

# 2021 年度バイオインフォマティクス技術者認定試験 問題解説について

バイオインフォマティクス技術者認定試験は、2021 年度から CBT 方式に移行しました。これに伴い、従来よりも幅広いトピックに関する問題も出題されるようになりました。

この問題/用語解説では、いくつかの新しいトピックや、これまでも出題されていたものの教科書には記載されていないトピック、過去問などについて、簡単な解説を行います。

## 【生命科学分野】

### 総論:

生物は多様で例外も多く、問題文や選択肢を読んでいる中で、厳密には当てはまらないと考えてしまう場合もあるかもしれません。そのような場合は大多数の生物にあてはまる選択肢や基本原理を選んで解答してください。

### 各論:

生命科学分野の問題は、参考書「バイオインフォマティクス入門」(慶応義塾大学出版会)を中心に、ホームページに掲載の出題範囲のキーワードに関する問題が出題されます。

2021年12月に参考書の第2版が出版され、以下の2項目が追加されました。

#### ○分子・細胞生物学実験技術

ゲノム編集、エピジェネティクス、ES細胞・iPS細胞、についてページが新たに追加されました。ゲノム編集やエピゲノムに関する過去の出題の頻度及び難易度は高くありませんでしたが、参考書ではCRISPR-Cas9によるゲノム編集のメカニズムが図によって詳しく紹介されていますので理解の参考にしてください。また、エピゲノムに関しては、参考書の記述について、DNAやゲノム、染色体の構造と合わせて理解するようにしてください。

#### ○生命科学研究のための法令・指針:

生命科学研究の実施にあたり遵守すべき法令・研究倫理指針のページが新たに追加されました。特にバイオインフォマティクスに関係の深い、「人を対象とする生命科学・医学系研究に関する倫理指針」と「個人情報の保護に関する法律」について最近の改正内容について解説されています。法令や指針の文書を暗記する必要はありませんが、技術者として知っておくべき事項として参考文献と合わせて必ず目を通すようにしましょう。

## 【構造分野】

### クライオ電子顕微鏡によるタンパク質立体構造解析

昨年度に引き続きクライオ電子顕微鏡法(クライオ EM)による立体構造決定に関する問題を出題しました。生体高分子の立体構造を決定する手法のひとつとして着実に定着しつつあり、今後ますます重要性は増していきます。電子顕微鏡では標的とする分子を含むサンプルに対して2次元の画像を得ることができます。この画像から3次元の空間データ(原子座標)を生成することで立体構造データベース(Protein Data Bank)に蓄積されているデータが得られます。このような2次元像から3次元像への変換には高度な情報科学技術が活用されており、タンパク質立体構造データを作成し登録する立場のみならず、データを利用する立場からも、この仕組みを理解しておくことが重要となります。

### ケモインフォマティクス

化学式等で表される化合物の構造や性質を情報科学的に扱う分野をケモインフォマティクスと言います。生命情報を主として扱うバイオインフォマティクスとは異なる分野として扱われることが多いですが、これら2つのインフォマティクス領域は極めて関連が深く、また部分的に重複しているとも言えます。特に医療、創薬、メタボロミクスなどでは生体高分子と低分子化合物の両者の関連を扱う必要があるため、ケモインフォマティクス技術が重要となります。代表的な技術として、通常はグラフとして表される化学構造式を線形表記する記法(SMILES 記法)、化合物の特徴を数値ベクトルとして表現し化合物同士の類似性を比較する手法などがあります。

### SARS-CoV-2

新型コロナウイルスの蔓延が長期化しています。ドキュメンタリ番組などでコロナウイルスのもつタンパク質のCG映像などを目にした方も多いのではないのでしょうか。2021年度はコロナウイルスのタンパク質に関連する問題を出題しました。コロナウイルスでは様々な変異株の存在が話題となりましたが、この「変異」とはタンパク質のアミノ酸残基の変化を意味し、立体構造と照らし合わせることでどこがどのように変わったのかを理解することができます。ヒトの個人差を考察する上でも、個人個人でどのようにタンパク質の配列上に変異があり、それが物理化学的にどのような帰結をもたらすのか、タンパク質立体構造データに基づいて考察することができます。疾病や様々な生命現象を理解するにあたって、タンパク質の配列、構造、機能の関連を理解することが重要となります。

## 【遺伝進化分野】

以下いくつかの過去問について解説します。

### 中立説・同義置換・非同義置換

問: 遺伝子の進化的性質の1つに分子時計がある。分子時計について述べた以下の記述のうち、もっとも不適切なものを選択肢の中から一つ選べ。

1. 分子時計とは、アミノ酸置換や塩基置換がほぼ一定の速度で起こる現象を指す。
2. 分子時計は、ゲノム配列に起こる中立な変異をもとに決定するため、遺伝子によらず同じ速度となる。
3. 分子時計によって分岐年代を推定するには、化石などの分子以外の証拠から分岐時間がわかっている生物種が一組は必要である。
4. 分子時計は、非常に離れた生物の間では、同じ塩基に複数回変異が入る可能性があるため精度は低下する。

解説: 分子時計は選択肢1に述べたとおり、選択的に(有利でも不利でもない)中立な遺伝子の変異が確率的過程で集団に固定することによって成立する現象である。この時、中立な変異の割合は遺伝子によって異なるので、分子時計の速度は遺伝子間では一定ではない。よって選択肢2が正解である。分子時計の速度が遺伝子ごとに異なるので、選択肢3の通り年代推定には化石データなどの外部参照が必要になる。

問: タンパク質をコードしている遺伝子の進化の過程における塩基置換は、同義置換と非同義置換に分類される。同義置換と非同義置換に関する以下の記述のうち、もっとも不適切なものを選択肢の中から一つ選べ。同義置換率は同義サイトあたりの置換数、非同義置換率は非同義サイトあたりの置換数である。

1. 同義置換はアミノ酸を変化させない置換で、非同義置換はアミノ酸を変化させる置換である。
2. 大部分の遺伝子では、同義置換率は非同義置換率よりも小さい。
3. ウイルスの表面抗原の抗体によって認識される部位では、抗体による認識を避けるためアミノ酸置換が加速されることがあり、非同義置換率が大きくなる。
4. マウスとヒトのオーソログ遺伝子を比較する場合、同義置換率は遺伝子によらずほぼ一定の値になるが、非同義置換率はコードされる遺伝子によって大きく異なる。

解説: 同義・非同義置換の定義は選択肢1のとおりである。また上記問題の解説のとおり、中立な変異の割合は遺伝子によって異なるので、選択肢3と4は正しい。ウイルスの表面抗原はアミノ酸変化により宿主免疫から逃れることができるので、非同義置換率が大きくなる典型例である。ただしそれ以外の大部

分の遺伝子では、非同義置換は遺伝子機能上の障害のため選択されにくく、同義置換よりも少ないので選択肢 2 が正解である。

### 遺伝子組換え

問: 遺伝子連鎖解析の結果、遺伝子座 A と B の間で 1000 回に 1 回の割合で組換えが観察された。この遺伝子座の間の遺伝学的距離として、もっとも適切なものを選択肢の中から一つ選べ。

1. 0.01 cM
2. 0.10 cM
3. 1.00 cM
4. 10.0 cM

解説: cM(センチモルガン)は遺伝子組換え頻度から推定されるゲノム上の遺伝子間距離である。100 回に 1 回の割合で起こる時が 1cM で、この値はおおむね遺伝子間距離に比例する(およそ 100 万塩基対/cM)。従って選択肢 2 が正解である。

問: 連鎖解析に関する以下の記述のうち、もっとも不適切なものを選択肢の中から一つ選べ。

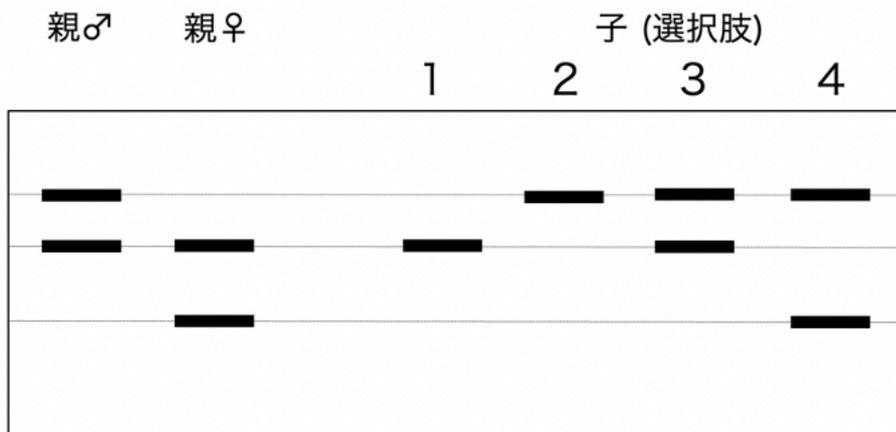
1. 連鎖不平衡とは、複数の遺伝子座の対立遺伝子の特定の組み合わせ頻度が有意に高くなる現象である。
2. 連鎖不平衡にある遺伝子座は同一染色体に存在する機会が多いが、交配がランダムでない集団などでは、異なる染色体上に存在する遺伝子座の間でも連鎖不平衡が認められる。
3. 連鎖不平衡にある複数の遺伝子座の対立遺伝子の特定の組み合わせがディプロタイプであり、相同染色体間でのディプロタイプの組み合わせをハプロタイプとよぶ。
4. LOD(logarithm of odds)スコアを用いた家系分析を行うことにより、遺伝的疾患の原因遺伝子が連鎖地図上でマーカー遺伝子からどの程度離れた位置にあるかを推定できる。

解説: 連鎖解析は、マーカー遺伝子との間の遺伝子組換え頻度から、目的の遺伝子のゲノム上の位置を推定する方法である。連鎖とは 2 つの遺伝子が組換えされずに子孫に伝わる現象であり、平均的組換え率から逸脱して連鎖が観測される現象を連鎖不平衡という。よって選択肢 1、4 の内容は正しい。連鎖は基本的に同一染色体上に存在する遺伝子間で起こる現象であるが、選択肢 2 の通りそもそも

交配がランダムでなく、特定個体に由来する染色体セットの頻度が高くなると見かけ上の連鎖不平衡が観測されうる。半数染色体上で連鎖する遺伝子の組合わせをハプロタイプ、二倍対染色体のハプロタイプの組合わせをディプロタイプというので、選択肢 3 は用語が逆であり、これが正解である。

### マイクロサテライト領域・遺伝子診断

問: 飼育しているマウスの親子関係が不明になったため、マイクロサテライト領域の多型を利用して DNA 鑑定を行った。下の図は、常染色体上の特定座位のマイクロサテライトを増幅するプライマーを用いて、両親(親♂および親♀)およびその両親の子供であることを検査するマウス(子 1~4)のゲノムに対して PCR 増幅を行い、増幅産物を電気泳動で分離したゲルの模式図である。この両親の子供でない可能性が最も高い個体を選択肢の中から一つ選べ。ただし、生殖過程における染色体の転座・転移・重複、およびマイクロサテライト反復数の変化は起こっていないものと仮定する。



解説: 子は両親からそれぞれ半数染色体分の多型 DNA を継承していることから、子の多型 DNA は両親のゲノム中にも存在し、また両親の持つ多型 DNA の少なくとも 1 つを子は保有しているはずである。選択肢 2 の個体のみが、親♀の多型 DNA を継承していないので、この個体がこの両親の子でない可能性が最も高い。