

「世界最高水準のライフサイエンス基盤整備」

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

中間評価 報告書

平成20年7月

「統合データベースプロジェクト」中間評価委員会

## 目 次

はじめに	1
I プロジェクトの概要	2
II 中間評価の概要	6
III 中間評価結果	
1 全体評価	8
2 個別評価	
(1) 中核機関	10
(2) 分担機関	
①京都大学	11
②東京医科歯科大学グループ	12
③東京大学グループ	13
(3) 補完課題実施機関	
①理化学研究所	13
②産業技術総合研究所	14
③国立遺伝学研究所	14
④九州工業大学	15
おわりに	16
(参考資料)	17
(1) データベース整備戦略作業部会報告書	18
(2) プロジェクト実施体制	50
(3) 平成18年度公募・選定の状況	51
(4) 平成19年度公募・選定の状況	56
(5) 平成19年度補完課題の公募・選定の状況	60
(6) 平成18年度研究運営委員会／戦略作業部会 委員一覧	63
(7) 平成19年度研究運営委員会／作業部会 委員一覧	64
(8) 平成18年度研究成果報告書	65
(9) 中間評価委員会設置要綱	106
(10) 中間評価委員会 委員名簿	107
(11) 中間成果実績一覧	108

## はじめに

現在、我が国は、第3期「科学技術基本計画」（平成18年3月28日閣議決定）の下に、「科学技術創造立国」を目指して諸施策を実施している。

同基本計画においては、「抜本的な科学技術システム改革」が求められており、その中で2010年に世界最高水準を目指してデータベースを含む「知的基盤の戦略的な重点整備」を進めることとされている。同基本計画に基づき、総合科学技術会議が策定したライフサイエンス分野の推進戦略では、戦略重点科学技術の一つとして「世界最高水準のライフサイエンス基盤整備」が掲げられている。

生命情報の統合化データベースはライフサイエンス研究を支える基盤であり、そ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なうため、文部科学省では平成18年度より「ライフサイエンス分野の統合データベース整備事業」（以降、「統合データベースプロジェクト」）を推進している。

平成20年度は、プロジェクト開始から3年目にあたることから、本プロジェクトに係る中間評価を実施し、本報告書を取りまとめた。

一方、総合科学技術会議「生命科学の基礎・基盤」連携施策群（平成17年度及び18年度は「ポストゲノム連携施策群」に相当）においては、同基本計画における分野別推進戦略（ライフサイエンス分野）の戦略重点科学技術の中で、関係府省間の連携を強化して研究体制の構築を行う課題のうち、統合的なデータベース整備に向けた研究開発は、他の研究領域の最も基盤となるものと位置づけて推進している。その一部は、統合的なデータベースの整備に向けた研究を補完的課題として公募・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」、研究代表者：大久保公策教授（国立遺伝学研究所）、平成17年度～平成19年度）が行われた。本調査報告を受けて、今後具体的な制度設計やロードマップ等について平成20年度末を目途に取りまとめ、データベース整備を推進して行く予定である。

本統合データベースプロジェクトは、同連携施策群の施策と密接に連携して進めてきたものであり、関係府省と共に引き続き連携をとって進めて行くところである。

## I プロジェクトの概要

### (目的)

現在、我が国のライフサイエンス分野のデータベースとしてDDBJ<sup>†</sup>（国立遺伝学研究所）、PDBj<sup>‡</sup>（大阪大学蛋白質研究所）、KEGG<sup>§</sup>（京都大学化学研究所）などが国際的に高い評価を受けている一方で、多くのデータベースについて、各機関や各プロジェクトで個々にデータベースが作られ、これらを関連付けて使おうとしたときに使い勝手が悪い、基本的サービスの多くが海外に依存していて、継続的に維持されない等の指摘も寄せられており、国内主要データベースの統合化と継続的な維持方策の必要性が指摘されている。

本プロジェクトは、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、我が国のライフサイエンス関係データベース整備戦略の立案・評価支援、データベース統合化及び利活用のための基盤技術開発、ポータルサイトの整備等を行い、統合化を推進することを目的としており、文部科学省が委託事業として実施しているものである。

### (期待される効果)

本プロジェクトを通じて将来整備される「生命情報の統合化データベース」は、個々の分子生物学研究において蓄積されたデータが戦略的に統合され、付加価値の高いデータベースとして整備されるもので、幅広いライフサイエンス分野の研究者等がこれを活用し、今後の我が国におけるライフサイエンス分野の科学技術の進展に大きく貢献していくことが期待される。

これまでの研究成果の蓄積を網羅的・安定的に利用できるようになり、ライフサイエンス研究の発展に不可欠な基盤となる。

また用法や様式をまたいだ検索機能の開発等による既存データの新たな活用や、産業界・医学関係者等による応用利用を通して新たな知見が得られる。

### (実施期間)

平成18年度～平成22年度（開始後3年度目に中間評価を実施）

### (予算)

平成18年度 2.5億円

---

<sup>†</sup> DDBJ (DNA Data Bank of Japan。国際塩基配列データベースを構築している拠点の一つ。)

<sup>‡</sup> PDBj (Protein Data Bank Japan。生体高分子の立体構造データベースを国際的に統一化されたアーカイブとして運営すると共に、様々な解析ツールを提供。)

<sup>§</sup> KEGG (Kyoto Encyclopedia of Genes and Genomes。遺伝子、タンパク質、また代謝やシグナル伝達などの分子間ネットワークに関する情報を統合したデータベース。)

平成19年度	16億円
平成20年度	11億円

(実施体制)

背景となる国の考え方については、「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 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報告書(参考資料(1))。以下「データベース整備戦略作業部会報告書」という。)を参考としている。

平成18年度は、大学やさまざまな研究機関に蓄積されている生命科学関連の情報を横断的に利用可能とする統合データベースの構築が将来のライフサイエンス研究を支える基盤であるという考えに基づき、フィージビリティ・スタディとしてそ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、「戦略立案支援・実行評価支援」、「統合データベース共通基盤技術開発」及び「ポータルサイト整備・広報、普及啓発」の3つの柱について、それら全ての統合化を推進する機関を公募・選定した。

採択された機関名、課題名『カギカッコ内』は以下のとおりである。

《代表機関》

大学共同利用機関法人情報・システム研究機構

『統合戦略立案評価および統合化基盤技術開発』

《参画機関》

独立行政法人科学技術振興機構『ポータルサイト構築』

国立大学法人九州大学『データベース統合化基盤技術開発』

平成19年度からは、平成18年度の取り組みの成果を基に、統合データベースの開発・整備に向けて本格的に推進すべく、「戦略立案・実行評価」、「統合データベース開発」、「統合データベース支援」の3つの柱について実施することとした。

実施機関の選定にあたっては、事業の3つの柱の全てを担う中核機関、及び中核機関の下に「統合データベース開発」の一部を担う分担機関、さらに統合化を一層加速する観点から、中核機関の示す統合化方針に従い、自ら保有するデータ又はデータベースを中核機関に提供する補完課題実施機関を公募・選定した。

採択された機関名、課題名『カギカッコ内』は以下のとおりである。

【中核機関】

《代表機関》

大学共同利用機関法人情報・システム研究機構

『ライフサイエンス統合データベース開発運用

(戦略立案・実行評価／統合DB開発／統合DB支援)』

《参画機関》

(1) 独立行政法人科学技術振興機構『統合データベース支援：

意見集約システム運用／広報／データベース受入・運用』

- (2) 独立行政法人産業技術総合研究所『統合データベース開発：  
ワークフロー技術を用いた統合DB環境構築』
- (3) 財団法人かずさディー・エヌ・エー研究所『統合データベース開発：  
植物及び植物関連微生物のゲノム情報データベース統合と高度化』
- (4) 国立大学法人東京大学『統合データベース支援：DB構築者の養成』
- (5) 学校法人関西文理総合学園 長浜バイオ大学  
『統合データベース支援：アノテータ・キュレータの教育』
- (6) 国立大学法人お茶の水女子大学  
『統合データベース支援：DB高度利用者の養成』
- (7) 国立大学法人奈良先端科学技術大学院大学『統合データベース開発：  
専門用語辞書管理システムと専門用語解析技術の開発』
- (8) 国立大学法人九州大学『統合データベース開発：多型知識表現技術開発』

#### 《分担機関》

- (1) 国立大学法人京都大学『ライフサイエンス知識の階層化・統合化事業』
- (2) (代表機関) 国立大学法人東京医科歯科大学  
『統合医科学データベース構築方式の開発』  
(参画機関) 国立大学法人大阪大学  
『統合医科学データベース構築方式の開発』
- (3) (代表機関) 国立大学法人東京大学  
『疾患解析から医療応用を実現するDB開発  
(ゲノムワイド関連解析のデータベース開発)』  
(参画機関) 国立大学法人東京大学医学部附属病院  
『疾患解析から医療応用を実現するDB開発  
(リシーケンスDBの開発)』  
学校法人東海大学『疾患解析から医療応用を実現するDB開発  
(ゲノムワイドSNPの統計遺伝学的解析手法の開発)』  
株式会社日立製作所  
『疾患解析から医療応用を実現するDB開発  
(ゲノムワイドSNPの疾患関連解析手法の開発)』

#### 《補完課題実施機関》

- (1) 独立行政法人理化学研究所『植物オミックス情報および蛋白質構造情報』
- (2) 独立行政法人産業技術総合研究所  
『糖鎖修飾情報とその構造解析データの統合  
(糖鎖科学統合データベースの構築)』
- (3) 大学共同利用機関法人情報・システム研究機構 国立遺伝学研究所  
『塩基配列アーカイブのデータベース構築と統合への貢献』
- (4) 国立大学法人九州工業大学『生体分子の熱力学データと構造データの統合』

(公募・選定の経緯)

本プロジェクトにおける実施機関の選定は、研究機関等を公募し、外部有識者により構成される選考委員会において書面および面接ヒアリングによる審査を実施し、選定している。但し、補完課題実施機関の選定においては、書面審査のみにて実施した。

これまでの公募・選定の経緯は以下のとおりである。

・平成18年度公募・選定

平成18年7月3日～同7月31日

一般公募の実施

平成18年8月29日 受託実施機関選考委員会 開催

(ヒアリング審査、採択候補の選定)

・平成19年度公募・選定

(中核機関、分担機関の選定)

平成18年12月27日 第1回受託実施機関選考委員会 開催

(公募課題、公募・審査方針の決定)

平成19年1月11日～同2月8日

一般公募の実施

平成19年2月20日 第2回受託実施機関選考委員会 開催

(ヒアリング対象課題の選考、ヒアリング審査方法の決定)

平成19年3月5日 第3回受託実施機関選考委員会 開催

(ヒアリング実施、採択課題の選考)

(補完課題実施機関の選定)

平成19年7月24日 第1回補完課題選考委員会 開催

(公募課題、公募・審査方針の決定)

平成19年8月1日～同8月30日

一般公募の実施

平成19年9月26日 第2回補完課題選考委員会 開催

(採択課題の選考)

## II 中間評価の概要

### (1) 目的

統合データベースプロジェクトの実施に当たり、各受託機関の実施業務の平成19年度末段階での中間成果を評価するとともに、プロジェクト全体の今後の方針等についての意見交換を行い、今後の事業展開に資することを目的とする。

### (2) 方法

#### ①書面評価および面接評価により行う（総合評価）。

但し、面接評価は中核機関および分担機関を対象に実施する。

#### ②評価対象

- ・プロジェクト全体の評価（以下の各受託機関の状況を踏まえて総合的に評価）
- ・中核機関、分担機関、補完課題実施機関のそれぞれについての評価  
（平成19年度末時点の進捗状況）

#### ③評価者 中間評価委員会（平成20年3月3日設置）

※設置要綱（参考資料（9））、委員一覧（参考資料（10））を参照。

#### ④評価結果の反映

- ・今後の事業展開に向けた検討すべき課題の指摘および助言
- ・次年度の予算配分の重点化等の提案

### (3) 評価事項

#### 1) プロジェクト全体の評価

- ・「データベース整備戦略作業部会報告書」（参考資料（1））を実現するために必要な目標設定になっており、ライフサイエンス分野のニーズに合致したものになっているか。
- ・プロジェクト全体の到達目標の下に、中核機関、分担機関および補完課題の各参画機関の役割に重複等の過不足がなく、全体が組織的に構成されているか。
- ・プロジェクトが中核機関を中心に円滑に推進され、効率的かつ効果的に進捗しており、プロジェクト実施期間内に期待する成果が達成可能であるか。

#### 2) 中核機関の評価

##### ①進捗・達成度（プロジェクト計画の妥当性）

- ・目標達成のために中核機関として必要かつ十分な施策展開が図られ、適切な計画が立案されているか。

##### ②事業の推進体制（プロジェクトマネジメントの妥当性）

###### (i) プロジェクトマネジメントに対する評価

- ・プロジェクトを円滑に遂行するためのマネジメント体制およびプロジェクト進捗管理の仕組みが整理され、日常的に適正に実施されているか。

###### (ii) 連携体制に対する評価

- ・分担機関および補完課題の参画機関との連携を図り、事業を円滑に実施しているか。

③全体総括（プロジェクトの意義、波及効果）

- ・プロジェクト全体を推進する上で、中核機関として適切に機能しているか。
- ・プロジェクト全体の活動や成果が適時に公開されているか。
- ・ユーザニーズに合致しているかどうかを検証するための評価を適宜実施しているか。
- ・中長期的観点から、人材育成などのソフト面での整備も含め、ライフサイエンス分野のデータベース基盤整備の実現に向けて着実に進んでいるか。
- ・オールジャパン体制の意識を持って関係・関連分野のデータを有する機関や研究者等の協力を広く働きかけた提案がなされ、それを実施する体制が整備できているか。

3) 分担機関、補完課題実施機関の評価

①進捗・達成度（課題計画の妥当性）

- ・目標達成のために分担機関又は補完課題実施機関として必要かつ十分な施策展開が図られ、適切な計画が立案されているか。
- ・課題実施期間内に、担当分野における量的並びに質的に実用に足り得るデータ又はデータベースの整備、統合化が実現できる計画となっているか。

②事業の推進体制（課題マネジメントの妥当性）

(i) 課題マネジメントに対する評価

- ・プロジェクトを円滑に遂行するためのマネジメント体制および課題進捗管理の仕組みが整理され、日常的に適正に実施されているか。

(ii) 連携体制に対する評価

- ・中核機関と適宜連携をとりながら、業務計画に則って適切に且つ効率的に実施項目が進捗し、活動や成果が適時に公開されているか。
- ・機関グループ内部の連携を図り、事業を円滑に実施しているか。

③全体総括（課題の意義、波及効果）

- ・プロジェクト全体を推進する上で、分担機関又は補完課題実施機関として適切な役割を担っているか。
- ・オールジャパン体制の意識を持って関係・関連分野のデータを有する機関や研究者等の協力を広く働きかけた提案がなされ、それを実施する体制が整備できているか。

以上の観点で評価を以下のとおり中間評価を実施した。

(4) 委員会開催実績

平成20年3月31日 第1回中間評価委員会 開催

(中間評価の視点、評価の進め方、評価方法の決定)

※中間成果報告（「調査票」）提出期限：平成20年4月21日）

※書面評価実施期間：平成20年4月23日～同年5月12日

平成20年5月21日 第2回中間評価委員会 開催

(面接ヒアリング、面接評価、総合評価)

### Ⅲ 中間評価結果

プロジェクト全体、及び各受託実施機関等について、書面評価及び面接評価による総合評価に基づく中間評価の結果をとりまとめると以下のとおりである。

#### 1 全体評価

##### (1) プロジェクト全体について

###### ○進捗・達成度について

本プロジェクトは、平成18年度から開始されたプロジェクトであり、初年度は中核機関である情報・システム研究機構を中心としてフェージビリティ・スタディを実施した後、平成19年度より分担機関の3機関が、さらに補完課題実施機関の4機関が、平成19年度途中より参画し、本格的な実施体制が構築された。

従って、補完課題実施機関に至っては約半年しか経っていないこともあり、全体的に見て目に見えて評価し得る成果が十分に出ているとは言い難いが、計画に対する進捗は順調以上に進んでいると評価される。

特に中核機関においては正味2年間の研究実施期間に満たないものの、国内の300以上の極めて雑多なライフサイエンス分野のデータベースの横断検索、文献検索、統合ツールなど、見える成果として、短期間で公開できる段階まで到達した点は大いに評価できる。

中核機関としては、プロジェクトの重要性や、その目指すべき方向性、さらにはそのために解決すべき課題については十分把握しており、このままの方向で事業を実施すればライフサイエンス分野のデータベースが現在抱えている問題の多くが解消されるものと期待する。

###### ○事業推進体制について

体制面において、中核機関、分担機関、補完課題実施機関の相互連携の姿や協力体制が、必ずしも円滑に図られておらず、一部で類似した開発項目や、中核機関からの期待に必ずしも符合しない開発項目が、分担機関等で実施されているといった機関間の意思疎通に係る問題が見受けられた。理想的な役割分担や業務計画の実現に向けて、本評価結果を踏まえたより具体的な議論が必要と思われる。

###### ○総論

平成19年度が実質的な立ち上げ段階ということもあり、今後、本格的な連携効果が出てくるものと思われる。しかし、これを加速させるためにも、より中核機関が主導的にプロジェクト全体を管理・運営できるような体制となるよう、後半期は見直すべきである。また、参画機関を含めた中核機関の研究体制を柔軟に見直す裁量を、中核機関に対して与えることが必要である。

さらに研究推進に当たっては、研究運営委員会をより効果的に活用することで、中核機関、分担機関、補完課題実施機関といった現行の組織構造のあり方についても十分に議論し、抜本的な体制の見直しに向けた検討を早急を実施するのが望ましい。

すなわち、中核機関の代表機関である「ライフサイエンス統合データベースセンター」の強力なイニシアチブの下、データベース戦略や方針に沿ってプロジェクト全体が統括・管理され、明確な参画機関各々の役割・機能により事業推進体制を整理することが、世界最高水準の研究基盤整備である統合データベースの実現に向けては不可欠であろう。

## (2) 全般的に見た今後の課題、助言等

### ○本プロジェクトの意義

- ・本プロジェクトは5年間であるが、単なる「プロジェクト」と見るのではなく、「日本国に必要なライフサイエンスデータベース確立の試行」として事業を進めることが必要である。

### ○プロジェクト終了後の今後あり方

- ・本プロジェクトは、日本のライフサイエンス研究の成果を集約させ、そこから新たなライフサイエンスの知識を生み出す場所の構築のため集中的に整備された。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。そのため、プロジェクトが終了した後、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とは言うまでもない。
- ・そのためには、昨年度の総合科学技術会議の優先付けにおける指摘事項（「（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推進センターについては、）一本化を含めた検討を行うことが必要」）\*\*を十分踏まえ、我が国のライフサイエンス基盤データベース（DDBJ、KEGG、PDB等）を支え、推進してきた科学技術振興機構（JST）のバイオインフォマティクス推進センターに主な経費を一本化し、本プロジェクトの中核機関である情報・システム研究機構による戦略立案機能と、研究開発独立行政法人の事業としての「統合データベース」の維持・運用・高度化等を、有機的に連携させるべきである。
- ・また、研究室単位のデータベースを研究費交付と連動させて統合化させる等、将来的な統合化に向けた工夫や、完成後の維持を考慮し、少ない経費で維持できるような仕組みの構築に向けた積極的な議論も望まれる。

### ○プロジェクト内の事業の整理等

- ・「統合データベース」として何をどこまでするのかという全体構想を参画者で共有した上で、現在の事業推進体制の中での分担機関のミッションや、個別のデータベース開発の必要性について再度精査し、中止することも必要である。
- ・特に、汎用的、包括的なデータベースと、特定研究に特化した専門性の高いデータベースとのバランスについては、予め検討を行うことが望ましい。さらに、汎用的

\*\* 総合科学技術会議（第70回、平成19年10月29日）資料1-2「平成20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ついて」（<http://www8.cao.go.jp/cstp/siryu/haihu70/siryu1-2-5.pdf>、p.11 特記事項：

○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。）

なものや、利用者が多く出ると予想されるものについては、商業的なベースを活用する等も検討すべきである。

- ・研究費の投入は新規データベース構築や新しい概念で既存データベースを組み替えるような作業を伴うものに選択的に投入されるべきである。つまり、既存データベースの維持管理については、事業化して民間資金の導入といった点も含めて検討してみてはどうか。
- ・統合データベースを支えて行くために必要とする人材育成は重要であるものの、本プロジェクトで実施している人材育成施策が「データベース整備戦略作業部会報告書」（参考資料（1））において求められている期待やアウトプットに沿っているのかの観点で再度見直すことが必要であり、限られた事業期間、予算内で対応すべき業務の優先順位を付けた上で、抜本的な整理を敢行すべきではないか。

#### ○その他の課題、助言等

- ・中核機関へのバックアップにより、省庁連携統合データベースを促進できる体制を担保すべきである。
- ・他省庁の主導するライフサイエンスのデータベース、特に医療や健康維持に関する情報とこれからどのように連結していくかと言う課題を念頭に置きつつ、検討を進めて行くべきである。
- ・メンバーに医学、医療の事情を熟知した研究者を加えると臨床情報との連結がより有効に行えるであろう。分子の統合データベースからライフサイエンス全体の統合データベースへと発展することを期待する。
- ・市民向けのコンテンツが少し不足しているように感じられる。昨今の市民意識の高まりを考えると、先端研究の内容を噛み砕いて市民向けに開放する作業が必要な時代になっていると強く感じるので、将来的な構想としてこうした観点からのアプローチも期待する。
- ・本データベースの利用者を増やし、利便性を高めるために、広くモニターを募って定期的に報告を求める制度を始めてはどうか。
- ・中核機関や分担機関の東京大学グループでは、計算機資源の不足問題が挙げられている。計算機リソースについて、公共の既施設を無償で（廉価に）使えるような仕組みや、支援する仕掛け、あるいは共用できる方策等について検討すべきである。

## 2 個別評価

### （1）中核機関（代表機関名：情報・システム研究機構、代表者：高木 利久）

#### ○進捗・達成度について

「辞書の整備」、「知識の整理棚」、文献のオープンアクセス化をにらんだ技術開発、日本語雑誌や学会用紙の検索、教育用教材、ポータルサイト作成、データベース受け入れなど様々な試みがなされ、利用者の利便性向上が緒につき、データベースの統合化が具体的に動き出したことは評価される。

またデータベース構築者のために公開されるデータ全てをダウンロード可能とするといった統合に向けた方針は、より多くの利用者と波及効果を期待できることから望ましい方向である。今後、質の高いデータをより多く蓄積するための取り組みが期待される。

我が国における中核機関として短期間にこれだけの進捗を見せていることを十分に評価するべきと考える。こうした取組を確実に進めていくためにも、中核機関に相応しい予算配分の再構成が必要である。

#### ○事業推進体制について

一方、中核機関の役割は、自らが成果を挙げると同時に、分担機関あるいは補完課題実施機関と上手く連携して相乗的に目標を達成することにある。しかし、一部の参画機関は必ずしも中核機関の認識どおりの意識にはなっていないようにも見受けられ、連携が不十分である。これはプロジェクト発足当初において、中核機関に指導、連携に関する権限が明確に与えられていないことによるものと思われる。

#### ○総論

非常に大きなミッションであり、きわめて重要な課題を扱っている。多様な参画機関、分担機関、補完課題実施機関のコントロールが必要であり、現在の中核機関のパワーだけに任せるのは不十分であり、より強いイニシアチブを持たせるべきである。問題と改善策に対しては、研究運営委員会における早急かつ十分な議論を望む。また中核に相応しい予算配分の再構成が必要である。

今後（プロジェクト終了後）統合データベースをどのような体制で進めるかの検討（コンセンサス）が必要である。この予算でできること、できないことを整理しておいた方が良い。

### (2) 分担機関

#### ①京都大学（代表者：金久 實）

##### ○進捗・達成度について

化合物に関するソフトウェア整備、医療医薬品、一般用医薬品に関するデータベースの公開、LinkDBの対象データベースの拡張、などについて妥当な進捗が見られる。また、本統合データベースはゲノムネット<sup>††</sup>とKEGGを分離し、前者のサポートということで、明確な切り分けをしておき、KEGGとのリンクは有機的に進められている。さらに国内（JAPIC<sup>‡‡</sup>、LipidBank）のみならず、国際データベースやIUPAC<sup>§§</sup>との連携も視野に入っており、継続支援は重要である。

<sup>††</sup> ゲノムネット（ゲノム情報を基盤とした新しい生命科学研究と創薬・医療・環境保全への応用を推進するために、京都大学化学研究所バイオインフォマティクスセンターが提供するインターネットサービス。http://www.genome.jp/）

<sup>‡‡</sup> JAPIC（財団法人日本医薬情報センター。国内外の医薬品に関する臨床的に有用な情報を収集・処理・提供することによって、薬剤の臨床使用の適正化を通じて製薬と医療の間のかけ橋の役目を果たすことを目的に設立された公益法人。）

<sup>§§</sup> IUPAC（International Union of Pure and Applied Chemistry. 国際純正・応用化学連合。元素名や化合物名についての国際基準（IUPAC命名法）を制定している国際学術機関。）

### ○事業推進体制について

中核機関との連携について、どのようになされているのか、姿が良く見えない。連携は発足当初に考えられていなかったもので、そのための話し合いで取り入れた医薬品、化合物データベース構築については、配分比率が高過ぎると思われる。

### ○総論

ゲノムネットの医薬品データベースは大学等の研究者または学生にとって、または薬に関心のある一般利用者には有用なデータベースであるかも知れない。しかし、創薬に係る研究者など医薬品企業等の産業界による利用まで考えるならば、JAPIC など廉価で手に入る情報とのリンクではなく、別のあるいは独自の情報とのリンクが必要ではないか。また、当該分野については、便利な有料のデータベースが揃っている点を踏まえつつ、魅力ある検索法を提案するなどの別の方策が必要である。

またプロジェクト全体予算（11億円）の中で、ゲノムネットの DBGET/LinkDB 解析ツールの開発費用が整合性を持って配分されているのか、また補完課題実施機関の産業技術総合研究所が担っている糖鎖関係のデータベースや中核機関が開発している検索エンジンの開発等において、各々の役割分担が曖昧になって整理されていない開発等が見受けられる点を考慮して、予算配分については適正に査定する必要がある。

さらに、当該機関が持つポテンシャルである KEGG の強固な基盤をどのように役立たせるかについてより深い議論が中核機関との間で必要である。

こういった観点も踏まえ、限られた全体予算の中で効果的な成果を上げるため、利用者のニーズに真に応えるためにどうすべきかを研究運営委員会等で議論すべきである。

### ②東京医科歯科大学グループ（代表者：田中 博）

#### ○進捗・達成度について

ライフサイエンスのデータを分子から臨床にいたる多階層の視点からデータベースを構築するという姿勢は評価できる。未だ例数が少ないので評価が難しい。

2つのデータベースをモデル的に構築しようとしている方針は評価できる。しかる後に、そのモデルをどのようにして国内に広めて行くのかについて確たる見通しが無いのが不安である。

特定の医師（場合によっては comedical）が限定しているデータベースとしては適当だが、中核機関の姿勢とそぐわないのではないか。

#### ○事業推進体制について

本データベースの構築に当たっては、疾患記述の標準化など日本全国にインパクトを与える可能性のある重要な事項が散在している他、中核機関との連携の姿が不明確である。

### ○総論

限られた予算規模及び時間で一定の成果を出すためには、取組内容について、優先順

位を考えた上で、フォーカスを絞る必要があり、インフォームドコンセントが必要となるなど、その統合に際して大きな困難が予想される医学情報のデータベース作りに向けて「自力で出来ること」を中心に進めるべきである。つまり、データベース開発を目指すのではなく、中核機関が目指すデータベースの統合化に向けて、臨床関係のデータベースはどう統合していくべきか、何をすべきか、何が重要かといった統合に当たっての課題に対する提言や、フィージビリティ・スタディ的なプロトタイプを小規模に作る等のロールモデルの提供に徹するべきである。

### ③東京大学グループ（代表者：徳永 勝士）

#### ○進捗・達成度について

GWAS (Genome-Wide Association Study) は世界の潮流であり、理化学研究所以外の GWAS 中核として、国内のデータ保有者と連携の下、多数の施設に亘る共同研究をベースとしたオールジャパン体制で進めており、ナショナルプロジェクトに相応しい。到達目標（アジアのハブ）も明快である。このまま継続されるべきである。

データの品質管理、及び標準化を行うという研究内容は評価できるものであり、限られた予算内で期待される達成度を満足している。

#### ○事業推進体制について

また、中核機関の方針に馴染んだマネジメント体制が取られていることも評価される。さらに、標準 SNP データベース、GWAS データベースとも平成 19 年度計画を十分達成しており、成果は着実に上がっている。

#### ○総論

今後は、一層のデータを集めるとともにデータを解析する努力が必要であり、特にこのデータベースにどの位のデポジットがなされ、どう発展させるかが重要である。

疾患解析のデータベース構築の際に発生しうる問題点を整理しつつ、小さくても「モデル的データベースの構築」を示して欲しい。

また、計算機資源、データ解析能力の充実に向けには、中核機関との密接な連携が必要であり、両者の一層の協調と協力体制の確立が必要であるが、計算機資源については、公共的なスーパーコンピュータの利用等のコスト便宜にも配慮すべきである。

### (3) 補完課題実施機関

#### ①理化学研究所（代表者：豊田 哲郎）

理化学研究所の多数あるデータベースを利用しやすいものとするとともに、付加価値を付け統合データベースに組み入れていくテストケースとして、当該取組みの意義は大きい。

主として理化学研究所で行われたシロイヌナズナのオミックス、タンパク 3000 プロジ

エクトからの高等動植物由来タンパク質の構造データ、微生物由来蛋白質構造データなどを用いて、そのアノテーションなどの構築、およびアノテーションシステムの開発運用の研究について、研究体制は適格に構築できており、時間的な問題も考えると、順調な準備段階であると言える。

今後は、分野毎の国内他機関との連携の構築と、理化学研究所内のあらゆるデータを積極的に公開する方向に進んでいただき、中核機関の期待に沿うよう盛り立てていただきたい。

### ②産業技術総合研究所（代表者：成松 久）

中核機関との役割分担が明確になって、最もすっきりした関係になっており、糖鎖関連という特定の領域で情報基盤の整備・有効利用が促進されることは関連分野の活性化にとどまらず、データベース統合化のモデルケースとしても評価できる。

利用者コミュニティが比較的まとまっているので、コミュニティとの意見交換を通して、利便性の高いデータベースを作成することが容易な状況にあると言える。一層の努力を期待する。

この種のデータベースにおいては、素材自体の重要性も考えられ、特に利用者を幅狭く想定しないのであれば、「自前」のデータのダウンロードなど様々な利用者のニーズに備えていただきたい。

今後は、プロジェクト終了までになるべく多くの関連データベースの統合を進めていただくとともに、中核機関と連携して、他の分野とも連携したより上位の統合の具体策を検討していただき、中核機関の期待に沿うよう盛り立てていただきたい。

### ③国立遺伝学研究所（代表者：五條堀 孝）

公開 ftp サイトと WWW サイトに関する開発については、国内主要機関との連携により統合システムの準備段階である。登録処理及び波形表示システムに関する開発についてはシステムの設計、プロトタイプシステムの完成、手法の開発のための情報の調査中の段階である。本プロジェクトが始まって時間が経っていないので、達成度は十分ではないものの概ね妥当である。

ただし、中核機関との連携においては、その目的から鑑みるに、例えば「新しい種類の、あるいは新しい発想に基づくデータベースの開発支援」のような新しい連携のあり方が構築できれば、一層統合データベースへの貢献がなされると考えられる。

今後は、新型 DNA シーケンサの出現により、今後塩基配列データの生産量あるいはその研究対象領域の急速な拡大が予想されるため、単に従来型のトレースアーカイブデータベースを構築するだけでなく、将来を見据えたシステムのあり方を検討していただき、中核機関の期待に沿うよう盛り立てていただきたい。

④九州工業大学（代表者：皿井 明倫）

データベース自体の重要性というよりは、小規模データベースと統合データベースの関係のモデルケースとして重要と思われる。

熱力学データと構造データの統合データベースを構築・提供することによって中核機関に対する保管機能は十分果たされている。XML化、オントロジー調査はサービスの質の向上に相当するが、データベースそのものの今後の改良とともに、中核機関と連携して中・小規模のデータベース構築者に適用しやすいシステムあるいはツールの開発も期待したい。

理化学研究所、大阪大学（PDBj）と打合わせを持ちながら課題を遂行しており、データベース構築も順調であることから課題マネジメントも問題はないと思われる。

今後は、巨大機関ではなく研究室単位のデータベース構築者としての統合データベース構築参加のモデルケースとしての役割を続けていただき、中核機関の期待に沿うよう盛り立てていただきたい。

## おわりに

本プロジェクトは、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、データベース整備戦略の立案・評価支援、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、ライフサイエンス関係データベースの統合的活用システムを構築・運用するという、これまでにない取り組みを行うものである。

平成18年度からの約半年間のフィージビリティ・スタディを経て、平成19年度より本格的な実施体制が敷かれて間もないため、全体的に見て中間評価として目に見えて評価し得る成果を要求することが厳しい中、中核機関である情報・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）を中心に精力的な活動がなされており、計画に対する進捗は順調以上に進んでいると評価したい。引き続き中核機関を核として分担機関ならびに補完課題実施機関との連携、協力をより密接に図ることにより、データベースの整備が順調に進み、いよいよ本格的に事業が加速、推進していく状況である。参画している各機関、研究者の今後の活動に大いに期待するところである。

報告書のまとめとして、今後、本プロジェクトをより強力に推進していくにあたって、留意すべき点を指摘しておきたい。

○中核機関の代表機関である「ライフサイエンス統合データベースセンター」がプロジェクト全体を統括、目標を管理し、参画機関各々の役割・機能の分担責任を明確にして、データベース戦略や方針に沿って、センター長が組織全体を引っ張るリーダーシップ、イニシアチブが思う存分発揮でき、プロジェクト全体をコントロールできるような体制となるよう、後半期は見直すべきである。

また分担機関、補完課題実施機関および中核機関の各参画機関は、「ライフサイエンス統合データベースセンター」との連携をより積極的に働きかけ、本統合データベースプロジェクトを一丸となって盛り立てていただきたい。

○本プロジェクトで築き上げられた統合データベースは、我が国のライフサイエンス上の財産の集積場所として、そこから新たな知識を生み出す場所として、長期的に維持する必要がある。プロジェクトが終了した後の管理運営体制の継続性の担保が必要である。また統合化に向けた仕組みをどのように維持するかという観点から、所在も見えてこない可能性のある研究室単位のデータベースをどのように統合するか等の将来像も必要である。これらの諸課題について研究運営委員会等の場を活用して十分に議論し、検討して解決策を見出していきたい。

**(参考資料)**

- (1) データベース整備戦略作業部会報告書
- (2) プロジェクト実施体制
- (3) 平成18年度公募・選定の状況
- (4) 平成19年度公募・選定の状況
- (5) 平成19年度補完課題の公募・選定の状況
- (6) 平成19年度研究運営委員会／戦略作業部会 委員一覧
- (7) 平成19年度研究運営委員会／作業部会 委員一覧
- (8) 平成18年度研究成果報告書
- (9) 中間評価委員会設置要綱
- (10) 中間評価委員会 委員一覧
- (11) 中間成果実績一覧