

平成 25 年 1 月 17 日

ライフサイエンスデータベース統合推進事業の進捗と今後の方向性について

文部科学省研究振興局
ライフサイエンス課

1. バイオサイエンスデータベースセンター設置から現在までの進捗状況（第一段階）

ライフサイエンス分野における我が国の恒久的かつ一元的な統合データベースの在り方については、平成 21 年 5 月に、総合科学技術会議ライフサイエンス PT により「統合データベースタスクフォース報告書」とりまとめられており、これを受けて、平成 23 年 4 月に、独立行政法人科学技術振興機構（JST）にバイオサイエンスデータベースセンター（NBDC）が設置され、ライフサイエンスデータベース統合推進事業を推進している。

これまで、4 省合同ポータルサイト*の開設、科学研究費補助金等により構築されたデータベース（DB）の公開促進（募集要項にデータ提供協力記載）、人体に由来するデータ共有のためのガイドラインの検討（平成 24 年度中に策定予定）を実施。また、DB 統合化に向けた基盤技術の研究開発等を実施し、各研究分野におけるデータベースの統合を推進している。

*文部科学省、厚生労働省、農林水産省、経済産業省の 4 省合同のポータルサイトを構築・運用。生命科学系 DB に関する様々なサービスを提供している。月間利用者数 約 9 万人、月間ページビュー 約 230 万（2011 年の最大値）

- ・総合目録である「生命科学系 DB カタログ」（約 1,200 件）（ ）内は登録件数。
- ・複数の DB を一度に検索できる「生命科学 DB 横断検索」（約 330 件）
- ・DB 単位でダウンロードが可能な「生命科学系 DB アーカイブ」（約 60 件）

2. 平成 26 年度以降（第二段階）の事業の方向性

上記の内容について、引き続き NBDC が中心となって現行体制で推進するとともに、以下についても新たに取り組む予定。

1) 人材育成

バイオインフォマティクス人材育成推進ネットワーク（仮称）の推進

2) データ量の飛躍的増大への対応

ゲノムコホート研究等との連携

3) JST 情報事業との連携強化

文献 DB、研究者情報 DB との連携

4) 既存の取組の更なる展開

合同ポータルサイトの 4 省を超えた連携、データ及びデータベースの提供を受ける対象事業の拡大、海外関連機関との連携の検討など

あわせて、これらの取組に必要な体制の整備を検討していく。

なお、本事業は、ライフサイエンス分野におけるオールジャパン体制で推進していくことが望まれるため、事業の実施に当たっては、総合科学技術会議の場を使って、関係府省と事業の進捗について情報共有を行う。



統合データベース タスクフォース 報告書の概要

平成21年5月21日
ライフサイエンスPT

統合データベース構築の必要性



【現状】

- ・ ライフサイエンス分野において、ゲノム解析プロジェクトやタンパク3000プロジェクト等多量のデータ蓄積型の研究事業を多数実施。
- ・ 今後のライフサイエンス研究の推進や新たな産業の創出のためには**産生されたデータの活用が不可欠**。
- ・ 現在、産出されたデータについては各研究プロジェクト毎にデータベースを維持・管理。
- ・ そうしたデータベースの結合化は整備途上にあり、研究プロジェクトとして実施。

【統合データベース構築に向けた課題】

- ・ わが国の研究開発基盤のさらなる強化のため、研究の成果として、産出されたデータを利用者の視点に立って統合化し、効率よく研究者、産業界、さらには国民に還元していく、**統合データベースの構築が必要**。
- ・ 恒久的なデータベースの維持・管理の**予算措置がとられていない**ため、プロジェクト終了後に、散逸してしまうことが危惧される。**国家的損失**につながりかねない。

我が国に
一元的かつ
恒久的な
ライフサイエンスの
統合データベースが
必要

【ゲノム・ポストゲノム主要プロジェクトの概要】

ゲノム・ポストゲノム 主要プロジェクト名	年度								プロジェクトの概要
	H12	H13	H14	H15	H16	H17	H18		
文部科学省	プロジェクトデータベースの増大増加								
ゲノムネットワーク	ゲノムネットワーク								遺伝子の発現調節機能に関わる網羅的な解析
タンパク3000	タンパク3000								主要タンパク質約3000種の基本構造及びその機能解明
遺伝子多型研究	遺伝子多型研究								ヒトゲノム遺伝子領域中のSNP関連情報の取得と解析
テラーメイド医療実現化	テラーメイド医療実現化								約30万人のSNPと薬剤の効果、副作用などとの関係解明
理研ゲノム、植物、遺伝子多型	理研ゲノム、植物、遺伝子多型								ヒト、マウス、植物のゲノム、cDNA解析、遺伝子多型解析
バイオインフォマティクス研究	バイオインフォマティクス研究								生命科学分野の基幹データベースの構築・高度化
統合データベース	統合データベース								生命科学分野DB戦略立案支援、ポータルサイト整備
経済産業省									
データベース結合	データベース結合								国内外の有用なヒトゲノム関連情報、解析ソフトの統合的整備
ゲノム情報統合	ゲノム情報統合								約3万のヒトの全長cDNA配列情報の取得と解析
完全長cDNA	完全長cDNA								創業支援のためのゲノム、タンパク、化合物一貫解析技術開発
生物システム制御基盤技術	生物システム制御基盤技術								膜タンパク質及び関連複合体の立体構造・機能解明
生体高分子立体構造	生体高分子立体構造								完全長cDNAの遺伝子発現頻度など多方面からの機能解析
蛋白質機能解析	蛋白質機能解析								ヒトのモデル疾患に関わる遺伝子多型情報の取得と解析
遺伝子多様性モデル解析	遺伝子多様性モデル解析								日本人集団768人に関するSNP15万種のアレル頻度の解析
標準SNP解析	標準SNP解析								
厚生労働省									
疾患ゲノムデータベース	疾患ゲノムデータベース								がん等5疾患のゲノムワイドなSNP解析などのデータベース化
トキシコゲノミクス	トキシコゲノミクス								遺伝子発現解析によるゲノムレベルでの毒性発現機構解明
疾患関連蛋白質	疾患関連蛋白質								主要疾患を対象とした疾患関連たんぱく質の探索、同定
農林水産省									
イネゲノム	イネゲノム								イネゲノム配列の解読および遺伝子の機能解明
家畜ゲノム	家畜ゲノム								ブタのcDNA配列情報、発現頻度、マーカー情報の取得と解析
蚕ゲノム	蚕ゲノム								蚕のゲノム、cDNA配列情報、連鎖地図情報の取得と解析
農林水産生物ゲノム情報統合DB	農林水産生物ゲノム情報統合DB								イネその他農林水産生物統合ゲノムデータベースの整備

<ライフサイエンスPT: 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」大久保公策 発表資料
平成20年5月7日 抜粋>

【欧米の代表的な統合データベース運営機関】

NCBI (National Center for Biotechnology Information) :

- ・ 予算: 約80億円 (2006年)
- ・ 人員: 約400名
- ・ 運営形態: 根拠法に基づきNLM(米国国立医学図書館) 予算で運営

EBI (European Bioinformatics Institute) :

- ・ 予算: 約45億円 (2006年)
- ・ 人員: 約300名
- ・ 運営形態: EMBL*が半分負担、残りは外部資金

【日本の主たるデータベース運営機関】

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

- ・ 予算: 約6億円 (2008年)
- ・ 人員: 約30名
- ・ 運営形態: 文部科学省委託プロジェクトで実施

BIRD ((独) 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) (Institute for Bioinformatics Research and Development)

- ・ 予算: 約17億円 (2008年)
- ・ 人員: 約90名
- ・ 運営形態: JST運営交付金で実施

* EMBL (European Molecular Biology Laboratory):
欧州19か国の出資により設立された分子生物学の研究所

統合データベース事業への取組み



総合科学技術会議における取組み

関係府省の取組み

平成
17年度

- ・科学技術連携施策群¹⁾「生命科学基礎・基盤」の開始（主監 五條堀 孝）
- ・補完的研究課題「生命科学データベース統合に関する調査研究」の実施（研究代表者 大久保 公策）

- ・関係省庁における検討

平成
18年度

- ・第3期科学技術基本計画²⁾の戦略重点科学技術³⁾「世界最高水準のライフサイエンス基盤整備」として統合データベースを位置づけ。
- ・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

- ・文部科学省の統合データベースプロジェクトの開始
- ・農林水産省の統合データベースプロジェクトの開始

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
- ・科学技術連携施策群の最終取りまとめ（終了）

平成
20年度

- ・総合科学技術会議ライフサイエンスPTに「統合DBタスクフォース」の設置（座長 五條堀 孝）

- ・経済産業省の統合データベースプロジェクトの開始。
- ・文部科学省「ライフサイエンス情報基盤整備作業部会」報告書策定。

平成
21年度

- ・「統合DBタスクフォース」の報告書の策定（4月10日）

平成21年度概算要求の優先順位付けにおいて、「文部科学省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推進センター（BIRD）との一体化を目指して、加速して事業を実施する必要がある。」と具体的な取組みを評価

1) 科学技術連携施策群：各府省の縦割りの施策に横串を刺す取り組み
2) 第3期科学技術基本計画：平成18年度から5年間の我が国における科学技術の推進戦略
3) 戦略重点科学技術：特に重要な技術を選択し、研究資金を集中させるべき分野

新しい統合データベースの組織体制(第一段階)



【統合データベース推進本部(仮称)の設置】

- ・「統合データベース推進本部(仮称)」はオールジャパンの視点で、**統合データベースの運営に関する司令塔**として機能。
- ・総合科学技術会議のライフサイエンスPTに附属して設置。

【統合データベース評価委員会(仮称)の設置】

- ・「統合データベース評価委員会(仮称)」は「統合データベースセンター(仮称)」の**運営全般について評価**する機能。
- ・「統合データベース推進本部(仮称)」の諮問機関として設置。

【統合データベースセンター(仮称)の設置】

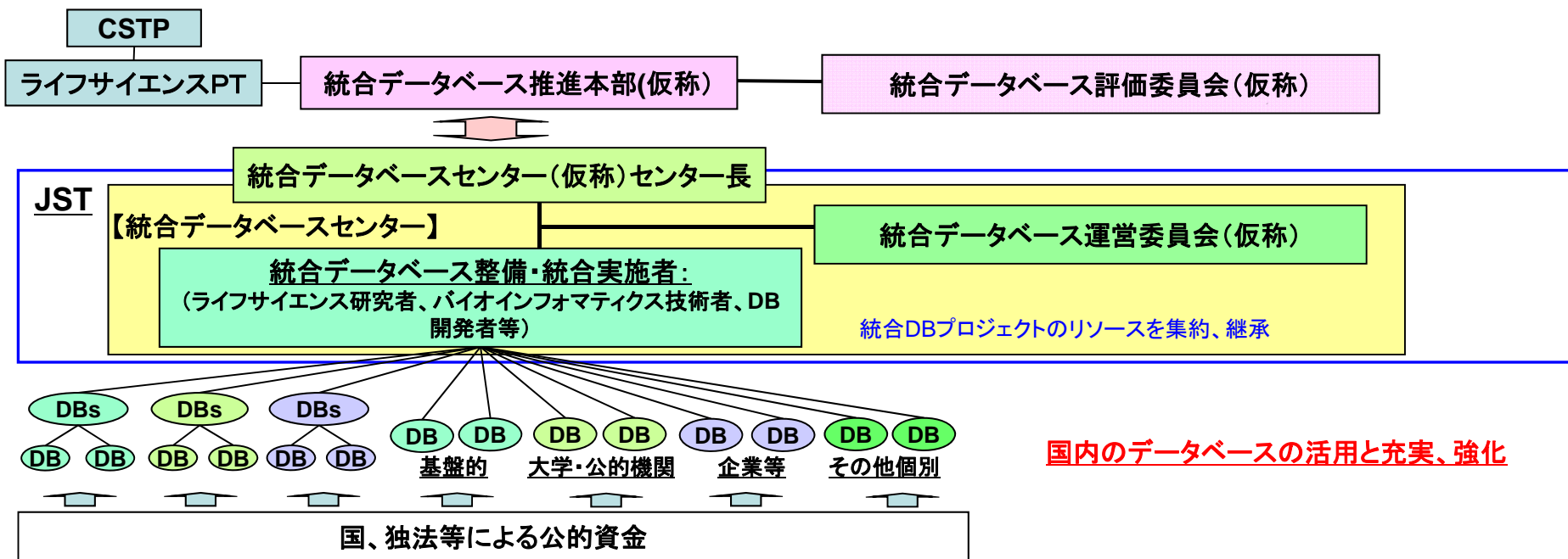
- ・「統合データベースセンター(仮称)」は、**センター長のリーダーシップの下**、データベース統合に必要な調査、標準化、システムの構築・維持・管理、ポータルサイト構築、DBの受入れ・管理・更新、データベースの品質管理等の業務を実施する機能。
- ・大学共同利用機関法人 情報・システム研究機構(ROIS)に設置したライフサイエンス統合データベースセンターを中心に実施してきた文部科学省統合DBプロジェクトと独立行政法人 科学技術振興機構(JST)のバイオインフォマティクス推進センター事業とを一本化し、**JSTの新たな組織として整備**。

【統合データベース運営委員会(仮称)の設置】

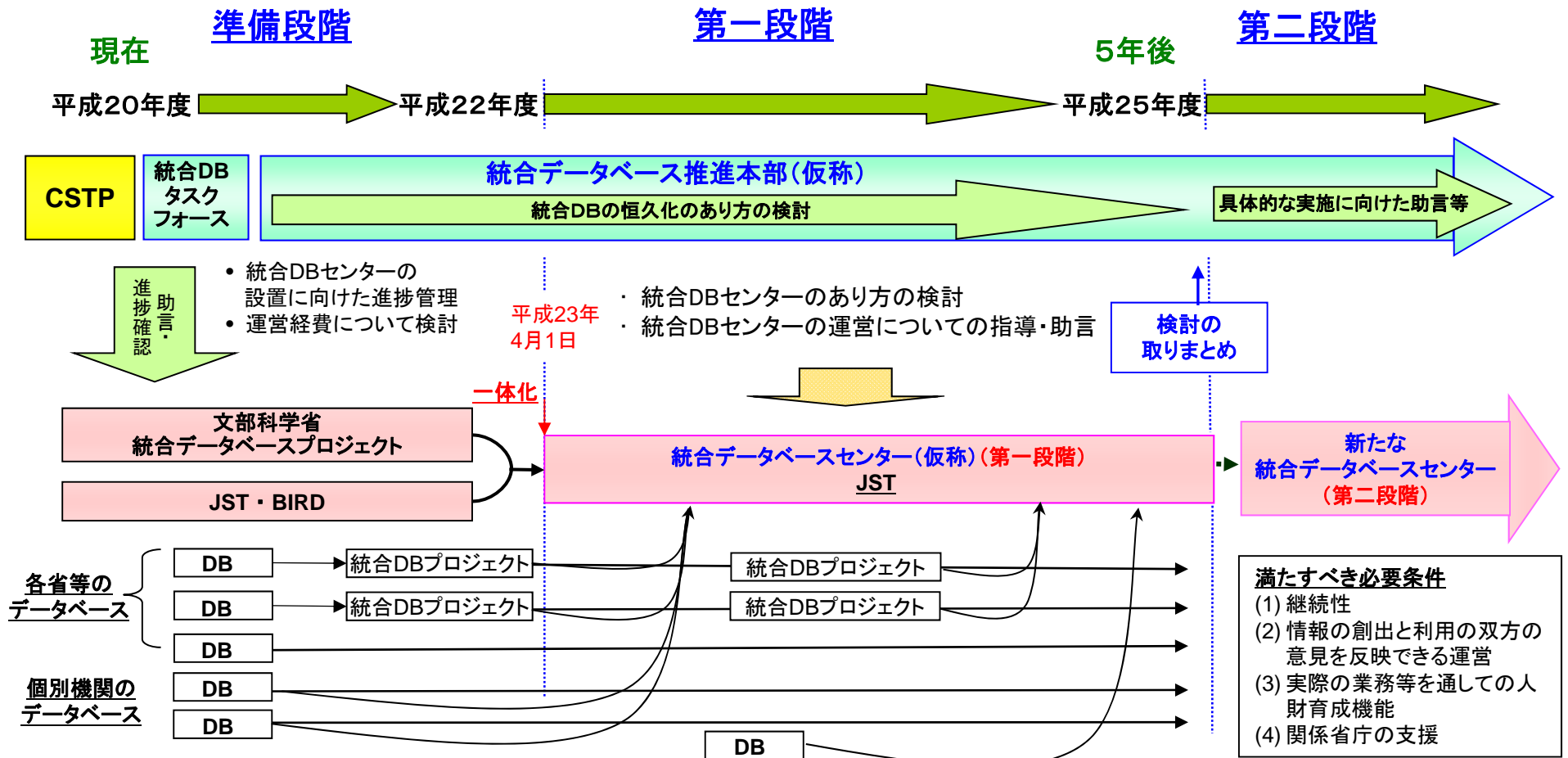
- ・「統合データベースセンター(仮称)」の具体的な事業に関する**協議・調整**の機能。
- ・「統合データベースセンター(仮称)」に設置。

【国内のDBの活用と充実・強化】

これまで関係省庁、研究機関(事業)、企業等で取り組まれてきたDB(既存の各省の統合DBを含む)についても引き続き適切な整備を図る。



統合データベース整備のロードマップ





ライフサイエンス分野の統合データベース整備の 第二段階のあり方について(報告)

(概要)

平成 24 年 7 月

独立行政法人 科学技術振興機構

バイオサイエンスデータベースセンター運営委員会



ライフサイエンス分野の統合データベース整備の第二段階のあり方について(報告)(概要) その1

1. 第一段階(平成23年度～25年度)
 ライフサイエンス分野のデータベース統合を進めることによってデータの価値を最大化することを目指し、平成23年4月にJSTにバイオサイエンスデータベースセンター(NBDC)が設置された。

2. 第二段階(平成26年度以降)
 引き続きJST内の組織として更に充実強化を図り組織を編成していくこと
 分野的に近い機関に移行すること
 新たな機関を設置すること
 などの実現可能性について、総合科学技術会議の場において今後検討される予定

しかし、現在の国を取り巻く状況を踏まえると、
 については第二段階が始まる平成26年度までに実現可能な方策はなく、
 については、NBDCがスタートしてから3年で分野的に近い別の機関に移行することも困難である
 の形で引き続きJSTが引き受け、恒久化を図ることが、現時点では現実的な選択であると考えられる。

第一段階の体制
 予算 約16億円、人数 約20名

総合科学技術会議の助言の下、オールジャパンでのデータベース統合を進める

1) 戦略の立案
 2) ポータルサイトの構築・運用
 3) データベース統合化基盤技術の研究開発
 4) バイオ関連データベース統合化の推進

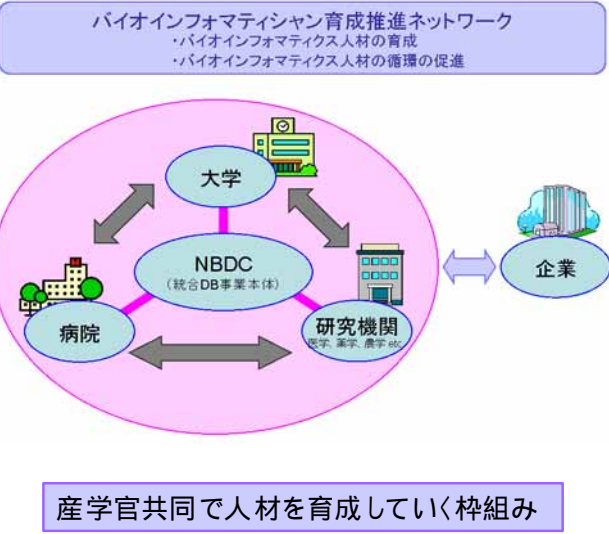


第二段階の体制
 予算 約30億円 人数 約100名
 JSTとROISによる一体的運営

【新たな要因】
 ・ゲノムコホート研究への対応
 ・ビッグデータへの対応
 (次世代シーケンサ、イメージデータ...)

ライフサイエンス分野の統合データベース整備の第二段階のあり方について(報告)(概要) その2

3. バイオインフォマティクス分野の人材育成 = 我が国全体で取り組むべき課題



- ・ ライフサイエンス分野のデータベースを活用するには、データベースの作成、ならびに利用の場においてバイオインフォマティクスが不可欠。
- ・ バイオインフォマティクス育成推進ネットワークではNBDCがコアとなり、バイオインフォマティクス研究を行う大学や研究機関と共にバイオインフォマティクス育成推進ネットワークを形成する。
- ・ 「人材の育成」と「人材の循環の促進」の2つの機能を持つことにより、人材育成とキャリアパスの確立の2つを実現することを目指す。
- ・ 人材の育成では、情報科学の知識を持った人材が、医学、薬学、農学などの様々な知識を身につける、あるいは、逆に医学等の知識を持った人材が情報科学の知識を獲得することによって「I型人材」となり、それぞれの分野で活躍できるような教育システムを構築する。
- ・ 人材の循環の促進では、ネットワーク内で人材が一定年限毎に循環することにより、スキルを身につけ、経験を積み、大学、研究機関や企業などで管理職クラスのポジションにつけるような仕組みを立ち上げる。

NBDC 運営委員会 委員名簿

委員長

末松 誠 慶應義塾大学 医学部長

委員(五十音順)

春日 雅人 独立行政法人 国立国際医療研究センター 総長

金岡 昌治 大日本住友製薬株式会社
常務執行役員 研究本部長 兼 知的財産・IT企画
推進担当

久原 哲 国立大学法人 九州大学大学院
農学研究院 生命機能科学部門 教授

小原 雄治 大学共同利用機関法人 情報・システム研究機構 理事
国立遺伝学研究所 所長

篠崎 和子 国立大学法人 東京大学大学院
農学生命科学研究科 応用生命化学専攻 教授

城石 俊彦 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 系統生物研究センター センター長
哺乳動物遺伝研究室 教授

大力 修 新日鉄ソリューションズ株式会社
フェロー（常務取締役待遇）

武田 英明 大学共同利用機関法人 情報・システム研究機構
国立情報学研究所 情報学プリンシプル研究系 教授

松田 秀雄 国立大学法人 大阪大学大学院
情報科学研究科 バイオ情報工学専攻 教授

NBDC 運営委員会 開催実績

第1回 平成23年10月11日

第2回 平成23年12月5日

第3回 平成24年2月29日

第4回 平成24年5月11日

第5回 平成24年6月5日