## 自然科学社会実装概論(数学/情報工学)

バイオインフォマティックスと テイラーメイド医療

田端俊英 博士(医学),教授

工学部知能情報工学コース

#### 世界の最先端の科学の中心はバイオ

- Nature, Scienceなど有力科学雑誌の記事の6割がバイオ関連
- シリコンバレーのIT長者が投資しているのもバイオ関連ベンチャー
- 数理,情報,機械,電気,化学…あらゆる分野の技術が必要とされる
- 我が国はバイオの基礎力は強いが、医療・福祉機器・サービスなどは輸入超過
  → 伸び代がある
- 医療・福祉機器・サービスの製品は命がかかっているので高くても売れる(付加価値が高い)
  - → 我が国もヘルスケアで経済発展を目指している

# <u>遺伝子DNAはタンパク質の設計図</u>



約2万種類の遺伝子 ↓ 約2万種類のタンパク質

### <u>遺伝子DNAは4種類の文字から成る暗号文</u>



 4種類の塩基 アデニンA チミンT グアニンG シトシンC

# <u>3塩基で1アミノ酸または「終止」を指定</u>



Ala/A	<b>アラニン</b> GCT, GCC, GCA, GCG	Leu/L	<b>ロイシン</b> TTA, TTG, CTT, CTC, CTA, CTG
Arg/R	<b>アルギニン</b> CGT, CGC, CGA, CGG, AGA, AGG	Lys/K	<b>リシン</b> AAA, AAG
Asn/N	<b>アスパラギン</b> AAT, AAC	Met/M	メ <del>チオニン</del> ATG
Asp/D	<b>アスパラギン酸</b> GAT, GAC	Phe/F	<b>フェニルアラニン</b> TTT, TTC
Cys/C	<b>システイン</b> UGT, TGC	Pro/P	<b>プロリン</b> CCT, CCC, CCA, CCG
Gln/Q	<b>グルタミン</b> CAA, CAG	Ser/S	<b>セリン</b> TCT, TCC, TCA, TCG, AGT, AGC
Glu/E	<b>グルタミン酸</b> GAA, GAG	Thr/T	トレオニン ACT, ACC, ACA, ACG
Gly/G	グリシン	Trp/W	トリプトファン

#### Bioinformatics 情報学を駆使して生物学を支援

- 生物学の情報を検索,収集
- 遺伝子の情報を検索,収集,(その産物を)可視化
- 遺伝子編集の方略立案
- 疾患の情報をマイニング
- 遺伝子情報から個人に最適化した予防・治療法の提案
- 病原体,疾患因子等(主としてタンパク質)の立体構造予測
- 治療薬の分子設計
- 専門家へのスキル、ツールの提供
- 一般市民への情報提供

生物学の情報を検索,収集 遺伝子の情報を検索,収集

Public Medline (PubMed) National Library of Medicine, National Institute of Health (NIH), USA

無料配布ツール 論文,ゲノム情報,タンパク質のアミノ酸配列情報,…

<u>英語で記述された</u>情報が網羅されている

このほか画像処理アプリなどの情報ツールが無料配布

### 遺伝子の(産物の)情報を可視化

遺伝子 → アミノ酸配列 → タンパク質

Chimera University of California, San Francisco (UCSF), USA

無料配布ツール アミノ酸の配列と相対的な位置情報からタンパク質の立体構造を可視化 病原体,疾患因子等(主としてタンパク質)の立体構造予測 アミノ酸1次配列からタンパク質立体構造を予測することは極めて難しい → 人類に残された難問の一つ

SWISS MODEL University of Basel, Switzerland

無料でブラウザーから利用できる X線構造結晶解析データから推測した代表的なタンパク質の立体構造を テンプレートとして,遺伝子・アミノ酸配列は判明しているが,立体構造 は分かっていないタンパク質の立体構造を予測する。 因みにX線構造結晶解析データの解析がスーパーコンピューター京・富岳の主目的

この他,アミノ酸配列だけからタンパク質の立体構造を予測するab-initio型のア プリも開発されているが,まだまだ発展途上 疾患の情報をマイニング → 電子カルテのビッグデータ解析

遺伝子情報から個人に最適化した予防・治療法の提案 → personalized medicine (テーラーメイド医療)

in-silico病原性予測アプリ single-nucleotide polymorphism (SNP)などの遺伝子多型や遺伝子突然変異の病 原性を予測する SIFT PolyPhen-2 MutationTaster-2 CADD

現時点では予測が全く当たらない





#### 専門家へのスキル、ツールの開発・提供

NIH Image J NIH, USA

ToGoTV <u>https://togotv.dbcls.jp</u> 大学共同利用機関法人 情報・システム研究機構データサイエンス共 同利用基盤施設**ライフサイエンス統合データベースセンター**,日本 一般市民への情報提供

Medline Plus National Library of Medicine, NIH, USA Bioinformaticsに関連した仕事に就くためには

- 英語力を少しずつ鍛錬する
- Bioinformaticsを開発または多用する大学研究室で修行する
- ライフサイエンス統合データベースセンターのToGoTVを視聴する
- 教科書で独習し、できれば資格(バイオインフォマティクス技術者認定試験)
  をとる

広川貴次先生→





←技術者認定試験 参考書

-----

#### Bioinformaticsを含む情報化社会全体で格差が拡がる可能性

英語ができるヒト → 情報強者英語ができないヒト → 情報弱者

しかしながら,

- 英米ではさほど問題にならない(一般市民も英語ができる)
- 諸外国では大卒者にはさほど問題にならない(大卒生は英語ができる)
- 日本では有力大学以外の大卒者では深刻な問題(文科省では、今後は英語ができない者は大学を卒業させないようにする方針)
- 英語力が必要とされることはこれから先も変わらない/変えられない、変えようとしようと考える人も世界では少数派
- ・ 学習方法を工夫して、少なくとも情報ツールを使いこなす英語力を修得することが重要(自動翻訳機はほとんど役に立たない) → 今後の我が国の発展を左右する → 教師が生徒に危機感を伝え、自らも学ぶ姿勢が必要

#### 課題

興味を持った話題に関するアプリをダウンロードし(またはブラウザー で操作し),実際に操作する。

上記アプリの動作原理、アルゴリズムを調べ、簡潔な文章でまとめる。

上記のアプリ以外にライバル·アプリや発展形のアプリを探し、その概要 をまとめる。

A4用紙1枚にまとめ(ワードファイル), <u>メールでttabata@eng.u-</u> toyama.ac.jpに提出する(期限:授業から1週間以内)。