




# 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

ライフイノベーション戦略協議会(第5回)  
平成24年8月30日

 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 (NBDC)

<http://biosciencedbc.jp>

# 1. ライフサイエンスにおける情報爆発

## ● データ爆発

- 超高速ゲノム配列決定装置 100億文字/台/日
- コンピューターの性能の伸び(ムーアの法則)を凌駕
- 1万を超える生物のゲノムプロジェクト進行中
- 画像、動画データも急増

## ● 知識爆発

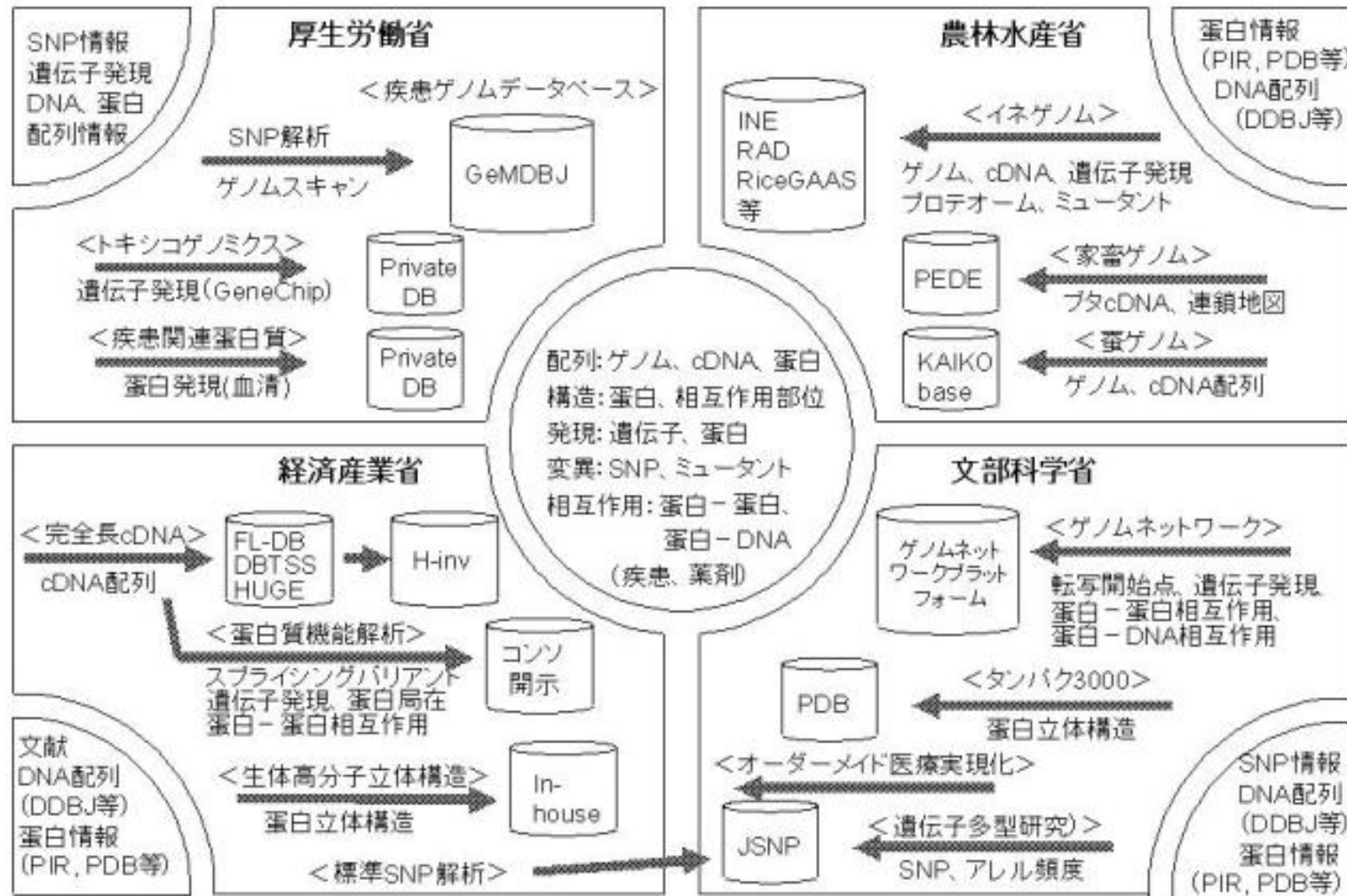
- 論文数 2,000万件
- フルペーパーも利用可

## ● データベース爆発

- 世界 1万
- 日本 1千
- 解析ツール 2千
- 内容も多様化

ライフサイエンスは  
**peta (10<sup>15</sup>=千兆)**オーダー  
 のデータ量の時代に突入

## 2. データベースを指向した国家プロジェクト



科学技術振興調整費  
 科学技術連携施策群の効果的・効率的な推進  
 「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～19年度)より



### 3. ライフ分野におけるデータの共有・統合の重要性

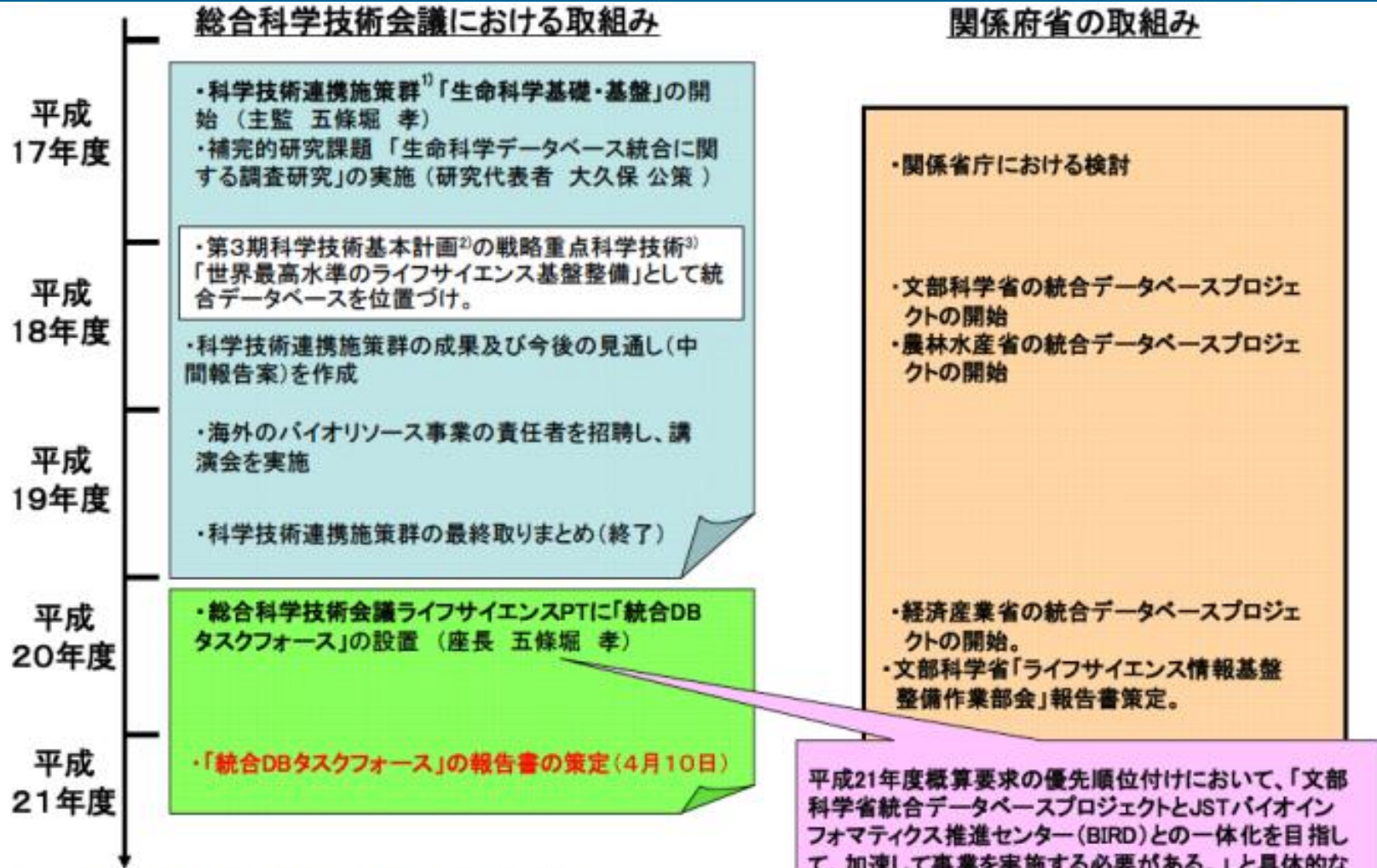
- データベースを指向した種々の大規模プロジェクト
  - 多額の研究費がデータに化けている
- 仮説駆動型からデータ駆動型(データ中心科学)へ
  - 法人化、バイドール法 → データの囲い込み
- データや知識の分断、断片化
- イノベーションには共有・統合化必要
  - ブラウズ(検索)や一部のデータ公開では不十分
  - データ生産者以外の不特定多数のイノベーターの存在
  - 再利用、転用、マッシュアップが自由にできることが不可欠



## 4. 我が国におけるライフサイエンスDBの問題点

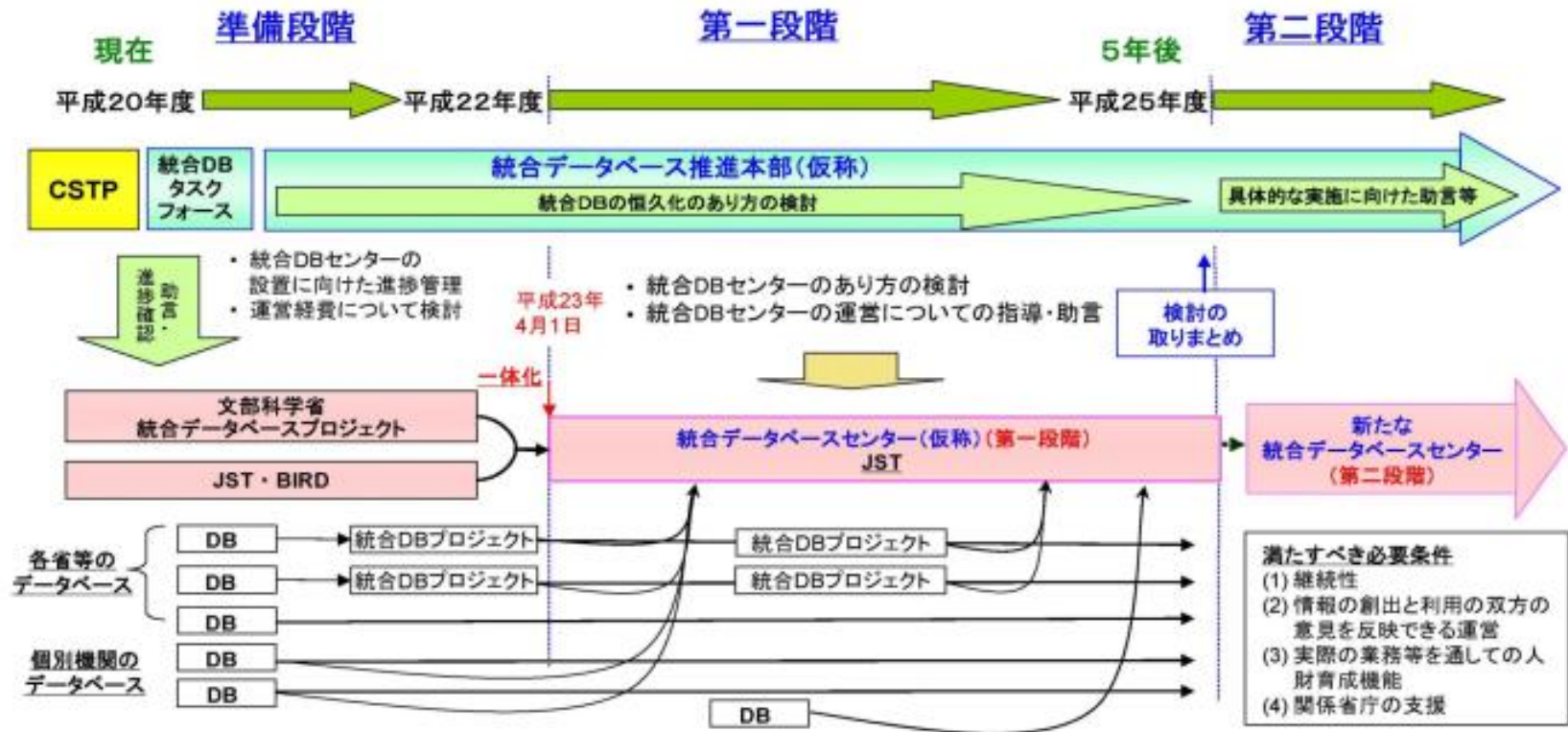
- DBが散在していて所在情報や利用法が分からない
  - 我が国で1,000、世界で10,000ものDB
  - 似たようなものがいくつもありどれを使ってよいか分からない
- DBやDBのエントリに信頼性の高い注釈がついていない
  - DB構築、維持を行える人材不足、DB構築への評価の低さ
- 大型プロジェクトの成果公開、共有が不十分
  - 我が国のライフ関係予算は年間3,000億円以上
  - 公開されているものもプロジェクトが終了すると更新ストップ
- ばらばらに構築・管理されていて検索・解析・応用が困難
  - 現在の統合化は分子レベルで行われていて医療などへの応用困難
  - 日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
- **我が国にはDB戦略、中核センターがない**

# 5. 統合データベース事業への取り組み



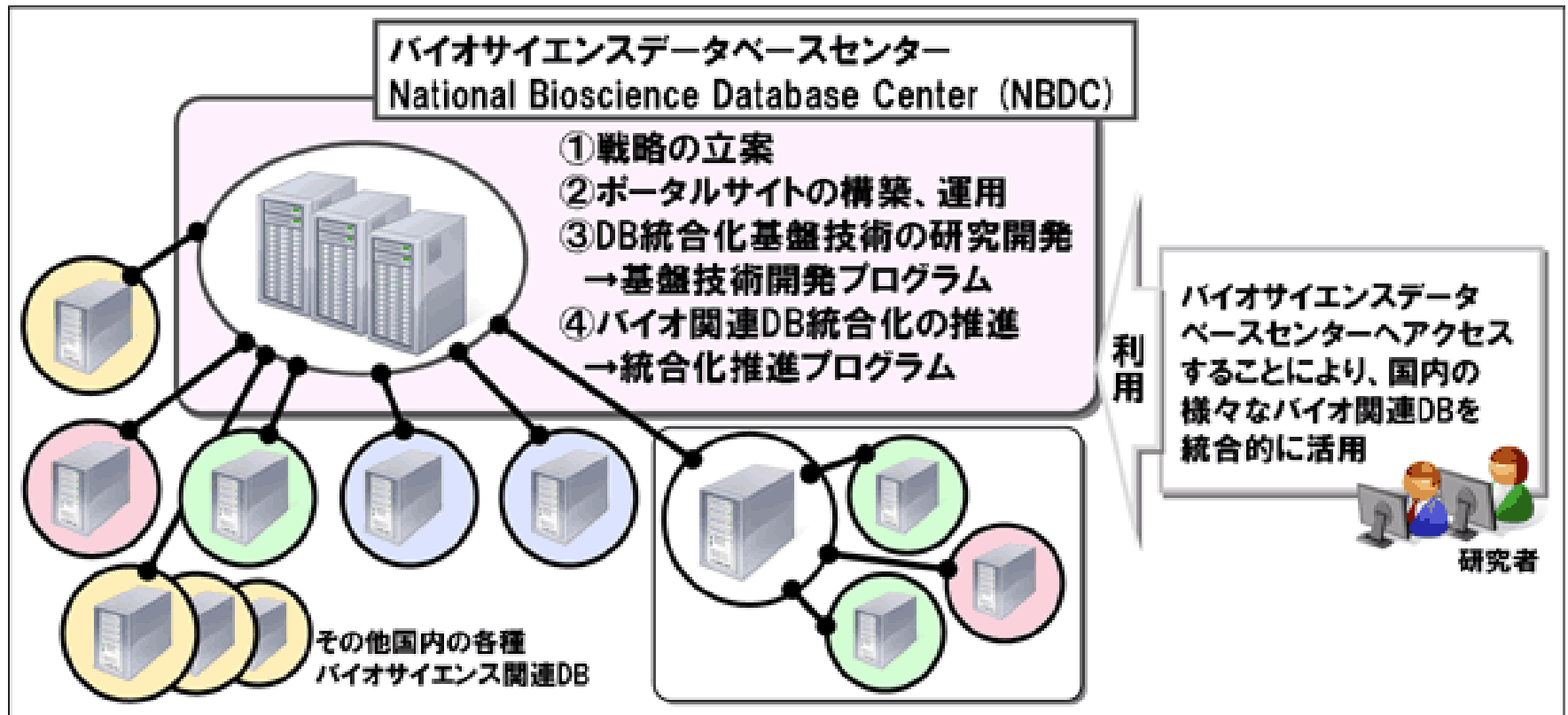
1) 科学技術連携施策群:各府省の縦割りの施策に横串を刺す取り組み  
 2) 第3期科学技術基本計画:平成18年度から5年間の我が国における科学技術の推進戦略  
 3) 戦略重点科学技術:特に重要な技術を選択し、研究資金を集中させるべき分野

# 6. 統合データベース整備のロードマップ



# 7. バイオサイエンスデータベースセンター発足

平成23年4月 バイオサイエンスデータベースセンター(NBDC)発足





# 8. NBDCの体制

