

2015年9月29日  
数学イノベーション委員会

# バイオインフォマティクスの人材育成 -これまでの取り組みと今後の課題-

東京大学大学院理学系研究科  
生物科学専攻

高木利久

# ゲノム科学とバイオインフォマティクスの歴史

地図作成  
配列決定(微生物)

基盤的データベース  
ホモロジー検索

配列決定(ヒトなど)  
SNP、比較ゲノム  
遺伝子の発現解析

遺伝子発見・アセンブリ  
統計解析、進化解析  
クラスタリング、知識発見

構造ゲノム科学  
相互作用解析  
機能ゲノム学  
システムの理解

機能予測、創薬支援  
ネットワーク・パスウェイ推定  
テキストマイニング、オントロジー  
シミュレーション、ネットワーク解析

ゲノム科学は基礎生物学から医学、薬学、農学などへ展開<sub>2</sub>

# バイオインフォマティクス

統合データベース  
ホモロジー検索  
プロモータ予測  
比較ゲノム  
遺伝子発見・アセンブリ  
統計解析、進化解析  
クラスタリング、知識発見  
機能予測、創薬  
ネットワークの推定  
オントロジー  
シミュレーション

# 情報科学

多変量解析、統計モデル  
確率モデル、ベイズ統計  
最適化、情報理論  
アルゴリズムとデータ構造  
形式言語とオートマトン  
文字列処理、数値計算  
自然言語処理、情報検索  
ウェブ検索、セマンティックウェブ  
データベース、プログラム言語  
インターネット、クラウド  
暗号化、匿名化  
並列処理、分散処理  
機械学習、深層学習  
グラフ、画像処理  
推論、探索、知識処理

# 最近のバイオインフォマティクスを取り巻く状況

- NGS(次世代ゲノムシーケンサー)
- ポスト・ポストゲノムはゲノム
- ビッグデータ(ペタバイトオーダー)
- イメージインフォマティクス
- 定量データ(生命動態科学)
- パーソナルゲノム、臨床データ
- ゲノムコホート、ゲノム医療
- 資源植物の分子デザイン

# 大規模ゲノムコホート研究 (ヒト生命情報統合研究)

収集・整理・共有・統合・分析

形態学的データ

疾患罹患情報

生理学的データ

オミックス

ゲノム

タンパク質

文献

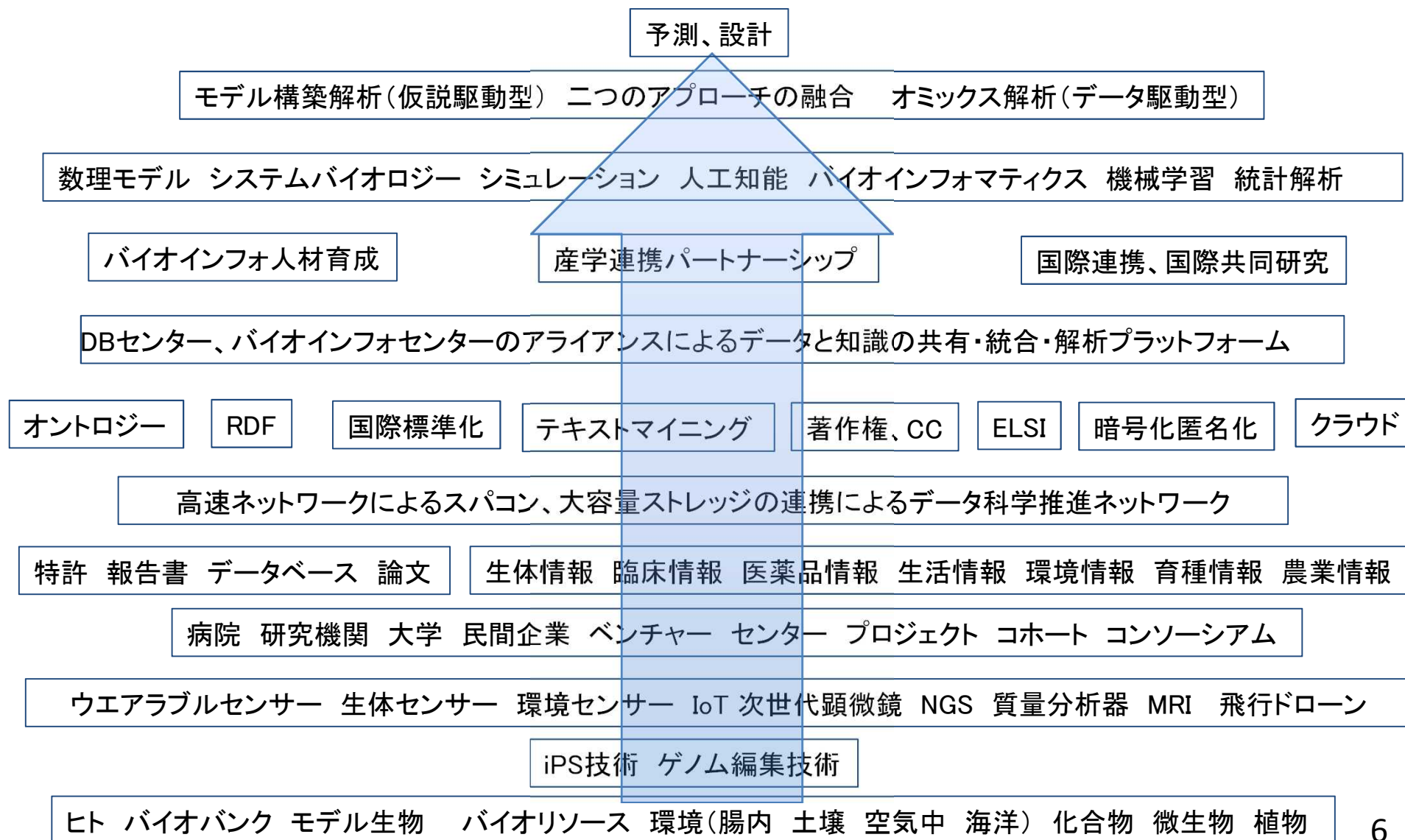
生活習慣情報

医薬品情報

環境情報

# データの産出、共有、統合、解析の流れ

健康長寿社会 予防医学 ゲノム医療 高効率高機能農業



# バイオインフォマティクス人材育成

- 生物情報学適塾
- JST-BIRDでの取り組み
- 科研費(ゲノム情報、ゲノム特定)での取り組み
- 科学技術振興調整費新興分野人材養成
- いくつかの大学における学科、専攻の設置
- 21世紀COE、GCOE
- 統合データベースプロジェクト、JST-NBDC
- 博士課程教育リーディングプログラム
  
- (参考)日本バイオインフォマティクス学会より
  - バイオインフォマティクス教育カリキュラム(2002,2006)
  - <https://www.jsbi.org/about/curriculum/>

# 生物情報科学適塾

- 開催年：2001年春、2002年春
- 場所：国際高等研究所
- 学生：十数名（全国から公募）
- 講師：約20名
- 講義形式：20日間の合宿形式で講義と実習
- テーマ：  
配列解析、発現解析、相互作用解析、  
タンパク質立体構造解析、進化解析、並列処理、  
細胞シミュレーション、パスウェイ解析、  
データベース、テキストマイニング、など

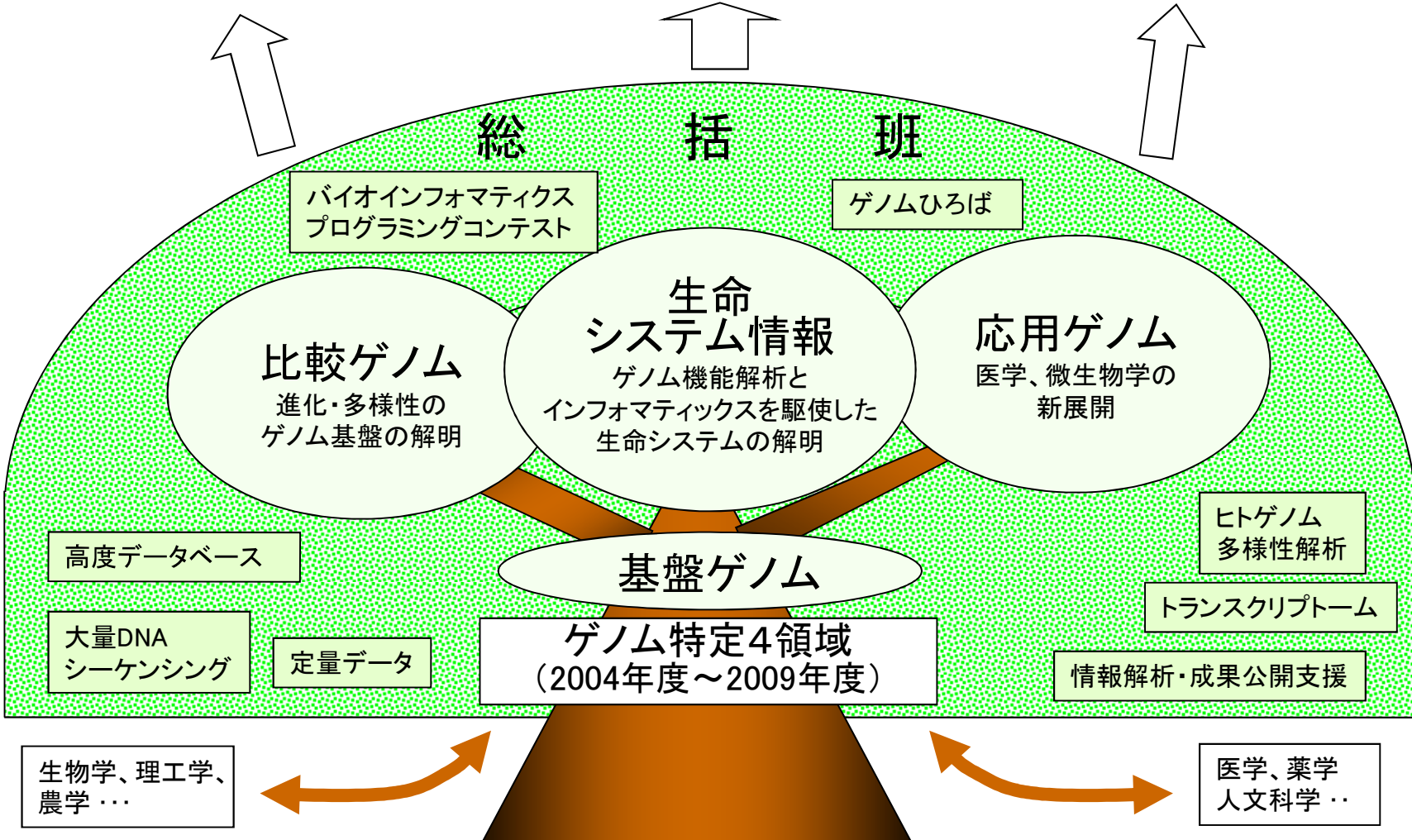


# 2005年頃までの人材育成取り組み

- 1991 科研費(松原新プロ、金久重点)
- 1991 東大医科研ヒトゲノム解析センター設立  
(当初設置の3講座のうち2講座がバイオインフォ系)
- 1996 科研費(榊重点)
- 2000 ミレニアムプロジェクト(科研費ゲノム特定4領域)
- 2000 ヒトゲノム概要配列決定(論文は2001に)
- 2000 科学技術会議ライフサイエンス部会ゲノム科学委員会  
「ゲノム情報科学におけるわが国の戦略について」  
人材養成、データベース、情報解析の3つの推進策
- 2001～2002 生物情報学適塾
- 2001 JSTにBIRD(バイオインフォ推進センター)設立  
ゲノムリテラシー講座開始
- 2001 科学技術振興調整費新興分野人材養成
- 2003 ヒトゲノム完全配列決定
- 2005 科研費ゲノム特定4領域



ゲノムから生命・生物システムの解明、医療・生活の向上へ



ミレニアムプロジェクト(特定領域研究4領域)(2000-2004) 「統合ゲノム」「ゲノム医科学」「ゲノム生物学」「ゲノム情報科学」
特定領域研究「ゲノムサイエンス」(1996-2000)
創成的基礎研究「ヒト・ゲノム解析研究」、重点領域研究「ゲノム情報」(1991-1995)
総合研究「ヒトゲノムの推進に関する研究」(1989-1990)

科研費ゲノム研究の歩み

# 科研費での取り組み

- ゲノム情報チュートリアル(金久重点)
- 学術振興会特別枠(ミレニアムプロジェクト)
- 合同班会議(ミレニアム、ゲノム特定)
- インフォマティクス研究者と医学研究者の交流会(ゲノム特定)
- ゲノム解析支援(ゲノム支援)

# 科学技術振興調整費新興分野人材養成 バイオインフォマティクス、バイオスタティスティクス

## 2001年度～

- 生物情報科学学部教育特別ユニット(東大)
- 産総研 生命情報科学人材養成コース(産総研)
- システム生物学者育成プログラム(慶應大)
- 蛋白質機能予測学人材養成ユニット(奈良先端大)

## 2002年度～

- ゲノム情報科学研究教育機構(京大)
- クリニカルバイオインフォマティクス人材養成ユニット(東大)

## 2003年度～

- クリニカルバイオスタテスティックスコア人材養成ユニット(久留米大)
- システム生命科学人材養成ユニット(九大)

## 2004年度～

- クリニカル・ゲノム・インフォマティクス(神戸大)
- 農学生命情報科学の大学院教育研究ユニット(東大)

## 2005年度～

- クリニカルバイオメディカル情報科学マスターコース(東海大)
- バイオ医療オミックス情報学人材養成プログラム(東京医科歯科大)

# 東大・理・生物情報科学科

- 科学技術振興調整費新興人材養成(2001～2005)
  - 生物情報科学学部教育特別プログラム
  - 夏休み、土曜日、夕方などを利用して講義、演習
  - 3年生9科目、4年生8科目、実験2科目
  - 理学部長より修了書発行
- 大学院新領域創成科学研究科情報生命科学専攻(2003～)
- 生物情報科学学部教育プログラム(2006～2008)
  - 修了生(2003～2006) 70名
- 生物情報科学科(2007～)
  - 学生定員10名
  - 2009より3年生受け入れ
  - 教授2、助教授1、助手1(総長室より)、その後少し増員
  - 理・生化、新領域・情報生命、情報理工・コンピュータの協力

# 東大・新領域・情報生命科学専攻

(2015よりメディカルゲノム専攻と合併してメディカル情報生命に)

- 2003年 新領域創成科学研究科(柏キャンパス)に
- 入学定員:修士24名、博士11名
- 教員:教授6名、准教授1名、助教1名
- 入試:選択問題(生物系、情報系どちらかでOK)
- 講義:生物系、情報系 両方
- 演習、実験:生物系、情報系 両方
- プログラミングは初心者向けと熟練者向けの2つ
- これまでは受験生少なかった
- 就職先は、さまざま(バイオインフォ以外多い)

# ライフサイエンス分野の統合DBプロジェクト

- 内閣府主導の統合データベースプロジェクト(2006～)
  - 文科省、経産省、農水省、厚労省で実施
  - 2011年12月に四省連携のポータルサイト
- 文科省の統合データベースプロジェクト(2006～)
  - データを公共財化・コモンズ化し、その価値を最大化
  - 中核センターの設立
    - 2007～ライフサイエンス統合データベースセンターDBCLS
    - 2011～バイオサイエンスデータベースセンターNBDC
- データベース活用人材の育成
  - RAの雇用、キュレータの人材育成
  - 教材(動画など)の作成、公開
  - 種々の講習会の開催(地方開催も)
  - 必要人材の整理(アンケート、インタビュー)
    - どんな人材が不足なのか、その解消に必要なことは？
    - <http://biosciencedbc.jp/gadget/chousa/Questionnaire.pdf>
  - 人材育成策の提言

No.	カテゴリー	能力
1	基礎/応用研究者 (ドライ)	自分で生物の問題を発見し、定式化し、必要に応じて新規のアルゴリズム、情報技術や DB を開発し、問題を解くことができる。
2	基礎/応用研究者 (ドライ)	新しい情報技術、DB、アルゴリズムを開発できる。 生物系の研究者と共同研究して問題を解ける。
3	基礎/応用研究者 (ドライ)	既存の情報技術、DB を使って問題を解ける。 生物系の研究者と共同研究して問題を解ける。
4	基礎/応用研究者 (ドライ+ウェット)	自分でウェットの研究開発を行い、新しい情報技術、DB、アルゴリズムを開発できる。
5	基礎/応用研究者 (ドライ+ウェット)	自分でウェットの研究開発を行い、既存の情報技術、DB を使って問題を解ける。
6	基礎/応用研究者 (ウェット)	自分で生物の問題を発見したり、定式化したりできる。 情報系の研究者と共同研究して問題を解ける。
7	基礎/応用研究者 (ウェット)	自分で生物の問題を発見したり、定式化したりできる。 情報系の企業にデータの解析を依頼して問題を解ける。
8	支援的研究者 (プログラマー)	カテゴリー1, 2, 3, 4, 5の研究者と協力して、プログラムを作り、支援的な研究開発ができる。
9	支援的研究者	ツールや DB を使ってカテゴリー4, 5, 6, 7の研究者の支援的研究ができる。
10	支援的研究者 (アノテータ、キュレータ)	カテゴリー1, 2, 3, 4, 5, 6, 7の研究者と協力して、データのアノテーション、DB のキュレーションなどの研究開発ができる。
11	支援者(SE)	DB や情報インフラの管理を通じて研究支援ができる。



# 海外におけるバイオインフォマティクス教育

大学等におけるバイオインフォマティクス教育プログラム数

エリア	取得できる学位等			
	学士	修士	博士	その他
米国及びカナダ	30	22	14	10
欧州	4	49	10	0
アジア	1	9	1	0
その他 (アフリカ、中東、豪州、南米)	4	4	0	0

<http://www.iscb.org/iscb-degree-certificate-programs>より集計

# これまでの総括とこれからの課題

- 本格的な教育体制の不足、未整備
  - これまでも多くの人材育成の取り組みあり
  - いずれも小規模で単発的なもの多し
  - 学ぶ学生の負担が大きいのに、得るものが少ない？
  - 大学院教育では遅い？
  - 教員の負担も大きい
  - 必要な人材像の具体化、明確化が必要
  - 人材は育ったが、需要の伸びが大きかった！
- キャリアパスとポジションの問題
  - アカデミックおよび産業界でのポジション不足
  - さまざまな人材に応じたキャリアパスの提供
  - 支援人材、融合人材の評価の仕組み
- インフラ(DBやスパコン)の未整備
  - 生物系は専門用語や概念が多く、参入障壁に
  - そのための知識の整理や教材必要
  - データの囲い込みの解消