

2025年7月25日

令和7年度筑波大学公開講座

「生命現象を「計算」する—in silico 生物学の現在—」

体験！バイオインフォマティクス解析

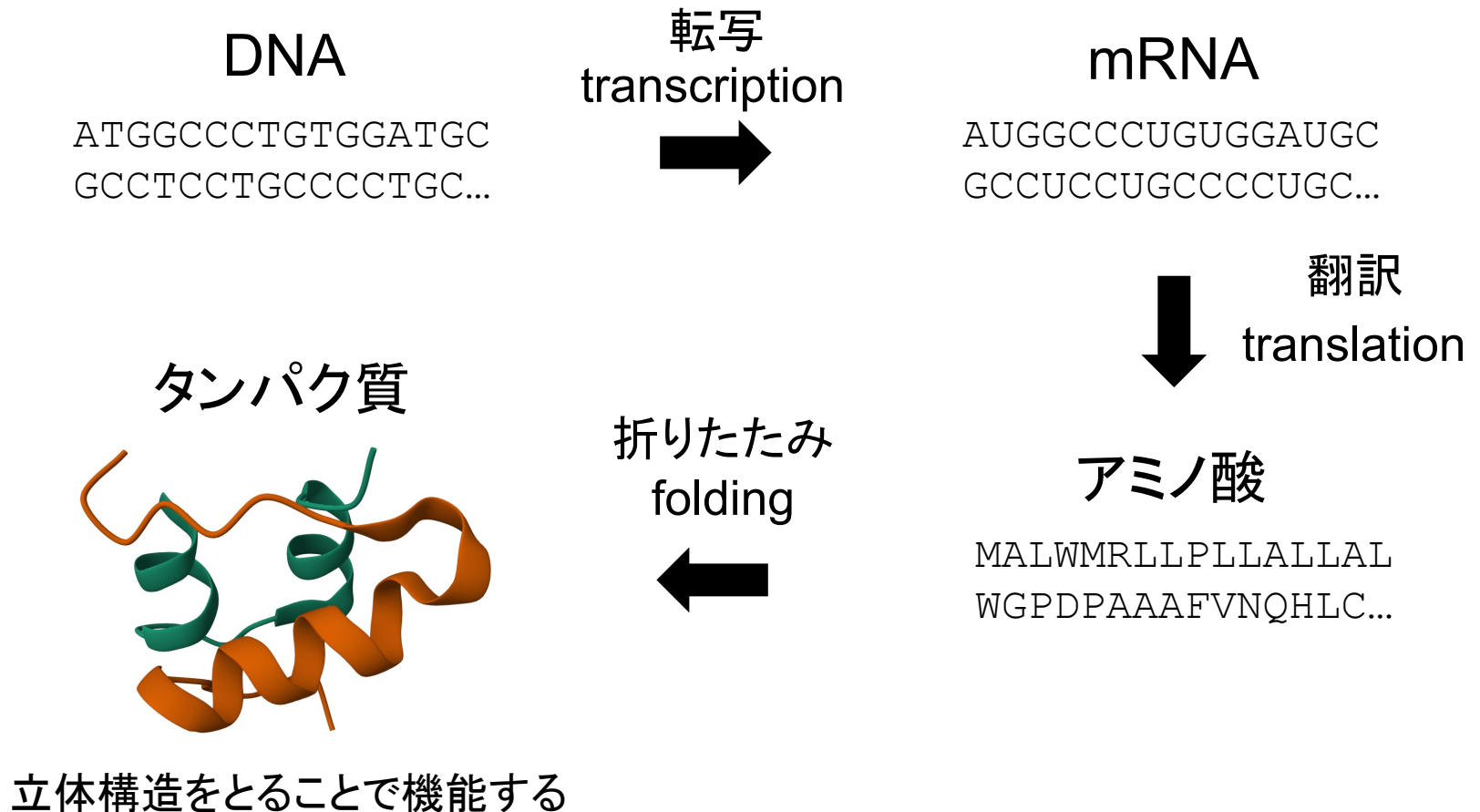
開催時間：13:15～14:30

- 本講義では、現代の生命科学研究に欠かせないバイオインフォマティクスを、実際に手を動かして体験していただきます。
- 操作性を考慮し、画面の大きなデバイス(パソコン>タブレット>スマートフォン)での参加を推奨します。
- 講義資料や演習課題は特設ホームページに公開されており、講義終了後もアクセスできます。https://github.com/ShuntaYorimoto/20250725_BioinfoHandson

筑波大学 生存ダイナミクス研究センター

頼本 隼汰

セントラルドグマ: DNAからタンパク質まで



DNAやタンパク質は「文字列」として表現することが可能

バイオインフォマティクスとは

- バイオ(生命科学) + インフォマティクス(情報科学)
- 生物学的データを計算機で解析する
- 生物情報をコンピュータで扱うために、DNAやタンパク質は決まった形式で記録・保存されている

FASTA形式

```
>NC_012920.1:5904-7445 MT-C01 [organism=Homo sapiens]
ATGTTGCGCCGACCGTTGACTATTCTCTACAAACCACAAAGACATTGGAACACTATACCTAT
TATTCGGCGCATGAGCTGGAGTCCTAGGCACAGCTCTAAGCCTCCTTATTCGAGCCGAGCT
GGGCCAGCCAGGCAACCTTCTAGGTAACGACCACATCTACAACGTTATCGTCACAGCCCAT
GCATTTGTAATAATCTTCTTCATAGTAATACCCATCATAATCGGAGGCTTTGGCAACTGAC
TAGTTCCCCTAATAATCGGTGCCCCCGATATGGCGTTTCCCCGCATAAACATAAGCTT
(以下省略)
```

DNAには様々な突然変異が生じる

突然変異:

DNA複製ミスや外的要因によって、遺伝物質の変化が引き起こされる

点変異

元配列: ATCG**A**TCGA

変異後: ATCG**T**TCGA

挿入変異

元配列: ATCGTCGA

変異後: ATCG**AAA**TCGA

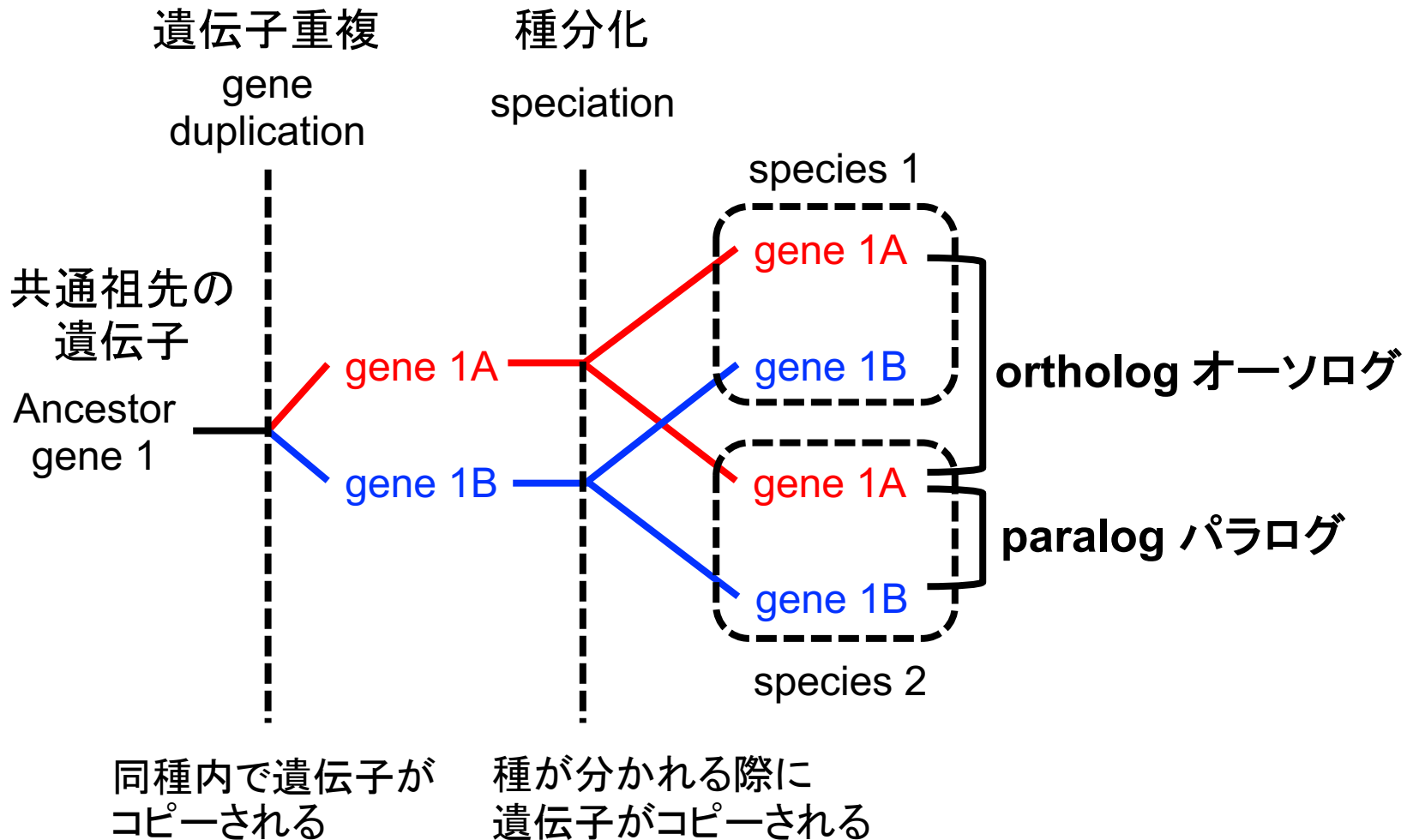
欠失変異

元配列: ATCG**AAA**TCGA

変異後: ATCGTCGA

これらの変異が子孫に受け継がれ、世代を重ねて変異が蓄積し、集団に広がることで生物の多様性が生まれる

遺伝子の進化: オースログとパラログ



配列を比較することで、進化や機能の関係を推定できる

配列の相同性解析

Pairwise alignment ペアワイズアライメント:

二つの塩基配列またはアミノ酸配列同士を比較し、最適な一致を見つけ、その類似性を調べる手法

例:

DNA

protein

配列1: GACTCTTACCT

配列1: QSDGNFLDED

配列2: GGC-CCTGCCC

配列2: QTDGNFLDDD

* * * * *

* : * * * * * : *

"-": gap 空白

"*": match 一致

":": conserved substitutions アミノ酸の性質が似ている

" ": mismatch 不一致

点数をつけて類似性を評価する

点数ルール1:

一致 = +1

不一致 = -1

空白 = -1

点数ルール2:

一致 = +1

不一致 = -1

空白 = 0

(A)

配列1: GACTCTTACCT

配列2: GGC-CCTGCCC

* * * * *

1

(A)

配列1: GACTCTTACCT

配列2: GGC-CCTGCCC

* * * * *

2

(B)

配列1: GACTCTTA-CCT

配列2: GGC-CCTGCCC-

* * * * *

0

(B)

配列1: GACTCTTA-CCT

配列2: GGC-CCTGCCC-

* * * * *

3

- 同じ配列でも、何を重視するかによって最適な並べ方は変わる
- どのツール・パラメータを使うかで結果が変わることがある

実際の配列はもっと長い...人間には無理！

- 人の目で確認するのは時間がかかるし、間違いやすい
- 高速で正確なコンピュータに任せよう！

```
>HBB_mRNA_normal NM_000518.5 Homo sapiens hemoglobin subunit beta (HBB), mRNA
ACATTTGCTTCTGACACAACCTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTGACTCCTGA
GGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGC
AGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATG
CTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGC
TCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGAT
CCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCA
CCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCA
CTAAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGTTTCCTTTGTTCCCTAAGTCCAACACTACTAACT
GGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCAA
```

```
>HBB_mRNA_mutated NM_000518.5 Homo sapiens hemoglobin subunit beta (HBB), mRNA
ACATTTGCTTCTGACACAACCTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTGACTCCTGT
GGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGC
AGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATG
CTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGC
TCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGAT
CCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCA
CCCCACCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAGTATCA
CTAAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGTTTCCTTTGTTCCCTAAGTCCAACACTACTAACT
GGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCAA
```


無料のWEBツールを用いて解析してみましょう！

- 練習問題1: Pairwise sequence alignment

https://github.com/ShuntaYorimoto/20250725_BioinfoHandson/wiki/exercise-1

- 練習問題2: BLAST

https://github.com/ShuntaYorimoto/20250725_BioinfoHandson/wiki/exercise-2

- 練習問題3: Multiple sequence alignment

https://github.com/ShuntaYorimoto/20250725_BioinfoHandson/wiki/exercise-3

まとめ

- バイオインフォマティクスとは、生物学的データを計算機で解析する分野である
- DNAやタンパク質の配列を比較することで、進化や機能の関係を推定できる
- 基礎生物学、医学・薬学、農林水産・畜産、食品科学、環境科学など、様々な分野でバイオインフォマティクスが利用されている