

我が国におけるバイオインフォマティクスの現状

東京大学大学院新領域創成科学研究科
情報生命科学専攻
高木 利久

ゲノム科学とバイオインフォマティクス

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

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オーダーのデータ量の時代に突入

最近のバイオインフォを取り巻く状況

- NGS(次世代シーケンサー)
- ポストポストゲノムはゲノム
- ビッグデータ
- イメージデータ(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(メディカルインフォマティクス)
- 人材不足がより顕著に
- **新たな展開に対応した研究開発、データベース構築、人材育成が必要**

海外のバイオインフォマティクス教育機関

| エリア | 取得できる学位等 | | | | 主な教育機関等 |
|------------------|----------|----|----|----|---|
| | 学 | 修 | 博 | 他 | |
| 米国およびカナダ | 20 | 40 | 43 | 18 | Universite de Montreal(加)、Indiana University、Ohio University、Oregon Health & Science University、University of California各校、University of Delaware |
| 欧州 | 3 | 54 | 4 | 7 | 英国が最も多く、ついで、ドイツ、スペインが多い。全体として大学が多いが、ドイツではMax Planckもコースを提供。 |
| アジア | 6 | 12 | 3 | 2 | ほとんどがインドの大学及び研究機関。そのほかでは韓国の研究機関が1件。 |
| 他(アフリカ、中東、豪州、南米) | 9 | 6 | 4 | 1 | オーストラリアでは各大学が、南アフリカではSouth African National Bioinformatics Institute、イスラエルではWeizmann Institute of Scienceが中心。 |

なぜ人材育成はうまく行かなかった？

- 人材は育ってきたが需要の拡大の方が大きかっただけ！
- なぜ需要を見越して対応できなかったのか？
 - 教育機関(大学)の問題
 - 任期制の問題
 - 産業界の問題
 - ライフ分野、研究機関の問題
 - 政府やファンディングエージェンシーの問題
 - 研究費の問題
 - 学生側の問題
 - 制度面の問題、データ囲い込み問題
 - ポジション、キャリアパスの問題
 - 評価の問題
- 切実感、切迫感がなかった？
- 需給のミスマッチが起きている

これからの課題

- これまでと同じことをしてはだめ！
- 人材と一口に言っても、、、
 - 新規アルゴリズムや新規データベースを開発
 - データベースやツールを使いこなして問題解決
 - キュレータ、アナレータ、SE
- これらの人材に応じたキャリアパス、評価の設計
- 学部レベルからの教育の重要性
- 新規参入を促すためのデータベース開発
- 研究開発、データベース、人材育成の一体設計
- 恒久ポジションの確保、人材の流通の仕組み
- 各プロジェクトとの連動
- 解決への動きも(NGS現場の会、生命情報若手の会)

統合データベースの重要性

- 他の分野からの参入障壁を取り除く
- 分野間の壁を取り払う
- 研究の加速、効率化
- オープンイノベーション、バイオインフォの基盤
 - データ生産者以外の不特定多数がイノベーターに
- 米国ビッグデータイニシアティブ
- 欧州ELIXIR
- 中国BGI
- 日本NBDC