続的にかつ柔軟性を持って続ける必要がある。

- ・ また、多種多様なライフサイエンスに関するデータの価値を最大限引き出すためには、ライフサイエンス DB 整備（統合・維持・運用）を通じた現存 DB の統合化やその維持・運用に留まることなく、研究の進展に対応した新たな DB や新しい発想に基づく DB の開発、さらには DB を活用したバイオインフォマティクス研究なども並行して着実に進めることが重要である。

2. 現状及び問題点

<現状>

- ・ 我が国全体のライフサイエンス DB 整備（統合・維持・運用）に当たっては、総合科学技術会議の指導の下、関係省が連携して推進することが重要である。
- ・ 文部科学省においては、「統合 DB プロジェクト」として、中核機関である大学共同利用機関法人情報・システム研究機構（ROIS）を中心にライフサイエンス DB 整備（統合・維持・運用）を担っている。
- ・ 一方、ライフサイエンス分野の基盤的 DB や新たな DB の開発を推進するためのファンディングを独立行政法人科学技術振興機構バイオインフォマティクス推進センター（JST-BIRD）が担っている。

<問題点>

- ・ 研究のロジスティックス（基盤整備）を維持するための、特殊なファンディング制度や、グラントの間接経費の一定率を自動的に集金するといったシステムが存在していない。
- ・ 研究の結果生じるデータの共有化に向けた意識が、研究者間で十分とは言えず、プロジェクト終了時に発生したデータを登録し、他の研究者や利用者の自由なアクセスを保証するような共有化推進に向けたルールも存在していない。
- ・ 統合化を担う中核機関のイニシアティブが発揮できるような状況（予算、権

限) に必ずしもなっていない。

- ・ 欧米の状況と比較すると、我が国のライフサイエンス DB 整備（統合・維持・運用）に対する予算規模が小さい（「統合 DB プロジェクト」の予算規模は 11 億円（平成 20 年度））。
- ・ 競争的資金によるプロジェクト毎に様々な DB が開発・運用されているが、恒久的な DB 維持のための安定的な財政措置が行われておらず、我が国の研究成果が生きないだけでなく、日本人固有のゲノム情報・疾患や薬品副作用などの情報が活かされていない。

3. 諸外国の状況

- ・ 米国では、1988 年に NIH (National Institute of Health、米国立衛生研究所) の NLM (National Library of Medicine、米国立医学図書館) の一部門として設立された NCBI (National Center of Biotechnology Information、国立バイオテクノロジー情報センター) を中心とした DB の集中管理を目指して、多額の予算・人材（予算規模は 85 億円、人員規模は約 400 名¹）を投入している。
- ・ 欧州では、1992 年に EMBL (European Molecular Biology Laboratory、欧州分子生物学研究所) から発足した EBI (European Bioinformatics Institute、欧州バイオインフォマティクス研究所) を中心として各国の DB を横断的に利用できる活動を開始している（予算規模は 32 億円、人員規模は約 300 名²）。

4. DB の統合・維持・運用の在り方 ～機能～

- ・ DB の開発は各省庁で行うものの、利用の観点から、まずは文部科学省がライフサイエンス DB 整備（統合・維持・運用）を率先して進め、その後の省庁

¹ 「我が国におけるライフサイエンス分野のデータベース整備戦略のあり方について」（平成 18 年 5 月 ライフサイエンス委員会データベース整備戦略作業部会）

² 「我が国におけるライフサイエンス分野のデータベース整備戦略のあり方について」（平成 18 年 5 月 ライフサイエンス委員会データベース整備戦略作業部会）

連携の在り方については、総合科学技術会議における検討を十分に踏まえていくことが必要である。

- ・ ライフサイエンス DB 整備（統合・維持・運用）を担う機関は、DB のカタログ作りや関連 DB の横断的検索にとどまることなく、他機関と連携し種々の DB を将来的には日本全体として一つに集約し、継続的・主体的に運営する機能を持つことが期待される。また、統合化や高度な検索のための技術開発（インデックス、辞書、データフォーマットなどの構築）も着実にを行い、利用者が真に使いやすく新たな知識発見を可能とする統合 DB の構築を目指すことが期待される。さらに分子データのみならず、文献や特許の情報についても併せて統合化を進めていくことが重要である。
- ・ ライフサイエンス DB 整備（統合・維持・運用）に係る事業が軌道に乗るまでの間、DB 開発者が積極的に DB 統合化に協力していくことが重要であり、その支障とならないように、それを担う機関は、自らの生物学的知識の発見につながるような DB 開発と競合する研究開発機能を持たずに、DB 統合化に向けた事業に徹していくことが望ましい。
- ・ DB 統合化に当たっては、統合化に必要な技術開発のみならず、個々の DB やデータの権利関係処理、個人情報秘匿のための措置などの法的、制度的な問題にも同時に取り組んでいくことが要求される。
- ・ DB 開発者に対する支援と DB 利用者に対する支援という 2 つの機能が損なわれずに両立すること、さらには、DB 開発コミュニティと産業界も含めた DB 利用コミュニティのいずれもが積極的に協力していく体制の構築が必要である。
- ・ 統合化を担う機関は、「統合 DB プロジェクト」により培われた成果や実績を継承し、これまでに育成された高度な専門性を備えた人材が効果的に活かされる環境、研究者コミュニティの意見を十分に反映できるような開かれた運営、高度な専門性を備えた人材の継続的な育成といった機能の重要性に留意するとともに、ライフサイエンス DB 整備（統合・維持・運用）に係る事業

に携わる人材の評価システムを構築し、大学等と連携してそのキャリアパスを考えていくことが期待される。

5. DBの統合・維持・運用の在り方 ～体制～

- ・ ライフサイエンス DB 整備（統合・維持・運用）にあたっては、時限的な競争的資金を活用した運用ではなく、海外の事例のように独立的で継続的な機関として、4. で提示した機能を有する我が国のライフサイエンス DB センターを新たに設置し、事業として運用を行うことが考えられる。
- ・ しかし、行財政改革が行われている中、新たな機関等を設立することは、既存の機関の廃止を含め、既存の機関や研究者コミュニティが新しい機関の構想及び具体的計画を作成し、調整を進め、現在の「統合 DB プロジェクト」が終了する平成 22 年度までに実現させることは極めて困難である。
- ・ また一方で、「統合 DB プロジェクト」を有期の事業として 5 年ごとに要求することは、事業が認められて予算を獲得できるかどうかについて不安定であり、継続性に欠けることになる。
- ・ したがって、本作業部会としては、将来的な理想像の実現に向けた積極的な今後の検討を、政策推進の司令塔として先見性と機動性を持って、府省を超えた国家戦略を示す役割を担う総合科学技術会議 に強く期待すると同時に、現時点においては、同会議による「平成 21 年度概算要求における科学技術関係施策の優先度判定等」における指摘『「統合 DB プロジェクト」及び「バイオインフォマティクス推進センター（JST-BIRD）」に関し、「総合科学技術会議としては、最重要課題の一つとして、本事業の成果を重視。恒久的な体制整備に向けて、統合 DB プロジェクトと JST-BIRD との一本化を目標として具体的検討を進めるべきである。』を十分踏まえ、以下のような体制を検討することが適切であると判断した。
 - 「統合 DB プロジェクト」の中核機関である ROIS においてそのプロジェクト推進を担っているライフサイエンス統合 DB センター（ROIS-DBCLS）

と、これまで我が国のライフサイエンス基盤の DB を支え、推進してきた JST-BIRD、さらに DB を有する関係機関とが、コミュニティの意向を踏まえ、それぞれの機関としての特徴を踏まえた役割分担を図り、一体的な運用を行う。

- ▶ そのため、これまで DB 開発者に対する支援を行ってきたファンディング事業としてのノウハウと、文献情報等の DB 利用者に対する支援を行ってきた情報事業としてのノウハウの両方を兼ね備えている JST が新たな組織を設置し、そこで ROIS-DBCLS を始めとする関係機関各々が持つポテンシャルを最大限活かしつつ、柔軟な運用を可能とする仕組みを構築し、DB の統合・維持・運用を図る。
- 新たな組織の具体的な運営の在り方については、今後、広く研究者コミュニティの意見を踏まえつつ設置主体の JST において DB 運営の基本的な内容を定める計画の策定に向けた検討が進められていくことが必要である。この策定に当たっては、広く研究者コミュニティの意見をふまえるとともに、DB は利用されることで一層価値が高まるものであることから、産業界を含めた研究者コミュニティの多くの利用者からのニーズに的確に応えられるように配慮することが重要である。また、現在の「統合 DB プロジェクト」で培われた成果を効果的に引き継ぐことが期待される。
- また、新たな組織が設置された後の運営にあたっては、広い産業界を含めた研究者コミュニティの意見を十分に反映できるような、透明性、公平性、客観性に十分配慮したオープンな運営体制を確保することが望ましい。
- さらに、新たな組織においては、DB を有する関係機関とのネットワークを密に図り、利用者本位のサービス提供を行うと共に、総合科学技術会議における検討結果も踏まえつつ、他機関との効果的な連携を重視し、統合 DB の整備に積極的に協力する研究者コミュニティに開かれた運営を行うことが期待される。

6. DBの統合・維持・運用の在り方 ～移行時期～

- ・ 平成 18 年度から 5 年間の期限付で進められている「統合 DB プロジェクト」終了時の平成 23 年度に向けた段階的移行を進めるため、平成 21 年度の予算編成より経過的措置を講じていくことが望ましい。

(参 考)

これまでの経緯

- 平成 12 年 11 月、当時の科学技術会議ライフサイエンス部会ゲノム科学委員会において、「ゲノム情報科学におけるわが国の戦略について」が取りまとめられ、バイオインフォマティクスの人材養成、研究開発の振興、データベース整備戦略の 3 つの課題に関して推進方策が提言されたのを受け、平成 13 年度より科学技術振興機構（JST）にバイオインフォマティクス推進センター（BIRD）が設立された。
- その後、BIRD による支援を通じて、バイオインフォマティクス研究の推進や、大阪大学の PDBj(Protein Data Bank Japan)、京都大学の KEGG(Kyoto Encyclopedia of Genes and Genomes)、国立遺伝学研究所の DDBJ(DNA Data Bank of Japan)といった世界的に定評のあるデータベース構築等が進められた。
- 一方で、平成 12 年当時の予想を超え、
 - 研究の進展によりデータベースが多く作られるようになったこと（そのためのポータルサイトや統合化などの必要性が増したこと）
 - 大型プロジェクトのデータベースの受け皿作りの必要性が高まったこと
 - オントロジーや文献の知識などの重要性が増したこと等に応えるべく、データ整備戦略の見直しの必要性が出てきたことから、平成 17 年 8 月、ライフサイエンス委員会の下にデータベース整備戦略作業部会が設置され、そこでの議論を経て平成 18 年 5 月「我が国におけるライフサイエンス分野のデータベース整備戦略のあり方について」が取りまとめられた。
- これを元に、平成 18 年度より 5 年間の時限付きプロジェクトとして、大学共同利用機関法人情報・システム研究機構(ROIS)を中核機関とした「ライフサイエンス分野における統合データベースプロジェクト（統合 DB プロジェクト）」が開始された。

ライフサイエンス情報基盤整備作業部会における審議の過程

○第1回 平成20年10月14日

- ・ライフサイエンス分野の情報基盤の現状について

○第2回 平成20年11月14日

- ・ライフサイエンスデータベースの統合・維持・運用の在り方について

○第3回 平成20年12月15日

- ・ライフサイエンスデータベースの統合・維持・運用の在り方について

ライフサイエンス情報基盤整備作業部会委員名簿

(委員)

- 秋山 泰 東京工業大学大学院情報理工学研究科教授
- 飯島 貞代 三菱ケミカルホールディングス株式会社ヘルスケア戦略室部長
- 漆原 秀子 筑波大学大学院生命環境科学研究科教授
- ◎勝木 元也 自然科学研究機構理事
- 五條堀 孝 情報・システム研究機構国立遺伝学研究所副所長
- 末松 誠 慶應義塾大学医学部長
- 諏訪 牧子 独立行政法人産業技術総合研究所生命情報工学研究センター
主幹研究員
- 豊田 哲郎 独立行政法人理化学研究所横浜研究所生命情報基盤研究部門長
- 長洲 毅志 エーザイ株式会社理事 研究開発担当付担当部長
- 深海 薫 独立行政法人理化学研究所筑波研究所バイオリソースセンター
情報解析技術室長
- 堀田 凱樹 情報・システム研究機構長
- 水上 政之 独立行政法人科学技術振興機構理事
- 水澤 博 独立行政法人医薬基盤研究所生物資源研究部長
- 吉田 光昭 東京大学大学院新領域創成科学研究科客員教授

計：14名
(敬称略 50音順)

◎は主査

(オブザーバ)

- 菅原 秀明 情報・システム研究機構国立遺伝学研究所特任教授
- 高木 利久 情報・システム研究機構ライフサイエンス統合DBセンター長

ライフサイエンス情報基盤整備作業部会の設置について

平成20年8月22日
科学技術・学術審議会
研究計画・評価分科会
ライフサイエンス委員会

1. 設置の趣旨

生命情報の統合化データベースは今後のライフサイエンス研究を支える基盤であり、我が国がライフサイエンス研究において国際的に伍していくためには、取組を強化していく必要があり、その必要性は、ライフサイエンス分野の推進戦略(平成18年3月。総合科学技術会議策定)の中でも掲げられている。

こうした中、文部科学省では、統合化データベースの整備を進めるために必要な戦略の検討と技術開発を行うため、平成18年度から「ライフサイエンス分野の統合データベース整備事業(「統合データベースプロジェクト」)」を開始したところであるが、同プロジェクトは5年間の時限付きで進められ、平成22年度末に終了することとなっている。しかしながら、データベースは継続的に整備され、利用されることにより大きな価値が認められることから、その後も引き続き、整備された統合データベースを維持し発展させることが肝要である。

このため、ライフサイエンス委員会運営規則第2条第1項に基づき、ライフサイエンス委員会にライフサイエンス情報基盤整備作業部会を設置し、ライフサイエンス分野における情報基盤整備を進めるに当たって必要な検討を行う。

2. 検討事項

- ・今後推進すべきバイオインフォマティクス研究開発の方向性
- ・データベースの統合・維持・運用の在り方

3. 設置期間

作業部会の設置が決定された日から平成21年1月31日までとする。

4. その他

本作業部会の庶務は研究振興局ライフサイエンス課が処理する。

「我が国のデータベース整備」

—科学技術連携施策群としての活動報告—

国立遺伝学研究所

五條堀 孝

(代理:菅原秀明)

ライフサイエンス統合データベースに関する経緯

(内閣及び総合科学技術会議)

(文部科学省) (農林水産省) (経済産業省) (厚生労働省)

～16年度

BT戦略会議
「ミレニアムプロジェクトの成果や活発に展開されるポストゲノムプロジェクト等の成果を実用化研究に活用できるように、一元的に集約・統合したデータベースの構築が不可欠」(第7回BT戦略会議委員(平田正 協和発酵工業株式会社代表取締役会社社長)提言)

DDBJ, PDBj, KEGG 他

イネゲノムDB (RAP) 他

H-Inv 他

MDBj 他

17年度

総合科学技術会議科学技術連携施策群「ポストゲノムー健康科学の推進」の設置
“ライフサイエンス統合データベース統合化の取組”を補完的課題として取り上げ、調査を開始。

統合データベースプロジェクト
・ポータルサイト
・一括検索システム
他

農林水産物ゲノム情報統合データベースプロジェクト
イネゲノム他の統合DBの構築

18年度

第3期科学技術基本計画(18年3月)
ライフサイエンス分野の戦略重点科学技術「世界最高水準のライフサイエンス基盤整備」として選定

19年度

総合科学技術会議科学技術連携施策群「ポストゲノムー健康科学の推進」の終了

統合データベースプロジェクト
経産省関連機関の研究成果を統合したデータベースの構築

20年度

総合科学技術会議ライフサイエンスPTの下にライフサイエンス統合DBに関するWG設置

21年度

22年度

23年度～

各省連携したデータベースの構築を目指す

科学技術連携施策群

各府省の縦割りの施策に横串を通す観点から、国家的・社会的に重要であって関係府省の連携の下に推進すべきテーマを定め、科学技術連携施策群として積極的に推進
(総合科学技術会議決定 平成16年7月23日)

【目標、主な関係府省】

**ポストゲノム
—健康科学の
推進—**

テーラーメイド医療やゲノム創薬、予防医学などの確立を目指す

文部科学省、厚生労働省
農林水産省、経済産業省

**新興・再興
感染症**

新興・再興感染症から国民の安心・安全を守る研究体制の確立を図る

内閣府、文部科学省、厚生労働省
農林水産省

**ユビキタス
ネットワーク
—電子タグ技術
等の展開—**

ユビキタスネットワーク社会実現の上で中核的な技術基盤の確立を図る

総務省、文部科学省、経済産業省
国土交通省

**次世代
ロボット
—共通プラット
フォーム技術
の確立—**

次世代ロボットのさまざまな応用分野に共通のプラットフォーム技術の確立を図る

総務省、文部科学省、国土交通省
経済産業省、農林水産省

【目標、主な関係府省】

**バイオマス
利活用**

バイオマス利用、燃料転換等の技術開発により循環型社会形成を目指す

総務省、文部科学省、農林水産省、
経済産業省、国土交通省、環境省

**水素利用/
燃料電池**

水素エネルギー社会実現のため水素利用、燃料電池技術の確立を目指す

総務省、文部科学省、経済産業省、
国土交通省、環境省

**ナノバイオ
テクノロジー**

ナノとバイオの融合領域研究により健康寿命延伸等安心安全な社会を目指す

文部科学省、厚生労働省
農林水産省、経済産業省、環境省

**地域科学技術
クラスター**

地域における革新技術・新産業創出を通じた地域経済の活性化を図る

内閣府、総務省、文部科学省、厚生労働省
農林水産省、経済産業省、国土交通省、環境省

「生命科学の基礎・基盤」連携施策群の位置付け

専門性を高める為に
改組を行った

補完的課題を継承

連携施策群
「ポストゲノム」

連携施策群
「生命科学の基礎・基盤」
コーディネーター
五條堀 孝(国立遺伝学研究所 副所長)

対象となる戦略重点科学技術
「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」

連携施策群
「食料・生物生産研究」
コーディネーター
小川 奎(農業・食品産業技術総合研究
機構 理事)

対象となる戦略重点科学技術
「国際競争力を向上させる安全な食料の生
産・供給科学技術」
「生物機能活用による物質生産・環境改善科
学技術」

連携施策群
「臨床研究・臨床への橋渡し研究」
コーディネーター案
松澤 佑次(住友病院 院長)

対象となる戦略重点科学技術
「臨床研究・臨床への橋渡し研究」
「標的治療等の革新的がん医療技術」

平成17, 18年度

平成19年度

目標と対象

目標)

世界最高水準の研究開発環境を提供するライフサイエンスデータベースの構築を中心とした、ライフサイエンス研究における国際的優位性の確保を目標とする。

対象とする戦略重点科学技術と対象施策の例

世界最高水準のライフサイエンス基盤整備

対象施策)

- ・統合データベースプロジェクト(文)
 - ・経産省統合データベースプロジェクト(経)
 - ・創薬基盤総合研究(厚)
 - ・農林水産生物ゲノム情報統合データベースプロジェクト(農)
 - ・ナショナルバイオリソースプロジェクト(文)
- 等

生命プログラム再現科学技術

対象施策)

- ・ゲノム機能解析の推進(文)
 - ・ターゲットタンパク研究プログラム(文)
 - ・糖鎖機能活用技術開発(経)
 - ・機能性RNAプロジェクト(経)
- 等

活動と成果

活動

主な成果

平成
17年度

- ・ 登録施策の精査を行い、連携強化すべき課題を抽出した(統合DB)。
- ・ 重要な研究開発課題について検討した。



- ・ 連携強化すべき研究領域、施策が明確となり意識共有が図れた。
- ・ 補完的課題「統合DB」の選定。
- ・ 分野別推進戦略の策定に寄与した。

平成
18年度

- ・ 登録施策に関して俯瞰図を作成した。
- ・ 科学技術連携施策群の成果及び今後の見通し(中間報告案)を作成した。



- ・ 俯瞰図により、各省施策の個々の位置付けを明確化でき、概算要求に反映させることができた。

平成
19年度

- ・ 海外のバイオリソース事業の責任者を招聘し、講演会を実施した。
- ・ 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」)を開催した。



- ・ バイオリソース事業・データベース統合化への施策に関して、学术界や産業界のより一層の理解を得ることができた。
- ・ 補完課題「統合DB」

現在、関係各省において統合データベース事業が開始され、将来の統合的なデータベース化を目指した、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

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」 —研究代表者 大久保 公策教授（国立遺伝学研究所）—

現状の生命科学データベース(DB)に関する課題
大規模なデータ産出型の施策の成果は広く利用されておらず、十分に活用できる状態にはなっていない。

調査・研究



国内外のDB調査



産業界等からの
ヒアリング



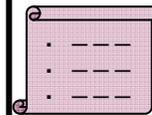
DB統合に向けた
技術的FS



DB統合に向けた
制度調査



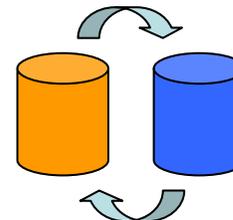
主な成果の例



国内最大の生命科学分野
データベースリスト



創薬企業研究者の
要望のとりまとめ



省をまたいだ
データベース
連携



DB統合が進む米国
等の制度の調査

＜連携施策群ポストゲノムにおけるライフサイエンス統合データベースに関する議論＞

- 平成17年8月 平成17年度のポストゲノム連携群の補完課題を「ライフサイエンス分野のデータベース統合化に関する調査研究」と決定
- 平成17年10月 科学技術振興調整費審査部会に於いて補完課題の採択決定
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19年度まで)
研究代表者:情報・システム研究機構 国立遺伝学研究所 教授 大久保公策
参画機関:国立がんセンター、(独)産業技術総合研究所、
(独)農業生物資源研究所(平成18年度から東京大学)
- 平成17年12月 平成17年度第1回タスクフォース 議事内容:補完課題の新年度体制について
- 平成18年 3月 分野別推進戦略 ライフサイエンス分野
戦略重点科学技術「世界最高水準のライフサイエンス基盤整備」
重要な研究開発課題「生命情報統合化データベースの構築に関する研究開発
- 平成18年 5月 平成18年度第1回タスクフォース 議事内容:補完課題の平成17年度成果報告及び今後の進め方について
- 平成18年11月 平成18年度第2回タスクフォース 議事内容:補完課題の平成17年度中間報告及び今後の進め方について
- 平成19年 2月 平成18年度第3回タスクフォース 議事内容:統合データベースの概念設計について(1)
- 平成19年 2月 平成18年度第4回タスクフォース 議事内容:統合データベースの概念設計について(2)
- 平成19年 2月 平成18年度第5回タスクフォース 議事内容:統合データベースの概念設計について(3)
- :
- (以後 平成20年3月まで議論を続ける)
- (一文科省ライフサイエンス統合DB運営委員会でも各省オブザーバーで出席)
- 平成20年3月 科学技術振興調整費審査部会補完課題「生命科学データベース統合に関する調査研究」の終了とともに、連携施策群「ポストゲノム」終了。
総合科学技術会議ライフサイエンスPTの下に統合データベースWGの設置。

今後の課題

世界最高水準のライフサイエンス基盤整備

生命プログラム再現科学技術

バイオリソース

データベース

現状

情報共有

各省での統合への
取組開始

基礎データの取得

課題

・継続的維持
・開発

・情報環境整備
・各省連携強化

連携

目標

世界最高水準のライフサイエンス基盤の構築

生命プログラムの理解

統合的なデータベース整備やバイオリソース維持等の継続的に検討が必要な課題は、
ライフサイエンスPTのもと検討を継続して行う。

ライフサイエンス情報基盤整備作業部会資料

統合データベースプロジェクト

ープロジェクト発足の経緯と背景及びその目的と現状ー

2008年10月14日

高木 利久

ライフサイエンス統合データベースセンター(DBCLS)
大学共同利用機関法人 情報・システム研究機構(ROIS)

我が国におけるライフサイエンスDBの問題点

- DBが散在していて所在情報や利用法が分からない
 - 我が国で300、世界で10,000ものDB
 - 似たようなものがいくつもありどれを使ってよいか分からない
- DBやDBのエントリに信頼性の高い注釈がついていない
 - DB構築、維持を行える人材不足、DB構築への評価の低さ
- 大型プロジェクトの成果公開が不十分
 - 我が国のライフ関係予算は年間3,000億円以上
 - 公開されているものもプロジェクトが終了すると更新ストップ
- ばらばらに構築・管理されていて検索・解析・応用が困難
 - 現在の統合化は分子レベルで行われていて医療などへの応用困難
 - 日本語化されていないので研究動向や分野の状況の把握困難
- 不可欠な基盤なのに我が国にはDB戦略がない

文科省での取り組み

内閣府連携施策群での取り組み
振興調整費「LSDB統合に関する
調査研究」(H17から19)
(代表 大久保公策)

- **ライフサイエンス委員会**
 - データベース整備戦略作業部会
 - 平成18年5月17日報告書
 - 国際動向を踏まえた問題点整理と推進方策(10項目)
- **統合データベース整備事業の公募**
 - 平成18年度(戦略立案、ポータルサイト、統合化技術)
 - 平成19年度(~22年度の4年間)
- **情報・システム研究機構での取り組み**
 - 統合データベースセンター設置(19年4月)

データベース整備戦略作業部会で提言された推進方策(1)

①現状調査、評価、戦略立案

データベースの現状や動向の定常的な調査および既存の戦略や活動の弛まぬ評価に立脚して、省庁の枠を超えて国家的視野に立って、ライフサイエンス研究全般やバイオ産業全般を見渡した戦略立案する。

②基盤データベースの支援

世界的競争力の確保に向けて戦略的に重要なデータベースなどについては、安定的、永続的に支援することが必要である。

③ポータルサイトの構築と運営

我が国では数多くのデータベースが日々作られている。これらを十分に活用するためには、常に最新の情報を保持したポータルサイトが不可欠である。

④統合データベース研究開発

バラバラだったデータベースを統合化することができれば、これまで別々のデータベースに収められていたデータ間の潜在的な関係を見出すことが可能になる。

⑤維持困難なデータベース受入れ

プロジェクトや科研費などの研究費が終了するなどして維持が困難になったデータベースの受け皿を国として用意する必要がある。

データベース整備戦略作業部会で提言された推進方策(2)

⑥文献情報との連携

機能情報の多くは論文の中にテキストとして記述されていることから、文献中に記述されたデータや知識と、配列や立体構造などの実験データとの連携と統合に今後取り組まなければならない。

⑦アノテーションの実施

未解析、未解釈のまま放置されている種々の実験データの意味付け(生物学的、医学的な解釈)を強かに推進すべきである。

⑧新たなデータベース構築への投資

ライフサイエンス研究の進展に対応した新たなデータベース、新たな発想に基づくデータベースの構築にも投資すべき。

⑨データベースを活用した研究の促進

データベース構築への支援と並行して、それを活用する技術の研究開発、いわゆるバイオインフォマティクスの促進も図る必要がある。

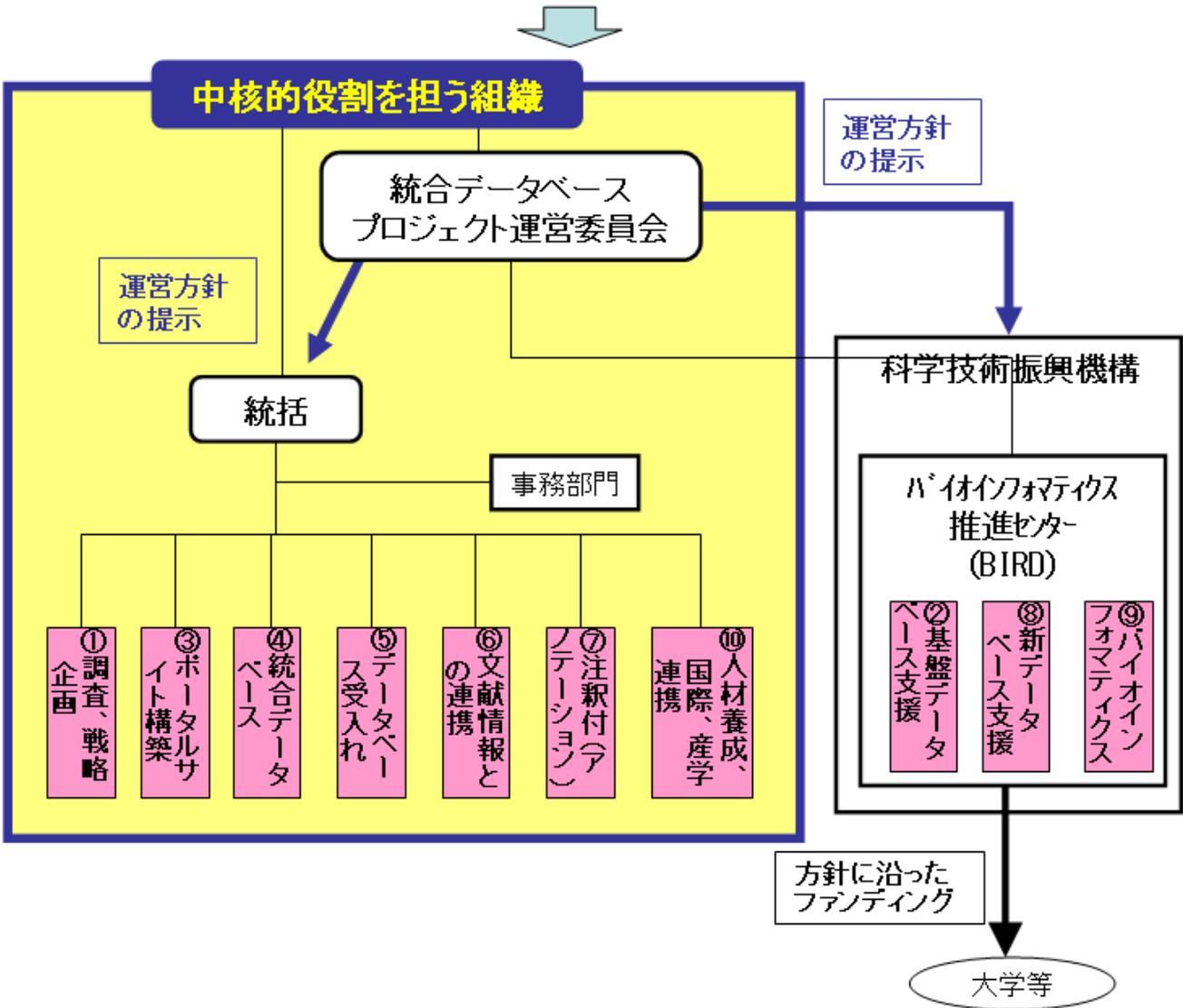
⑩データベース開発のための人材育成

有用なデータベースを開発するには、まずはアノテータやキュレータの安定的な職を数多く確保するとともに、それに相応しい人を養成することが不可欠であり、そのための体制を早急に確立する必要がある。

運営体制

文部科学省科学技術・学術審議会
ライフサイエンス委員会データベース整備戦略作業部会
文部科学省としての統合化戦略の提示

- 委託費の範囲
- 作業部会報告書の取り組むべき課題



生命科学データベース統合に関する調査研究(H17-19)(概要)

DB統合が望まれる背景 ~研究内部の変化と研究環境の変化~

- ・ 巨大観察データの登場によるライフサイエンス全体のデータ(ベース)依存型への変貌
- ・ 高額な巨大データに関して大学・企業研究開発の公共データ(ベース)への依存増
- ・ 科学の民営化(データ占有による知財獲得、データベースサービスも競争的)

多面的かつ
複雑な背景

課題：「生命科学データベース統合に関する調査研究」

- 実施内容：①国内外のDB調査、産業界等からのヒアリング
②DB統合に向けた技術的FS：省をまたいだDB連携
③DB統合に向けた制度調査：DB統合が進む米国等の制度調査

技術的課題は
十分対応可能

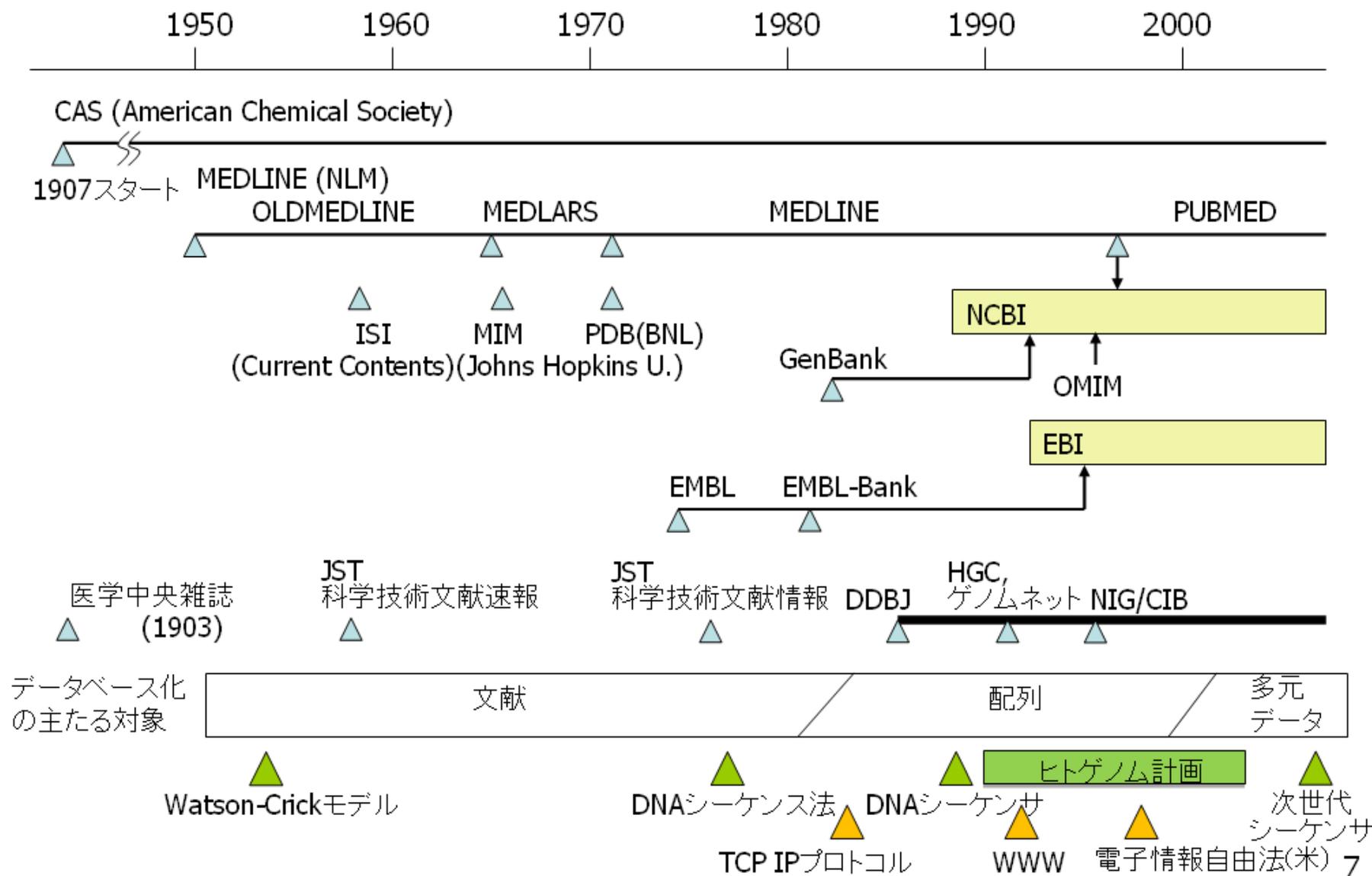
制度上の問題
を明確化

NCBIを代表とする米国での統合DBの成功例とそれを支える社会の仕組み(制度)の変化
OECD等国际的な公的機関によるデータ共有の必要性の認識

DB統合化に向けた提言に資する試案

1. 「政府資金によるデータ産生型プロジェクトのデータを我が国の研究社会で早期に共有するためのルール」が必要
2. 公共財としてのデータを保全・管理し、長期にわたるデータの育成と共有を行う公的機関を設ける(ルール運用の際の受け皿)

ライフサイエンス情報基盤の歴史



日米欧の現状比較

	NCBI	EBI	日本		
			DBCLS	DDBJ	BIRD
運営形態	根拠法に基づき NLM予算で運営	EMBLが半額負担、残りは外部資金	文科省委託プロジェクトで実施	国立遺伝研の運営費交付金	JST運営費交付金(バイオインフォマティクス推進事業)
予算	80億円	45億円	6億円(DBCLS分)	12億円(除く教官人件費)	17億円
人員	約400名 (正規職員約150名) サービス55% 研究 30% その他 15%	約300名 (原則有期雇用) サービス66% 研究 22% その他 12%	28名 (主にプロジェクト費用で雇用、客員教員、事務を含む)	58名(教員 10名、他はアナテータ、事務、SE) サービス79% 研究 12% その他 9%	87名(JST雇用者および委託費雇用者)
提供DB	35種類 (ほとんどが基幹DB相当)	60種類 (うち15種類が基幹DB)	10種類 (その他ツール、辞書15種類公開)	20種類	25種類 (BIRD事業で作成支援)
特徴	塩基配列バンク構築、PubMed等文献情報と配列等の分子情報の提供とEntrezによる統合、BLAST等解析ツール	塩基配列バンク構築、タンパク配列を基礎とした機能情報、ゲノム情報の統合サービス、Web service等へのイニシアティブ	DB戦略立案、統合DB開発、ポータル(DBカタログ、横断検索、統合検索)DB受入れ	塩基配列バンク構築、日米欧三極連携拠点、国プロDB構築の受け皿	基幹DB支援、新規DB開発支援、バイオインフォマティクス研究支援

欧州における新たな試み：ELIXIRプロジェクト

ELIXIR: a sustainable infrastructure for biological information in Europe

第7次欧州研究開発フレームワーク計画の枠組みで、ヨーロッパにおける今後10～20年の持続可能な生物情報基盤を構築し・運用することを目的に、EMBL-EBIが取りまとめ(コーディネータ Janet Thornton)の、13カ国32機関(企業を含む)によるコンソーシアム体制で推進。2007/11～2010/12が準備フェーズで、2011年から基盤構築フェーズに入る。予算は、準備フェーズが€4.5Mで、基盤構築フェーズが€500Mの見込。

情報基盤整備

- ヨーロッパの主要な生物情報データリソースの維持、運用とそれらへの持続的な資金拠出
 - EMBL-Bank, UniProt, The Macromolecular Structure Database, ArrayExpress, Ensemble
- European Biomolecular Data Centre(スパコンセンター)の建設を含めた施設のメジャーアップグレード
- バイオインフォマティクスツール統合に向けたインフラ構築
- 生物学情報やサービスプロバイダのための国家横断的なインフラ構築

標準化、評価、調整

- 下記項目を調整するためのヨーロッパの意志決定機関(資金提供も行う場合有)
 - 新しいデータリソース構築
 - 標準規格、オントロジー作成
 - 相互運用のための基準
 - データリソースアセスメント(評価)の基準
 - ヨーロッパとその他の国々との調整

他分野との連携、普及

- 文献、医学、環境、化学の他の分野のデータリソースの統合に向けたロードマップ作成
- 情報がより利用されるためのトレーニング・普及の戦略

運用・維持

- 本基盤を持続的に構築、運用するための規則、予算の枠組みの作成

プロジェクト遂行の基本方針

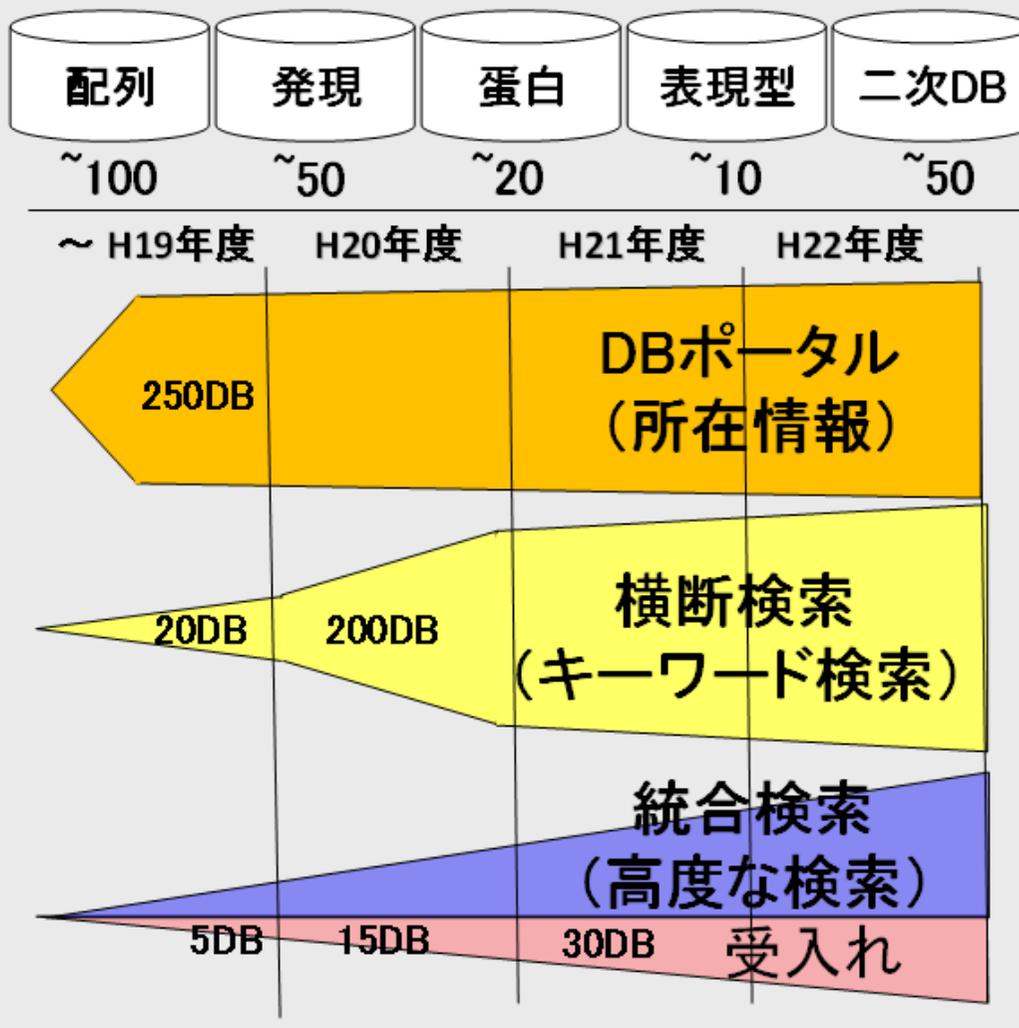
- 文科省だけでなく我が国の中核センターを指向
 - 日本版NCBI、ただし、まねはしない
 - 医学図書館やゲノムセンターの一部の機能も
 - 様々な立場、利用者、府省の人からなる委員会
- 研究ではなくサービス事業として推進
 - 我が国の生命研究の効率化と質の向上に寄与
 - それに必要なことは何でもする(技術面以外も)
 - 網羅性、利便性、持続性の追求
 - 自分では知識発見研究しない(縁の下の力持ち)

整備、統合に対する考え方

- 技術面だけではなく制度面での取り組み
 - データの権利関係、個人情報の取り扱い、なども
- 統合は手段であり、目的ではない
 - 物理的に一つのDBが目的ではない
 - 望ましい整備、統合は研究分野や利用者毎に異なる
 - 統合のあり方は研究の進展とともに変化する
- 目的は生命研究の研究開発の効率や質の向上
 - 必要十分なデータにすばやく辿り着ける
 - 公開ではなく共有化を図り、データマイニングを可能に
- 研究開発に関わるデータすべてを対象に
 - データだけでなく、文献(論文、総説)、図表、特許等も

統合化の年次計画

国内データベース(約250について)



文科ライフ委DB戦略WS報告

H18年度
戦略立案
ポータル
統合技術

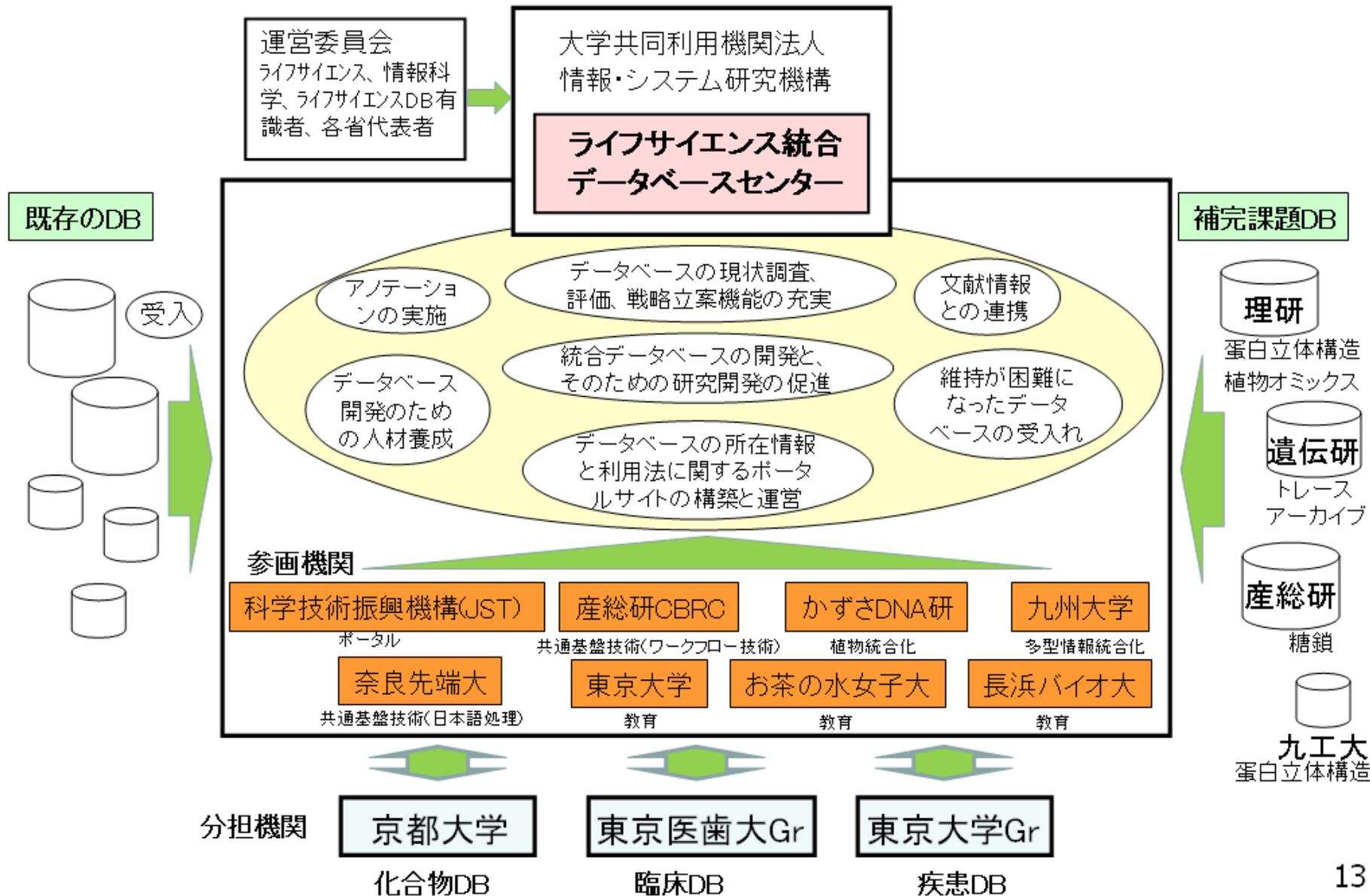
産業・臨床・育種
従事者の課題
解決への展開

4省
庁
統
合
DB

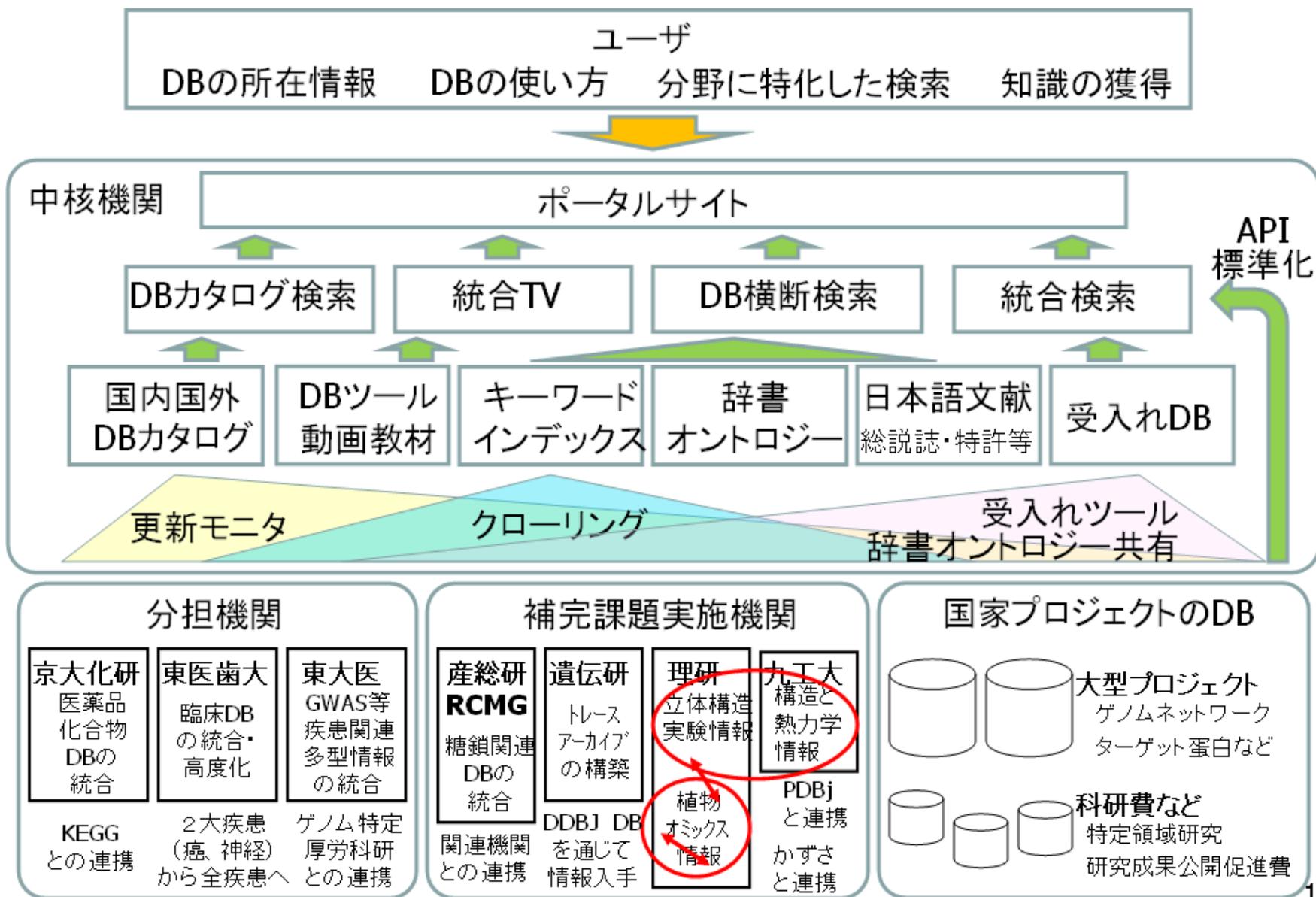
環境変化と知識
データ増大への
絶え間ない対応

統合データベースプロジェクトの体制図

中核機関



統合データベースプロジェクト連携の概要



統合ホームページへようこそ
 このウェブサイトは文部科学省「統合データベースプロジェクト」のポータルサイトです。



ポータル
[生命科学系データベースカタログ](#) 目録
[生命科学 学協会カタログ](#) 目録
[ゲノム・ポストゲノム主要プロジェクト一覧](#)



教材・人材育成
[統合TV\(DBやツールの動画教材\)](#) 目録
[MotDB \(教育・人材育成のサイト\)](#)



DB検索
[生命科学データベース横断検索](#) 目録
[DNAデータベース総覧と検索](#) 目録
 (DDBJ/EMBL/GenBank)
[遺伝子発現バンク\(GEO\)目録](#) 目録



基盤技術開発
[共通基盤技術開発の概要](#)
[誰でもデータベースが構築できる TogoDB](#)
[国内ウェブサービスの統合 TogoWS](#)
[OpenID 認証システム](#) 目録



文献検索
[蛋白質組解析 全文検索](#)
[文科省「ゲノム」研究報告書 全文検索](#) 目録
[学会要旨統合検索](#) 目録
[新聞記事検索](#) 目録



統合DB事業
[国内データベースの統合\(受入れ\)事業](#)
[文科省 統合データベース整備事業サイト](#)
[H18年度成果公開サイト](#)
[情報公開 \(委員会などの記録\)](#)



自然言語処理によるサービス
[QReFil \(オンラインリソースファインダー\)](#)
[Allie \(略語の正式名称を検索\)](#) 目録
[遺伝子名シノニマス検索](#)
[遺伝子名ゆらぎ取組システム](#) [遺伝子・タンパク質名総覧](#)



参画分担機関・補完課題の成果
[かずさアノテーション \(かずさDNA研究所\)](#)
[かずさNavigation](#)
[ゲノムネット医薬品データベース \(京大\)](#)
[統合医科学データベース \(医科歯科大グループ\)](#)
[疾患解析から医療応用を実現するDB開発 \(東大グループ\) **new**](#)
[WingPro \(JSTのDBポータル\)](#)
[Webリソースポータルサイト \(JST解析ツールポータル\)](#)
[持続可能型社会への貢献遺伝子データベース \(長浜バイオ大\)](#)
[日本糖鎖科学統合データベース \(JCGGDBワーキンググループ\) **new**](#)



統合ツール
[アナトモグラフィ/BodyParts3D\(解剖図理解権\)](#) 目録
[Wired-Marker\(参照情報共有ツール\)](#) 目録



国際連携
[BioHackathon2008](#)



リソース
[辞書](#)
[生物アイコン](#) 目録

新着情報

- ▶ [「生命科学データベース横断検索」にデータベース追加のお知らせ](#) 目録 2008-09-05 (Fri)
- ▶ [ゲノムワイド関連解析\(GWAS\)のデータベース公開のお知らせ](#) 目録 2008-08-26 (Tue)
- ▶ [糖鎖に関するデータベース公開のお知らせ](#) 目録 2008-08-15 (Fri)

LSDBブログ

- ▶ [統合ホームページへようこそ](#) 目録 2008-09-05 (Fri) 19:47:36

ニュース

- ▶ [細胞増殖に関する酵素の構造解明 抗がん剤への応用も](#)
- ▶ [精巣スライス アルバム樹に凍結 理研、マウスの精子保存に新方法](#)
- ▶ [GABA抑制促進が アルツ病の記憶障害に関与 理研などの研究グループ発見](#)

バナーリンク


 日本糖鎖科学統合データベース

 糖鎖関連遺伝子データベース

 グライコプロテインデータベース

 レクチンプロンティアデータベース

 ゲノムワイド関連解析(GWAS)データベース

 標準SNPデータベース

中核機関が開発、提供している主なサービス

生命科学系 DBカタログ	DBサイトのトップページをサムネイルで表示。DBの稼働状況をモニターして表示。利用者からのコメントを受け付ける。登録DB数約400。
生命科学DB 横断検索	生命科学分野のDBや文献の横断検索サービスで、生命科学分野の国内外のDBと文献を網羅的に検索。検索対象DB数37。
DNAデータバンク 総覧と検索	巨大な国際DNAバンクを高速に検索し生物種、分子種、プロジェクトに分類して表示。検索結果の対象データの一括ダウンロードも可能。対象レコード数8000万。
「蛋白質 核酸 酵素」 全文検索	国内商業出版社による総説誌である「蛋白質 核酸 酵素」バックナンバーの全文検索サービス。個別の記事はPDFファイルで閲覧可能。検索対象1985年～2005年。
学会要旨統合検索	学会要旨集の検索サービス。著者名や施設名の表記揺れを吸収する辞書を搭載し、研究者の研究履歴をたどることが可能で、キーワードのトレンドを表示させることもできる。
統合TV	DBやウェブツールの使い方などを動画で配信するウェブサイト。プロジェクトのサービスの利用法も配信。Videocastや動画配信サイトからもコンテンツを利用可能。コンテンツ数80。
MotDB (Master of the Database)	アノテータ、キュレータ、DB管理者向けの教材提供サイト。独自に作成した教材のPDFファイルもダウンロード可能。プロジェクト主催の講習会情報も掲載。
Allie	文献中に登場する略称とその展開型の組を検索するシステム。略称や名称の一部で検索するとその使われ方(展開型)と登場する文献を一覧表示。
アナトモグラフィ/ BodyParts3D	解剖学用語が示す人体の部品(臓器、器官)の位置や形状を3次元人体モデルで記述したDB(BodyParts3D)と、そこから人体部分のモデルを描くツール(アナトモグラフィ)。
Wired-Marker	オンラインで閲覧したページ本文内をハイライトし、記録できるFireFox用プラグイン。蓄積された情報はあとから閲覧・編集可能で、アノテーションのエビデンスとして利用可能。

生命科学系データベースカタログ

世界に散在するデータベースの情報を閲覧・検索可能にしたデータベースのデータベース。更新状況をモニタ。利用者のコメントや評価を受け付ける仕組みも。



DB型(329)
 カタログ(7)
 データバンク(19)
 プログラム(66)
 プロジェクト(131)
 リソース(8)
 解析サービス(15)
 辞典(38)
 知識モデル(27)
 注釈(18)

対象(284)
 DNAモチーフ(1)
 DNA制御領域(3)
 DNA多型(25)
 DNA配列(42)
 DNA立体構造(1)
 RNA(5)
 アノテーション(24)
 タンパク質、アミノ酸特性(1)
 タンパク質、プロテオーム(5)
 タンパク質、モチーフ(5)
 タンパク質、機能(8)
 タンパク質、局在(1)
 タンパク質、抗体(2)
 タンパク質配列(10)
 タンパク質、変異(1)
 タンパク質立体構造(21)
 マーカー分子(1)
 遺伝子
 画像(8)

表示形式: サムネイル 表示順: 名前順 検索

COG - Clusters of Orthologous Groups of proteins
 タンパク質オルソログ・グループ

dbCFC
 Cytokine Family cDNA Database (dbCFC)

HIV感染症総合データベース(HIV-DB)

Hepatitis Virus Database(肝炎ウイルスデータベース)
 肝炎ウイルス遺伝子系統解析結果

Micro

詳細画面

LsdbID	133
Title	COG - Clusters of Orthologous Groups of proteins
URL	http://www.ncbi.nlm.nih.gov/COG
LsdbName	タンパク質オルソログ・グループ
Category	DB型: プログラム 対象: タンパク質配列
Catalog	The Bio Netbook Genexpress NAR catalog MetaDB bioformatik
Organization	NCBI :: USA
Desc-J	ゲノムが解析された原核/単細胞真核生物のタンパク質を、オルソロググループに分類したデータベース。グループは機能カテゴリー(COアノテーション)ごとに分類されている。真核生物ゲノム上で予測されたオルソログを同様に分類したデータベースはKO00として別途公開されている。

H20年9月現在の登録状況
 国内350, 国外50

生命科学系データベース横断検索

生命科学分野のDBや文献の横断検索サービスで、生命科学分野の国内外の主要DBと特許や日本語総説誌などの文献を網羅的に検索。検索対象DB数37 (H20年9月現在)。



The screenshot shows the search results for 'p53' on the DBCLS website. The interface includes a search bar at the top with the query 'p53' and a 'Search!' button. Below the search bar, there are navigation tabs for different user groups: '一般向け' (General), 'ヒト・動物研究者向け' (Human/Animal researchers), '植物研究者向け' (Plant researchers), and '微生物研究者向け' (Microbiology researchers). The search results are displayed in a list format, with a '遺伝子リスト' (Gene list) section showing various p53-related genes and their identifiers (e.g., GENE:7157, UNP:P04637). A '検索結果' (Search results) section provides detailed information for the top result, including a description of the p53 protein and its function, and links to external resources like Wikipedia and PubMed. The interface also features a sidebar with 'データベース一覧' (Database list) and '特許' (Patents) sections, and a 'DBカテゴリー' (Database category) section at the bottom left.

DBカテゴリー

検索語入力

遺伝子名リスト

検索結果

分担機関、補完課題実施機関の取り組みと成果

分担機関	
京都大学	各種化合物、医薬品データベースの統合化と種々の検索技術の開発 JAPIC医薬品情報の統合化と日本語支援機能を持った検索システム公開
東京医科歯科大Gr	がん疾患、神経疾患を対象にした臨床・疾患データベースの実証的統合 がん、パーキンソン病各100例の試験公開と国内臨床疾患DBの包括調査完
東大医学部Gr	標準SNP DB、GWAS(ゲノムワイド関連解析)DB、リシーケンスDBの構築 標準SNP DB(健常者500名)、6疾患のケース-コントロールDBの公開
補完課題実施機関	
理化学研究所	シロイヌナズナ、タンパク質立体構造を対象にしたDB統合化と理研DBの統合化のためのモデルケース構築
産総研糖鎖医工学セ	糖鎖業界に散在するDBを集約し、糖鎖科学統合データベースを構築 糖鎖関連遺伝子、糖タンパク質、レクチンDB公開、参画DBの横断検索提供
国立遺伝研	トレースデータ用DBの開発、トレースデータ用登録システムの開発、及びトレースデータ活用のためのFTPサイトの整備
九工大	蛋白質の安定性や相互作用の網羅的な熱力学データを構造データと統合 XML等のデータ交換フォーマットの整備、オントロジー等の統合化技術開発

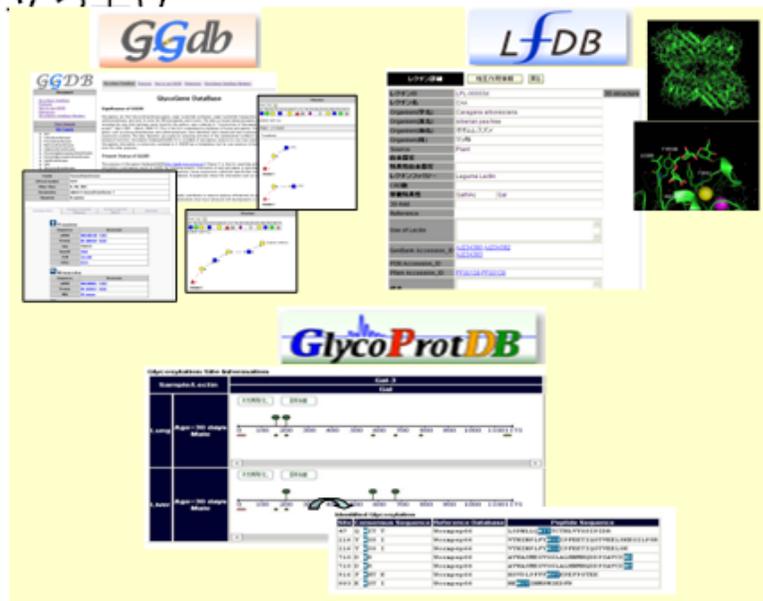
赤字は公開済成果 19

日本糖鎖科学統合データベース (JCGGDB)

— 産総研糖鎖医工学研究センターの成果 —

1) 日本糖鎖科学統合データベース(JCGGDB) の立ち上げ

- ・事務局 ; 糖鎖医工学研究センター
- ・20年度参画 ;
立命館大学(GlycoEpitope)
名古屋市立大学(GALAXY)
名古屋大学(KOマウスDB)
LipidBank構築委員会(LipidBank)



糖鎖センター内のDBの公開

2) 糖鎖センター内のDBの公開

- GGDB(糖鎖関連遺伝子データベース)
- LfDB(レクチンデータベース)
- GlycoProtDB(糖タンパク質データベース)

3) JCGGDB参画機関DBの横断検索

4) JCGGDB参画機関DBとのウェブサービスによる分散型の統合(構築中)

5) 中核機関との連携

- 横断検索における連携(共通インデックス仕様の採用)
- 中核で構築中のダウンロードサイトやDBカタログにおいても連携の作業中



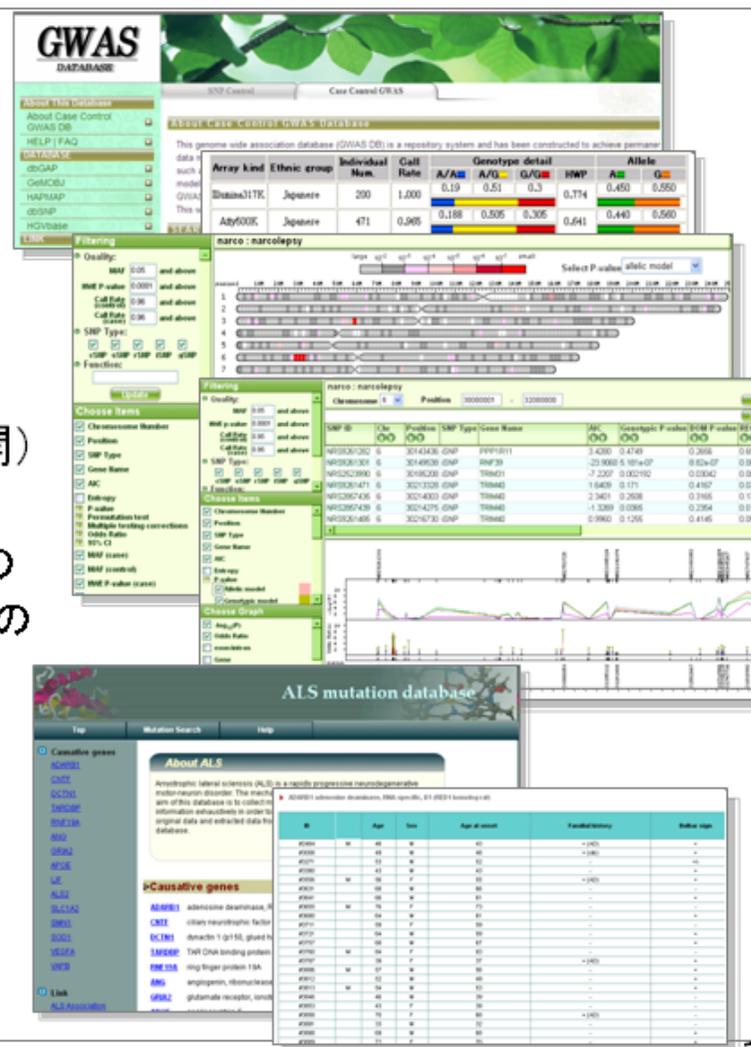
JCGGDB参画機関DBの横断検索

疾患解析から医療応用を実現するDB開発

— 東京大学医学部グループの成果 —

1. GWASデータベース

- ・標準SNP DBの構築 (08/08公開)
 - 健常者500名以上の30-100万SNPの遺伝子型頻度、アレル頻度、Call rate、ハーディーワインバーグ平衡検定値、連鎖不平衡値、ハプロタイプ頻度等
- ・Case-control DBの構築 (6疾患分を08/08公開)
 - 30-100万SNPの遺伝子型頻度、アレル頻度、ハーディーワインバーグ平衡検定値、Call rate等のP-value (genotype, allele)、Additive risk model等のP-value, OR, 95% CI, AICなどの遺伝統計値
 - SNPのアノテーション (機能、染色体上位置、同義/非同義など)



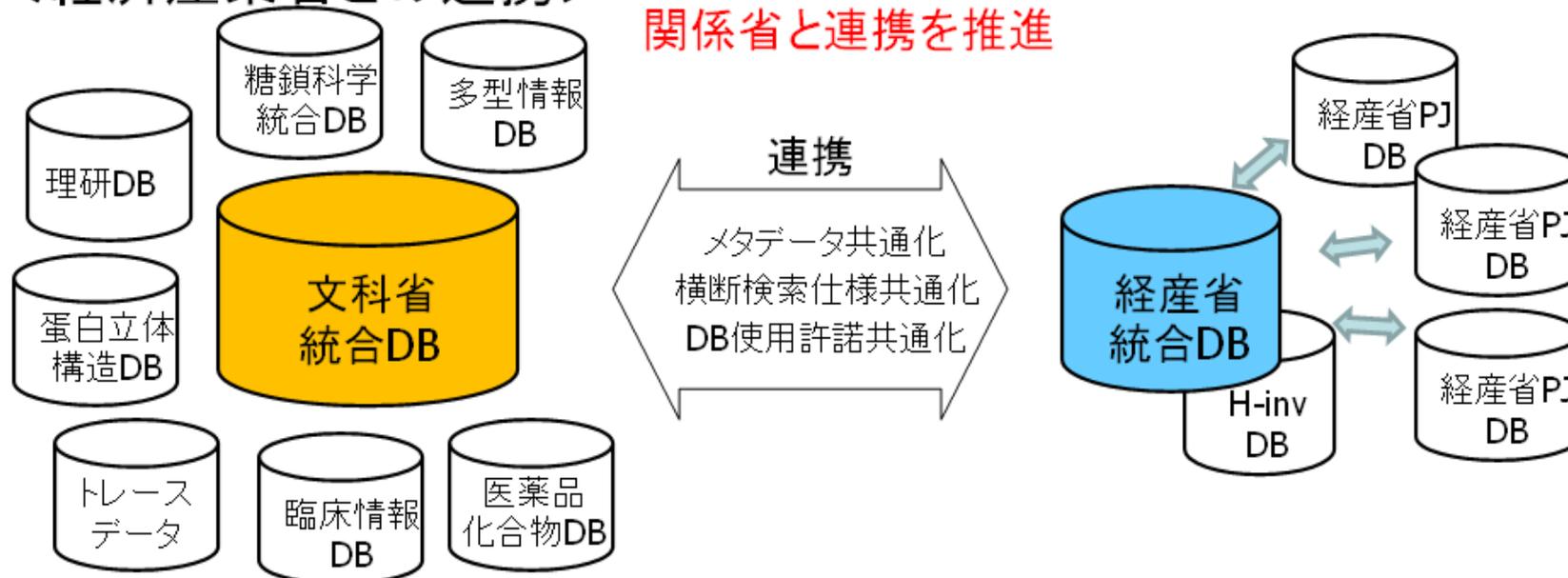
2. リシークエンスデータベース

- ・ALSリシークエンシングDB
mutation情報と臨床症状の関係を網羅的に収集

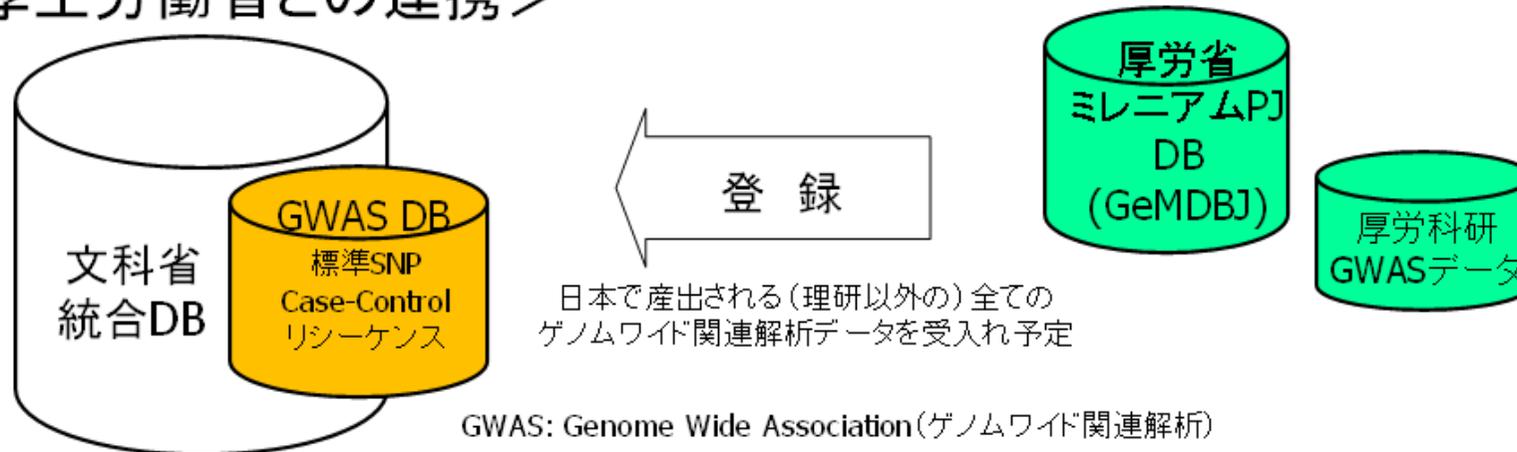
他省庁プロジェクトとの連携

<経済産業省との連携>

総合科学技術会議の指導の下、
関係省と連携を推進



<厚生労働省との連携>



GWAS: Genome Wide Association(ゲノムワイド関連解析)

進捗状況と今後の見通し

- 当初の計画通りサービスが実現できた
 - 中間評価(5月21日)で非常に高い評価
 - 参画各機関との連携には注文も
- 3年後にはプロジェクトの目標達成可能
 - 医療、医薬品関係、高度の意味付けは厳しい
- しかし、大幅な予算減による課題も
 - 計算機が容量不足
 - 次世代型シーケンサーの登場
 - 人件費不足で高度なアノテーションや辞書作り困難
 - 質の高い統合検索に影響も

今後の課題

- ・中核機関のイニシアティブがとりにくい体制
→ 体制や個別のミッションの見直し(中間評価)
- ・受け入れや共有化の推進の裏付けなし
→ データ共有化のためのルール作り
- ・ミッションと予算、権限、体制とのアンバランス
→ 生命科学の進展への柔軟な対応の担保
- ・プロジェクトがあと3年弱で終了
→ 恒久化のための体制構築
- ・データベースは国が支援すべき
→ 科学と産業振興のための国家的基盤作り



我が国のバイオインフォマティクスの現状 JSTバイオインフォマティクス推進センターについて

平成20年10月14日

独立行政法人科学技術振興機構(JST)

我が国のバイオインフォマティクスの現状

- バイオインフォマティクスに関連した政策動向①
バイオインフォマティクス振興策
- バイオインフォマティクスに関連した政策動向②
近年の関連プロジェクトの動向
- バイオインフォマティクスに関連した政策動向③
人材育成の取り組み
- バイオインフォマティクスの国際比較
- バイオインフォマティクスの産業貢献

JSTバイオインフォマティクス推進センター事業の取り組み

- 1.概要
- 2.生命情報データベースの高度化・標準化
- 3.バイオインフォマティクスの研究開発の推進
- 4.成果①、②、③、④
- 5.研究成果情報発信(ゲノムリテラシー講座)

我が国のバイオインフォマティクスの現状

バイオインフォマティクスに関連した政策動向①

-バイオインフォマティクス振興策-

第2期科学技術基本計画(2001-)

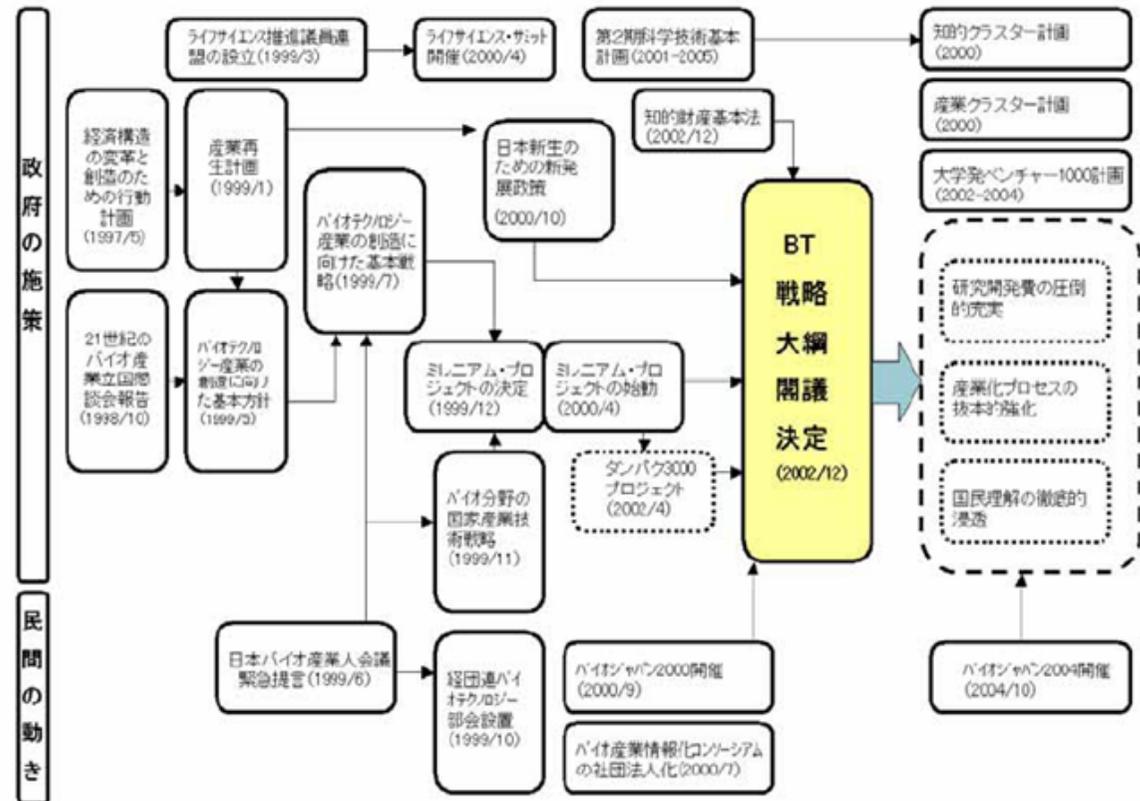
21世紀のバイオ産業立国懇談会報告(1998年):バイオインフォマティクスが重要技術として明文化。

バイオテクノロジー(BT)戦略大綱(2002年12月):バイオインフォマティクス振興策

直接的予算措置 (例)

- ・科研費 1991-
- ・未来開拓事業(ゲノム研究) 1996-1999(終了)
- ・JSTシーケンシングプロジェクト1995-2001終了
- ・ミレニアムプロジェクトの一部 2000-2004
- ・バイオインフォマティクス推進センター事業2001-

ライフサイエンス分野の政策の流れ



引用: 平成16年特許出願技術動向調査報告書バイオインフォマティクス(要約版).平成17年3月.特許庁
http://www.jpo.go.jp/shiryou/pdf/gidou-houkoku/16life_bio.pdf

我が国のバイオインフォマティクスの現状

バイオインフォマティクスに関連した政策動向②

-近年の関連プロジェクトの動向-

①各省の統合データベースプロジェクト(文科省2006-、農水省2006-、経産省2007-)

②各プロジェクト内にデータ管理機能等のインフォマティクス要素が組み入れられている。

例・橋渡し研究支援推進プログラム2007-

→データセンター機能

・ターゲットタンパク研究プログラム2007-

→研究情報を集約化して統合する「情報マネジメントシステム」を構築・運営

・ゲノムネットワークプロジェクト2004-2008

→共通リソースの整備、ゲノム機能情報の整備

・糖鎖機能活用技術開発2006-

・機能性RNAプロジェクト2005-

・ゲノム創薬加速化支援バイオ基盤技術開発プロジェクト

・理研の例

脳科学総合研究事業

→ニューロインフォマティクス 基盤プラットフォームシステムの開発

③バイオリソースの整備

例・ナショナルバイオリソースプロジェクト

・創薬基盤推進研究(生物資源・創薬モデル動物研究)

バイオインフォマティクスに関連した政策動向③

-人材育成の取り組み-

文部科学省科学技術振興調整費の新興分野人材養成ユニット等の取り組みにより、バイオインフォマティクス関連の講座、学科の新設が見られる。

例) 新興分野人材養成ユニット

大学院修士課程以上のレベルの実務者・研究者の養成

ライフサイエンス分野を中心とする総合領域		調整費支給年度	新設学科名等	設置年度
システム生物学者育成プログラム	慶應義塾大学	H13～17	理工学部生命情報学科	H14
生物情報科学学部教育特別ユニット	東京大学	H13～17	理学部生物情報科学科	H19
			大学院新領域創成科学研究科情報生命科学専攻	H11
奈良先端大蛋白質機能予測学人材養成ユニット	奈良先端科学技術大学院大学	H13～17	情報科学研究科情報生命科学専攻	H14
ゲノム情報科学研究教育機構	京都大学・東京大学	H14～18	京都大学大学院薬学研究科医薬創成情報科学専攻	H19
臨床的バイオスタティスティクスコア人材養成ユニット	久留米大学	H15～19	大学院医学研究科 バイオ統計学群	H16
システム生命科学人材養成ユニット	九州大学	H15～19	大学院システム生命科学府	H15

※企業等の研究者、技術者向けプログラム： 生命情報科学技術者養成コース 産総研 H17～21

バイオインフォマティクスの国際比較

◆融合分野 バイオインフォマティクス【ゲノム/配列DB、文献DB、配列解析アルゴリズム、立体構造解析、オミクス解析、パスウェイ解析】

国・地域	フェーズ	現状についての比較	近年のトレンド	留意事項などコメント全般
日本	研究水準	○	↘	バイオインフォマティクスが普及し一段落ついたと認識される傾向があり、研究者人口が減少。研究費が今後減少することも大きく影響。
	技術開発水準	○	↘	製薬やベンチャーが少なくもともと求人も少なかったが、最近になって三井情報や日立などが相次いでライフサイエンス部門から撤退。
	産業技術力	△	↘	基礎研究を行える体力を持つ製薬企業はもともと少ないが、最近の経営統合ブームで基礎研究からさらに撤退傾向に拍車。国内にある外資系企業の研究所も相次いで撤退（ロシュやグラクソに続き、今年ファイザーも中央研究所閉鎖）。
米国	研究水準	◎	→	1000ドルゲノムプロジェクト、人工ゲノム、ハプロタイプ解析などで一人勝ち状態がつづく。ゲノム配列読み取りで圧倒的な優位に立つため、関連する解析も全て米国中心。
	技術開発水準	◎	→	研究分野におけるニーズが非常に多く、ベンチャー企業の数、質ともに他国を圧倒。近年は超高速シーケンシング技術を独占。
	産業技術力	○	→	DNA チップ事業、DNA シーケンサーなどで圧倒。
欧州	研究水準	○	↗	旧ソ連や東欧圏において多くの研究者を輩出。優秀な人材の層が厚い。
	技術開発水準	○	↘	酵素科学の Zymogen などヨーロッパが優位な部門も多いが、一時のバイオベンチャーブームは去った。
	産業技術力	○	→	大手の製薬企業は当たり外れの多いバイオインフォマティクスやシステムバイオロジーに対して慎重姿勢。
中国	研究水準	×	↗	中国本土においてバイオインフォマティクスの研究は殆どみられない。近年、国内研究所へ米国等で優秀な成績を上げた研究者を呼び戻す運動が進んである。
	技術開発水準	×	↗	ヒトゲノムやヒトメタボロームに特化した形で国主導の大型解析が進行中。ただし開発力や成果は未知数。
	産業技術力	×	→	目立った活動をしている企業は無い。
韓国	研究水準	△	↗	バイオインフォマティクスの研究者が急速に増加しており、数年のうちに応用研究は日本に比肩するであろう。独自性という点では劣っている。国際会議等の誘致にも積極的。
	技術開発水準	×	→	英語による国外へのアピールに長けているが、システムバイオロジーやインフォマティクスという流行を追いかけているだけなので、長期的な視点にたった独自の技術は育っていない。
	産業技術力	×	→	目立った活動をしている企業はない。

全体コメント：測定、要素技術の開発では日欧が先行。出遅れた米国が急速に追いついた。特に植物メタボロームやリビドームで欧米が研究拠点を組織し、戦略的にこの分野を一気に推し進めようとする動きがある。中国も欧米型の拠点形成に動き出す。産業分野では、欧米で多くのベンチャーが誕生し、分析メーカーも参入した。日本の幾つかの発酵食品企業はメタボロームを研究開発に取り入れているが、製薬会社は様子見模様である。一方欧米の製薬企業は疾患マーカー探索や治療にメタボロームが有効か積極的に検討を始めている。

(註1) 現状について 【○：非常に進んでいる、○：進んでいる、△：遅れている、×：非常に遅れている】※

※我が国の現状を基準にした相対評価ではなく、絶対評価である。

(註2) 近年のトレンド 【↗：上昇傾向、→：現状維持、↘：下落傾向】

引用：ライフサイエンス分野科学技術・研究開発の国際比較2008年版.平成20年2月.独立行政法人科学技術振興機構研究開発戦略センター
http://crds.jst.go.jp/output/pdf/07ic02.pdf 2008-10-07

我が国のバイオインフォマティクスの現状

バイオインフォマティクスの産業貢献

バイオインフォマティクスにおける成果の活用・展開。

(例)

- ・創薬、試薬分野

 - ケモインフォマティクス 創薬支援

 - 遺伝子発現抑制 RNAi法。

 - アルゴリズムを用いたsiRNAの設計から合成

- ・測定技術・装置

 - シーケンサー、質量分析装置の発達

- ・受託解析

 - 高速シーケンス解析、遺伝子発現解析、質量分析 等

一方で、バイオインフォマティクス関係の部署の縮小や撤退。

(例) 複数の企業での関連部署や子会社の解散 2007年3月

1. 概要

目的

ゲノム情報等の生物情報データベースの構築、高度化、活用のための研究開発を行い、研究開発成果を情報発信することにより、情報基盤の整備の一翼を担うとともにライフサイエンス研究のさらなる進展に貢献する。

背景

- ・データベースの構築・維持更新、人材不足の問題が顕在化
- ・2000年11月報告書「ゲノム情報科学における我が国の戦略について」
(人材養成、研究開発促進、データベース整備戦略)
- ・第2期科学技術基本計画

設置

- ・平成13年4月

センター運営

- ・統括: 勝木元也 (自然科学研究機構 基礎生物学研究所 名誉教授)
- ・副統括: 高木利久 (東京大学 大学院新領域創成科学研究科 教授)
- ・委員会: 事業推進、研究開発課題の選考、評価等を行うため設置
- ・計算資源: 科学技術計算を行うための共用サーバを設置 (H19より縮小)
- ・契約・経理等事務業務

予算推移

H13年度	H14年度	H15年度	H16年度	H17年度	H18年度	H19年度	H20年度
20.0億円	21.5億円	20.2億円	17.7億円	18.6億円	18.0億円	16.8億円	16.8億円

2. 生命情報データベースの高度化・標準化

第Ⅰ期 研究開発課題: 4課題(研究開発期間:平成13年~17年度)

第Ⅱ期 研究開発課題: 7課題(研究開発期間:平成18年~22年度)

第Ⅱ期 代表研究者	研究開発課題
稲葉 一男(筑波大学 教授)	ホヤプロテイン統合データベースの構築(CIPRO)
金久 實 (京都大学化学研究所 教授)	ゲノムと環境の統合解析による生命システムの機能解読(KEGG)
菅原 秀明 (国立遺伝学研究所 特任教授)	バイオ基幹情報資源の高準化と共用化
高木 利久(東京大学 教授)	オントロジーによるパスウェイの高度化および国際標準化(INOH)
中村 春木 (大阪大学蛋白質研究所 教授)	蛋白質構造データバンクの国際的な構築と高度化(PDBj)
西岡 孝明 (慶應義塾大学 教授)	メタボロームMSスペクトル統合データベースの開発(MassBank)
森下 真一(東京大学 教授)	マルチモーダル統合バイオDB

3. バイオインフォマティクスの研究開発の推進

内容: 情報科学と生物科学の融合したアプローチにより、生物情報から
生物現象の原理や法則を発見し、体系化することを目指す研究開発

研究開発期間: 3年間、継続を認めた場合、2年間延長

平成13年度: 7課題採択 (応募101課題、14.4倍)
継続4課題。平成18年度終了

平成17年度: 6課題採択 (応募113課題、18.8倍)
継続4課題。平成20年10月開始

平成18年度: 5課題採択 (応募 79課題、15.8倍)

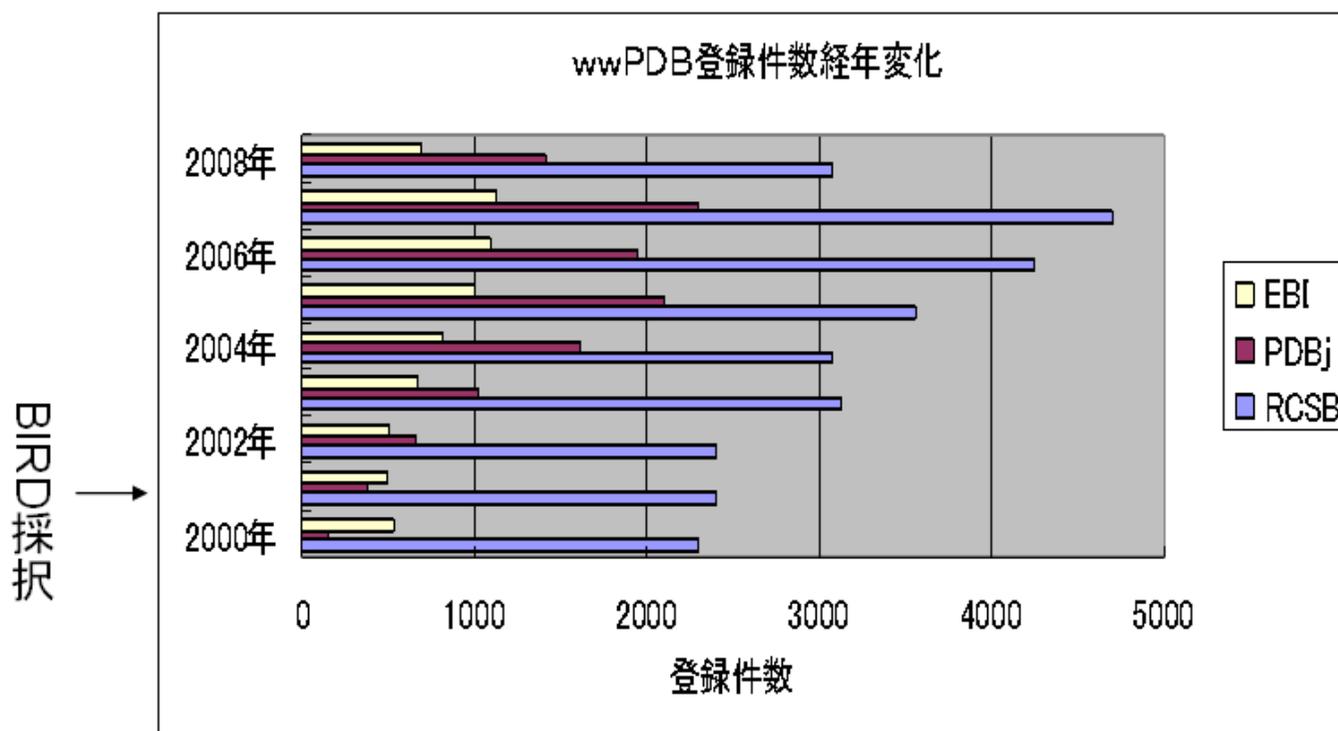
平成19年度: 4課題採択 (応募 45課題、11.3倍)

平成20年度: 採択なし

4. 成果①

・我が国の基盤データベースの支援→世界に誇れる我が国のデータベースへ発展

例1:PDBj(大阪大学蛋白質研究所)wwPDBデータ登録件数の推移
 日本(PDBj)、アメリカ(RCSB)、欧州(EBI)の三極で蛋白質構造データを登録
 2001年のBIRD採択後、PDBjデータ登録件数は増加
 世界全体の約30%を担当(日本、アジア、オセアニア地区)



wwPDB HP (<http://www.wwpdb.org/stats.html>) より2008.10.1現在

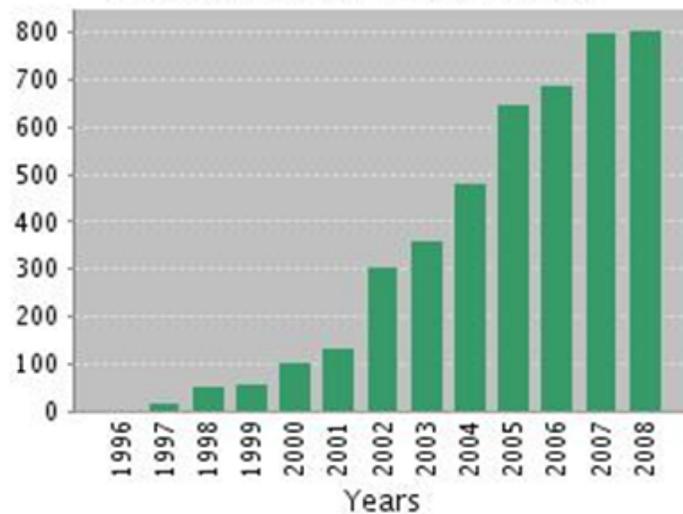
4. 成果②

・我が国の基盤データベースの支援→世界に誇れる我が国のデータベースへ発展

例2:KEGG(京都大学化学研究所)

金久實教授の発表論文の被引用論文数推移

Citations in Each Year



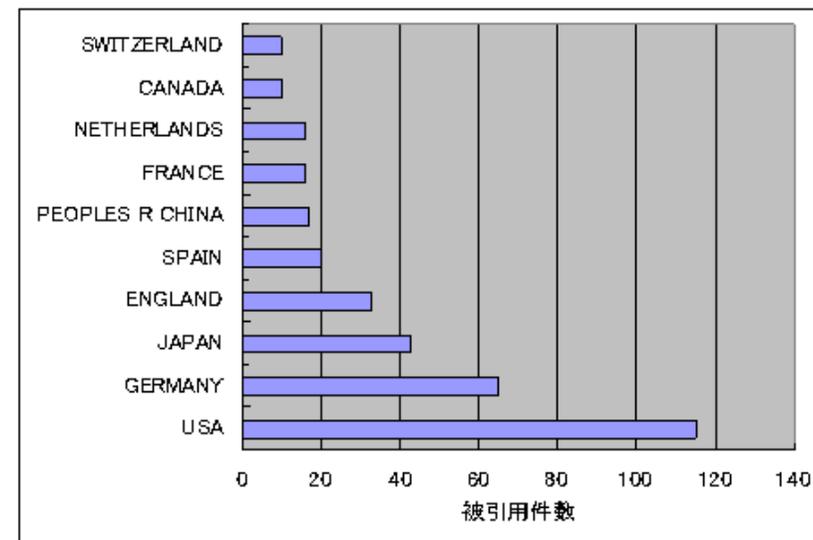
BIRD採択

Web of Science 検索結果

検索条件: AU=(Kanehisa M) 2008.10.1検索

被引用論文著者所属機関所在地

10件以上のみ表示 国外からの引用が86%以上



From genomics to chemical genomics: new developments in KEGG
Source: NUCLEIC ACIDS RESEARCH Volume: 34 Special

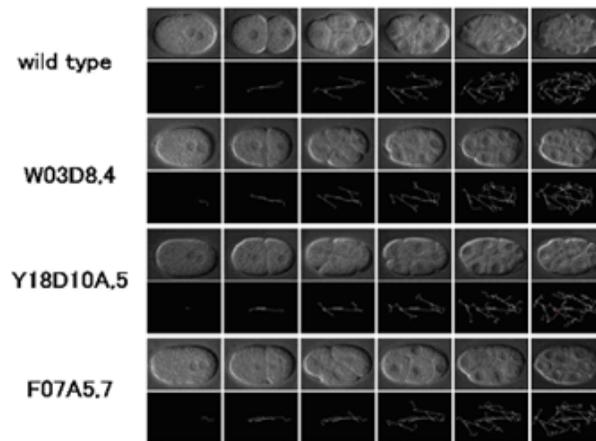
Issue: Sp. Iss. SI Pages: D354-D357 Published: JAN 1 2006

被引用件数: 329件

12

4. 成果③

・形態形成のインフォマティクスの成果→他の生物情報へ応用発展



●「線虫*C. elegans*発生過程のシステム解析」

研究期間: H13.10-16.9

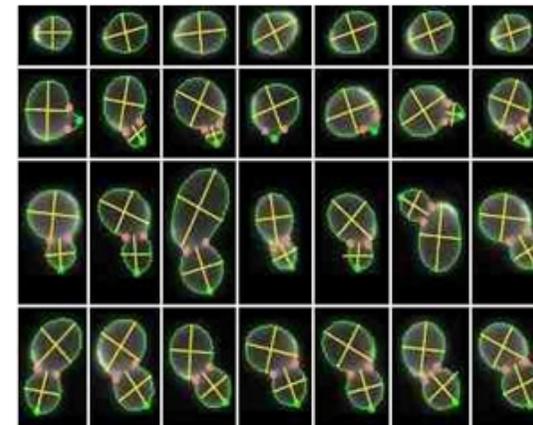
研究機関: 慶應義塾大学

代表研究者: 大浪 修一

線虫の細胞系譜データを収集し、多細胞生物の線虫初期発生システムの計算機科学的解析等を目指す。

このために、大規模突然変異体細胞系譜測定、細胞系譜データを用いた遺伝子機能ネットワーク推定アルゴリズムの開発、シミュレーション技術開発、パラメータ測定技術開発等を行う。

Kimura A, Onami S. (2005). Computer simulations and image processing reveal length-dependent pulling force as the primary mechanism for *C. elegans* male pronuclear migration. *Dev Cell.* 8:765-75.



●「遺伝子破壊株イメージ・マイニング」

研究期間: H13.10-H18.9

研究機関: 東京大学

代表研究者: 森下真一、大矢禎一

細胞のかたちから遺伝子の機能を予測
パンに使用される出芽酵母の遺伝子をひとつひとつ働かなくした4,800株の顕微鏡画像をコンピュータで解析し、形態的な特徴を数値化した。この結果を統計処理したところ、類似の機能を持つ遺伝子が破壊されると、形態も似てくることがわかった。発酵食品を含め、産業に利用できる酵母の遺伝子機能解明のあたらしい手法を提供。

4. 成果④

課題参加研究者の昇格状況 H13採択11課題の追跡調査より

平成13年度採択11課題全参加者数 294名

BIRD課題参加時 研究員～助教授クラス 108名

テニユアの准教授、教授にステップアップ 29名

内訳

研究員、助手、講師 →准教授 14名(2008.7)

研究員、助手、助教授 →教授 15名(2008.7)

参考:

JST さきがけ(研究者個人の独創性を活かした研究を推進)

1991年から2004年までに「さきがけ」研究を終了した研究者

さきがけ参加時 研究員～助教授クラス 396名

ステップアップ 170名

引用: 戦略的創造研究推進事業国際評価報告書2006年8月. 科学技術振興機構

<http://www.jst.go.jp/kisoken/pdf18/seika03-18.pdf>

5. 研究成果情報発信(ゲノムリテラシー講座)

平成13年度～18年度 バイオインフォマティクスの普及活動

バイオインフォマティクスの基礎から応用 講義・実習

テーマ: バイオインフォマティクス入門、ホモロジー検索、SNP解析、perl入門、蛋白質立体構造予測、プロテオミクス、テキストマイニング、E-Cell、生物進化系統樹の作成、細菌の比較ゲノム 等

○バイオインフォマティクス初級者、大学院生、企業研究者等が参加

平成19年度～ バイオインフォマティクス分野の研究情報発信

研究開発課題の代表研究者等が講義・実習、研究内容紹介

テーマ: ゲノムネットとKEGG、蛋白質構造予測の基礎、ケモインフォマティクス、転写制御領域の配列解析入門 等

○受講者層が拡大

例) バイオインフォマティクス授業担当教官、衛生試験場研究者等が参加

(参考) BIRD関連データベース



構造データベース

核酸配列

- [OASYS\(国立遺伝学研究所\)](#)
- [HGS\(JST\)](#)

蛋白質立体構造

- [PDBJ\(大阪大学蛋白質研究所\)](#)
- [eF-site\(東京大学医科学研究所\)](#)
- [eProtS\(大阪大学蛋白質研究所\)](#)
- [Promode\(早稲田大学\)](#)
- [BioMagResBank \(BMRB\)\(大阪大学蛋白質研究所\)](#)
- [GTOP\(国立遺伝学研究所\)](#)

一塩基多型

- [JSNP@東京大学医科学研究所\(JST\)](#)
- [多型情報ネットワーク\(JST\)](#)
- [HapMap](#)

遺伝子地図

- [HAL\(京都大学大学院情報学研究所\)](#)

細胞形態

- [SCMD\(東京大学大学院新領域創成科学研究科\)](#)

総合データベース

- [ヒトゲノム情報統合データベース\(HOWDY@JST\)](#)
- [ホヤプロテイン統合データベース\(CIPRO\)](#)

比較ゲノム

- [真核生物比較ゲノムブラウザー\(JST\)](#)

化合物

- [SILA\(JST\)](#)
- [MassBank\(慶應義塾大学\)](#)



関係性データベース

分子間相互作用

- [KEGG-BRITE\(京都大学化学研究所\)](#)

遺伝子発現

- [CIBEX\(国立遺伝学研究所\)](#)
- [BSD\(国立遺伝学研究所\)](#)
- [PGE\(JST\)](#)



機能データベース

シグナル伝達

- [INOH\(東京大学/産業技術総合研究所\)](#)



オントロジー

- [Ontology\(東京大学\)](#)

その他



地球規模生物多様性情報機構

- [GBIF\(GBIF/JST\)](#)

バイオインフォマティクス推進センター事業



背景

- ①第3期科学技術基本計画
 - ・データベース等知的基盤の戦略的な重点整備を進める
 - ・バイオインフォマティクスなど急速に発展している分野において、機動的な人材の養成・確保を推進する
- ②第3期科学技術基本計画 分野別推進戦略(ライフサイエンス分野)
 - ・重要な研究課題:「生命情報統合化データベースの構築に関する研究開発」
 - バイオインフォマティクスの展開に不可欠なデータベースの高度化・標準化等を促進
 - ・戦略的重点科学技術:「世界最高水準のライフサイエンス基盤整備」ゆるぎないライフサイエンス基盤を整備する
- ③「我が国におけるライフサイエンス分野のデータベース戦略のあり方について」
 - (H18年5月科学技術・学術審議会 研究計画・評価分科会 ライフサイエンス委員会 データベース整備戦略作業部会)
 - ・バイオインフォマティクスの研究開発の促進等が今後取り組むべき課題

目的

ゲノム情報等の生物情報データベースの構築、高度化、活用のための研究開発を行い、研究開発成果を情報発信することにより、情報基盤の整備の一翼を担うとともにライフサイエンス研究のさらなる進展に貢献する。

概要

バイオインフォマティクス推進センター

研究開発の推進

創造的研究開発

- ・情報科学と生物科学の融合型アプローチによる研究開発を支援(公募により実施)
- 研究期間:3年

生物学
医学

情報学

センター運営

統括、副統括
運営委員会

全体的方針
調整
評価

データベースの高度化・標準化等

- ・バイオインフォマティクスに不可欠なデータベースの構築・標準化を支援(公募により実施)
- (H18~22)

代謝 オンコロジー 核酸 蛋白 パスウェイ

生命情報データベース高機能化

成果の公表・発信

- ・最先端の研究者を講師に普及研修会の開催
- ・研究成果生命科学データベースの発信

多型情報

統合データベース連携

- ・ポータルサイトの運営等
- ・連携強化

連携・調整 国際的な連携・協力

世界的協力体制の構築

- ・GBIF等への参画

文科省統合データベース

- ・戦略立案、技術開発等

持続可能なライフサイエンス分野のデータベースの統合・維持・運用の在り方等について検討を行うため、ライフサイエンス情報基盤整備作業部会を設置

データ・活用研究

大学・研究機関

遺伝子間の新原理解明等

知識抽出

利便性向上

産業界

生命情報を基にした新薬の創製等

要望

国民生活

革新的医療、豊かな社会の実現

統合DBタスクフォース（仮称）の設置について

平成20年5月7日
ライフサイエンスPT

1. 主旨

第1期、第2期科学技術基本計画の基に、これまで、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。しかし、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。したがって、これらのデータを利用者の視点に立って統合化し、どの様に効率良く研究者、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いたため、平成17年度から、科学技術連携施策群「生命科学の基礎・基盤」の中で、補完的な課題として「生命科学データベース統合に関する調査研究」（3年間）を実施する等、関係省と連携した取組みを行ってきた。

これまでの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が進められる等、所定の成果が得られたところである。

今後は、さらに次の段階として、補完的課題の成果も活用して、これまで各省で整備されたデータベースの横断的な統合化を行い、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の整備が期待されている。そのためには、関係府省の役割分担を明確にしつつ一層の連携強化を図り、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行動計画を作成し、データベースの統合化を実施していくことが求められる。

以上のことから、連携施策群終了後においても、データベース統合化に向けた各省横断的な取組みを強力に推進して行くため、ライフサイエンスPTのもとに統合DBタスクフォースを設置し、1年間を目処として具体的な検討を行っていくこととする。

2. 検討内容

統合データベースの構築に向けて、以下の項目について検討を行っていく。

- (1) 各省のデータベースの統合化に向けた制度設計や環境整備について
- (2) 利用しやすいデータベースとするための環境整備について
- (3) 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していくための拠点のあり方について
- (4) バイオインフォマティクス研究者や技術者の人材養成のあり方について
- (5) その他、統合データベースの構築に必要な事項について

平成21年度概算要求における科学技術関係施策（ライフサイエンス分野）（継続案件）

（金額の単位：百万円）

施策名	所管	概算 要求額	前年度 予算額	戦略 重点	最重要 政策課 題	競争 的資金	施策の概要	改善・見直し指摘内容	特記事項	昨年度特記内容
【体制整備】										
統合データベース関連事業 (統合データベースプロジェクト)	文部科学省	2,691 (850)	2,782 (1,100)	○			我が国のライフサイエンス関係のデータベースの利便性の向上を図るため、データベースの統合化及び利活用のための基盤技術開発、人材育成等を行い、データベースの統合的活用システムを構築する。	○これまで積み重ねられてきたライフサイエンス分野の研究成果をデータベースとして整備することは、今後の研究開発に向けた基盤整備としての意義は大きく、成果も着実にあげられている。特に、統合データベースプロジェクトは、将来の4省統合のデータベースの基盤技術となるものであり、更なる事業の充実が求められる。 ○こうした期待に応えるためには、将来の4省統合のデータベース化を視野に入れた検討委員会のメンバーの充実や、他省のデータベースとの連携、特に農林水産省のデータベースとの連携を積極的に進める必要がある。また、恒常的な体制として、将来的な統合データベースをどのように整備して行くのかについても、検討が求められる。	○総合科学技術会議としては、最重要課題の一つとして、本事業の成果を重視。 ○恒久的な体制整備に向けて、統合データベースプロジェクトとJSTバイオインフォマティクス研究センター(BIRD)との一本化を目標として具体的検討を進めるべきである。	○継続性をいかに担保するかが重点課題である。 ○JST-BIRDとの連携について、将来的な一本化を含めた検討を行うことが必要である。 ○データベースを作るのみにとどまらず、常に改訂していくことが必要である。
統合データベース関連事業 (バイオインフォマティクス研究センター)	文部科学省 JST	(1,841)	(1,682)	○			膨大なゲノム情報等の解析の格段の効率化・省力化、利用の高度化等を実現するため、革新的なゲノム解析ツールの研究開発等、バイオインフォマティクス研究を推進する。	○将来的な統合データベースの整備に当たっては、統合データベースプロジェクトとバイオインフォマティクス研究事業との一体化が必須であり、そのための検討を含め、我が国としての統合データベースの整備を加速して実施する必要がある。		○統合データベースプロジェクトとの連携について、将来的な一本化を含めた検討を行うことが必要である。
農林水産生物ゲノム情報統合データベースの構築	農林水産省	707	707	○			イネ、カイコ、ブタ等農林水産生物のゲノムや遺伝子の情報等を統合したデータベースを整備し、大学や民間企業等の研究者に提供する。 また、他生物のゲノム情報をもつデータベース機関であるGenbank、EMBL等とリンクし、高精度に遺伝子情報の類似性検索を行うことが出来るシステムを構築する。	○我が国が優れているイネやカイコ等のゲノム情報に関するデータベースとしてはレベルが高く、そうした分野の研究基盤としては有用性が高い。 ○しかし、科学技術の基盤となるデータベースとして拡張していく必要がある。また、他省との統合データベース化に向けた観点からも、より積極的な取組が求められる。 ○農林水産生物のデータベースの整備については、国民の関心が高い、安心・安全な食料の開発や確保にとって重要であることから、将来的な統合データベース化を踏まえ、着実・効率的に実施する必要がある。	○農林水産省の事業に閉鎖されず、他省のデータベースとの統合化に向けて積極的に視野を広げるべき。特に、データベースの項目立てについて、広がりのある名称とすべきである。 ○データベースに収載する内容としても、ゲノム配列以外の分野の更なる充実に努めるべきである。	○統合データベースの構築のみならず、データの更新、修正にも配慮して実施体制を充実させること。
統合データベースプロジェクト	経済産業省	70	70	○			政府全体の“生命科学データベース統合化の取組”の一環として、経済産業省関連の公的資金研究から産出される研究データを、産業上の有用性を評価のうえ、統合化し、産業界等に提供する。	○経済産業省関連機関から産生されるヒト遺伝子関連のデータや、糖鎖や機能性RNA等に関するデータをデータベースとして整備し、その活用を図ることは、医薬品等の産業化にとって重要であり、本データベースと他省庁のデータベースを統合することによる成果も期待されることである。 ○文部科学省の統合データベースとの連携を強化しながら、着実・効率的に実施する必要がある。	○関係省の役割分担を含め、他省とも十分連携し、政府全体で臨床情報も含めた真のデータベース構築に向けて積極的に取り組むべきである。	○積極的に活用される統合化したデータベースを構築し、統合データベースの質の向上や維持には一定の経費が必要であり、予算の確保に努めること。 ○各省とのデータベースの統合を視野に入れて連携を図りながら実施すること。