平成27年度

日本バイオインフォマティクス学会 (JSBi) バイオインフォマティクス技術者認定試験 試験問題

注意事項

- 1. 試験開始の合図があるまで、問題冊子を開いてはいけません。
- 2. 問題数は 80 問 です。すべての問題に解答してください。
- 3. 答案用紙への受験番号などの記入は、試験監督の指示があってから始めてください。
- 4. 試験時間は、13:30 ~ 15:30 (2時間)です。
- 5. 試験中にトイレに行きたくなったり気分が悪くなったりした場合は、手を挙げて試験監督に合図してください。
- 6. 退出可能時間は 14:30 ~ 15:20 です。
 - 途中で退出する場合には、手を挙げて試験監督に合図をし、解答用紙が回収されてから静かに退出してください。この場合、試験終了まで再入場はできません。
- 7. 試験問題の内容に関する質問にはお答えできません。文意通りに解釈してください。
- 8. 解答用紙(マークシート)の記入に当たっては、次の指示に従ってください。
 - (1) 問80までの解答欄を利用してください。問81以降への記入は、無効です。
 - (2) HB かB の黒鉛筆もしくはシャープペンシルを使用してください。
 - (3) 受験番号は、受験票に印字されている通りに記入し、マークしてください。
 - (4) 氏名とフリガナは、受験票に印字されている通りに記入してください。
 - (5) 年月日欄には、生年月日を記入してください(年は、西暦で記入)。
 - (6) 解答は、各解答記入欄に一つずつマークしてください。
 - (7) 訂正の場合は、あとが残らないように消しゴムできれいに消し、消しくずを残さないでください。
- 9. 試験終了後、この問題用紙は持ち帰ることができます。
- 10. 解答用紙 (マークシート)は、白紙であっても必ず提出してください。

原核生物に関する記述についてもっとも不適切なものを選択肢の中から一つ選べ。

- 1 明確な核を持たない。
- 2 ゲノム DNA はほとんどの場合直鎖状である。
- 3 真正細菌と古細菌に分けられる。
- 4 90℃以上の高温で生存可能なものが存在する。

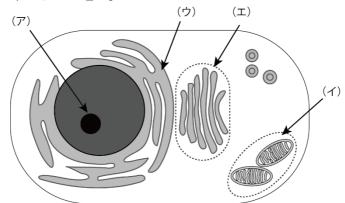
問 2

ウイルスに関する記述についてもっとも不適切なものを選択肢の中から一つ選べ。

- 1 細菌に感染するものがある。
- 2 ゲノムとして RNA を持つものがある。
- 3 ゲノムとして DNA を持つものがある。
- 4 分裂により増殖する。

問 3

次に示した動物細胞の模式図の(ア)~(エ)の細胞内小器官の名称と機能の組み合わせとしてもっと も不適切なものを選択肢の中から一つ選べ。



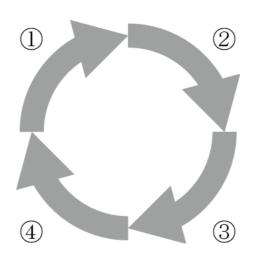
選択肢	記号	名称	機能
1	(ア)	核小体	rRNA の転写、リボソームの構築
2	(イ)	ミトコンドリア	解糖系によるエネルギーの産生
3	(ウ)	小胞体	タンパク質の修飾・輸送
4	(工)	ゴルジ体	タンパク質の分泌

ミトコンドリアに関する以下の記述について、もっとも不適切なものを選択肢の中から一つ選べ。

- 1 ミトコンドリアはほとんどすべての真核生物の細胞内に存在する。
- 2 ミトコンドリアは独自の直鎖状ゲノム DNA を持つ。
- 3 ミトコンドリアには核のゲノムにコードされたタンパク質も存在する。
- 4 ミトコンドリアは内膜と外膜の二重膜で形成されている。

問 5

細胞周期は大きく 4つの時期(図中の①~④)に分けられ、(①→②→③→④→)の順で繰り返される。 真核生物では図の②の時期には核構造が消失し、染色体と呼ばれる特徴的な棒状の構造体が観察される。 ①~④のそれぞれの時期の名称と細胞あたりの DNA の相対量について、正しい組み合わせを表中の選 択肢から一つ選べ。ただし、n は倍数性の基本単位である。



選択肢	図中の記号	周期名	DNA 量
1	1)	G ₂ 期	2n
2	2	G ₂ 期	4n から 2n に減少
3	3	S期	4n
4	4	S期	2n から 4n に増加

減数分裂に関する以下の記述のうちでもっとも不適切なものを選択肢の中から一つ選べ。

- 1 二倍体の細胞から一倍体の細胞を作る過程は減数分裂と呼ばれる。
- 2 減数分裂では二倍体の細胞が1回分裂し一倍体の細胞がつくられる。
- 3 減数分裂では1個の母細胞から4個の細胞が形成される。
- 4 減数分裂において相同染色体間で交差が起こり、遺伝子の組み換えが起こる。

問 7

連鎖していない 2 つの遺伝子座がある。それぞれの対立遺伝子を A および a、B および b とする。このとき遺伝子型 AaBB と遺伝子型 AABb の親の掛け合わせから生ずる子 (F1) の遺伝子型の組み合わせとして正しいものを選択肢の中から一つ選べ。

- 1 AABB, AaBB, aaBB, AABb
- 2 AABB, AABb, AaBB, AaBb
- 3 AABB, AABb, aaBB, aaBb
- 4 AABB, AABb, AAbb, AaBB, AaBb, Aabb, aaBB, aaBb, aabb

問8

免疫に関する以下の記述に関してもっとも不適切なものを選択肢の中から一つ選べ。

- 1 マクロファージは死んだ細胞や体内に入り込んだ異物を取り込んで消化し抗原提示を行う。
- 2 成熟した B 細胞は 1 細胞で様々な抗原を認識する抗体を産生する能力を獲得しており、そのまま記憶細胞として長期間生存し続ける。
- 3 抗体分子の多様性は主に V(D)J 組換えと呼ばれる遺伝子再構成によって生じる。
- 4 自然免疫は微生物などに固有の分子パターンを受容体が認識することで行われる。

神経伝達物質としてもっとも不適切なものを選択肢の中から一つ選べ。

- 1 リジン
- 2 ドーパミン
- 3 グルタミン酸
- 4 アセチルコリン

問 10

以下の生体膜に関する文章中の(a)~(d)に入る語句の組み合わせとしてもっとも適切なものを選択肢の中から一つ選べ。

生体膜の主成分は(a)であり、他にステロールや糖脂質が含まれる。(a)は(b)の部分と炭化水素鎖の(c)の部分からなる両親媒性の分子であり、水溶液中で(b)の部分を外側に(c)の部分を内側に向けて互いに集まり、脂質二重層やミセルを形成する。脂質二重層はそれ単独では(d)をほとんど通さないことから、膜の内外で(d)を受け渡すためには膜タンパク質の助けが必要となる。

1	(a)リン脂質	(b)親水性	(c)疎水性	(d)イオン
2	(a)糖鎖	(b)疎水性	(c)親水性	(d)イオン
3	(a)リン脂質	(b)親水性	(c)疎水性	(d)水分子
4	(a)糖鎖	(b)疎水性	(c)親水性	(d)水分子

問 11

DNA とそれを構成する塩基($A: \mathcal{P} \mathcal{F} = \mathcal{V} \setminus C: \mathcal{V} \mathcal{V} \setminus G: \mathcal{J} \mathcal{F} = \mathcal{V} \setminus T: \mathcal{F} \mathcal{F} \mathcal{V}$)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 AとGはプリンに分類される。
- 2 DNA は二重らせん構造を形成する。
- 3 DNA中のAはTと、GはCと対合する。

DNA の複製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 DNA は半保存的に複製される。
- 2 DNA ポリメラーゼは DNA の 5' 末端から 3' 末端の方向に合成を行う。
- 3 複製フォークが進行する方向と同一方向に DNA ポリメラーゼが連続して DNA を合成する鎖をリーディング鎖、その逆方向に向かって断続的に DNA 合成が行われる鎖をリバース鎖という。
- 4 複製フォークの進行とは逆方向に断続的に合成される DNA 断片を岡崎フラグメントというが、これは発見者の日本人にちなんで名付けられた。

間 13

真核生物のメンセンジャーRNA (mRNA) 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 ほとんどの mRNA は 3' 末端にポリA配列を持つ。
- 2 すべての mRNA の主な塩基成分はアデニン、グアニン、チミン、シトシンである。
- 3 成熟した mRNA は 5' 末端にキャップ構造を持つ。
- 4 mRNA 前駆体はイントロンを持つ。

問 14

アミノ酸の特性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 もっとも分子量の小さなアミノ酸はグリシンである。
- 2 生理条件下で側鎖が正に荷電しているアミノ酸としてリジンやアルギニンがある。
- 3 セリンやシステインはその側鎖に水酸基を持つ。
- 4 化学合成したアミノ酸と生体内で合成されたアミノ酸は光学特性が異なる場合がある。

タンパク質の生合成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 タンパク質の翻訳は mRNA の 5' 末端から 3' 末端の方向に進む。
- 2 タンパク質の合成はアミノ末端からカルボキシル末端方向に合成が進む。
- 3 リボソームの A 部位にはアミノアシル tRNA が結合し、P 部位にはペプチジル tRNA が結合する。
- 4 終止コドンに終止 tRNA が結合すると翻訳が停止する。

問 16

タンパク質の折りたたみ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 疎水的なアミノ酸残基はタンパク質の内側に、親水的なアミノ酸残基はタンパク質の外側に配置されていることが多い。
- 2 タンパク質の折りたたみはリボソームによる mRNA の翻訳と同時に進行する。
- 3 プロテアソームはタンパク質が正しく折りたたまれることを助ける。
- 4 変性したタンパク質はオートファジーなどの機構により分解される。

問 17

ヌクレオソーム構造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 ヌクレオソームは 4 種類のヒストンそれぞれ 2 分子ずつで形成される 8 量体に 146 塩基対の DNA が巻き付いた構造をしている。
- 2 ヌクレオソーム構造は、ヒストン 8 量体に DNA が巻き付いたヌクレオソームコアと呼ばれる領域 とヒストンが結合していない DNA のみのリンカーと呼ばれる領域に分けられる。
- 3 真核生物の細胞中の DNA は全てヌクレオソーム構造を形成している。
- 4 ヒストン8量体とDNAの相互作用の強さはDNAの塩基配列からある程度予測が可能である。

ポリメラーゼ連鎖反応 (PCR) 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 ごく微量の試料から DNA を増幅することが可能である。
- 2 高熱環境に生息するバクテリアから単離された耐熱性の DNA ポリメラーゼがよく利用される。
- 3 反応液の組成を変化させることで変性、アニーリング、伸長の3つのステップを繰り返し増幅を行う。
- 4 増幅反応過程の DNA 量を逐次検出することで、サンプル中の標的 DNA 量を定量することが出来る。

間 19

次世代シークエンサの用途としてもっとも不適切なものを選択肢の中から一つ選べ。

- 1 RNA-Seq は cDNA の塩基配列を大量に決定することでサンプル中に含まれる転写物を計測するトランスクリプトーム解析手法である。
- 2 ChIP-Seq はタンパク質を免疫沈降した際に同時に回収された DNA の塩基配列を決定することで、 当該タンパク質と相互作用しているゲノム領域を同定するエピゲノム解析手法である。
- 3 細菌の集団(細菌叢)から個々のクローンを単離したりせずに直接 DNA を精製し、その配列決定 を行うことでその集団中の微生物の組成を明らかにする解析手法をメタボローム解析という。
- 4 キャプチャープローブを用いてゲノム DNA から主にエキソン領域を濃縮・回収し、配列決定を行 う手法をエキソーム解析という。

問 20

ヒトのタンパク質をコードする遺伝子数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 2万~3万個
- 2 200 万~300 万個
- 3 2000 万~3000 万個
- 4 2 億~3 億個

近年のコンピュータシステムでは、システム上に物理的に実装されているメモリ量を超えるような大規模なメモリを使用するプログラムを実行することができる。これを実現する技術に関連する用語としてもっとも不適切なものを選択肢の中から一つ選べ。

- 1 仮想記憶
- 2 スワップファイル
- 3 デフラグ
- 4 ページング

問22 全員正解としました

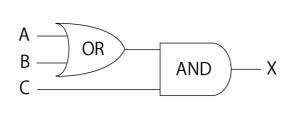
プログラミング言語の一つである Perl 言語は正規表現を柔軟に扱うことが可能である。正規表現 J[a-z]*[0-9]*-*[a-z]* にマッチする文字列として、不適切なものを選択肢の中から一つ選べ。な お、記号[] はその中で囲まれた文字のクラスを表し、その構成要素として、a-z は a から z までの 26 個の英字アルファベットを、0-9 は 0 から 9 までの 10 個のすべての数字を表す。また、記号 * は 直前の文字の 0 回以上の繰り返しを表す。

- 1 Japan2020-olympic
- 2 JSBi2015-test
- 3 Jlleague
- 4 Java-compiler

本来はパターンの完全一致を意図した問題でしたが、出題意図が十分に伝わらないという指摘があり、検討を重ねた結果この問題は全員正解としました。

すなわち、Perl の場合 $/^{J[a-z]*[0-9]*-*[a-z]*}/$ でしたら、2 が正解になりますが、 ^ および \$ がついていないので、すべての選択肢がマッチしてしまいます。

下記の図は、2つの論理素子を接続した論理回路を表現している。この回路では、A, B, Cが入力であり、Xが出力であるとする。この回路に対する入出力の結果によって真理値表を作成した。ここで、真理値表の(a) と(b) に入る値の組み合わせとして適切なものを、選択肢の中から一つ選べ。



論理回路図

A	B	C	X
0	0	0	0
0	0	1	(a)
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	(b)
1	1	1	1

真理値表

1	(a)	0	(b)	0
2	(a)	0	(b)	1
3	(a)	1	(b)	0
1	(a)	1	(b)	1

問 24

クラウドコンピューティングに関する以下の記述の中から、もっとも不適切なものを選択肢の中から一つ選べ。

- 1 クラウドコンピューティングとは、従来手元のコンピュータで管理・利用していたソフトウエアや データをインターネットなどのネットワークを通じて利用する方式である。
- 2 クラウドコンピューティングにより、あらかじめ数多くの計算機を購入しておくのではなく、必要 に応じて仮想計算機を借りることが可能となる。
- 3 必要に応じて借りられる代わりに、仮想計算機には必要なソフトウエアが一切インストールされて いないので、設定を全て自身で行う必要がある。
- 4 同一の仮想計算機イメージを利用することで、同一の計算機環境を複数構築できるため、計算機実 験の再現性を向上させることができる。

以下の(\mathbf{a})~(\mathbf{d})に入る語句の組み合わせとして、もっとも適切なものを選択肢の中から一つ選べ。

以下の図 1 は(a)言語の一種である(b)とよばれるプログラミング言語で書かれたコードの一部である。一方、図 2 は(c)言語の一種である(d)とよばれるプログラミング言語で書かれたコードの一部である。(c)言語はソースコードを機械語に翻訳してから実行するのに対し、(a)言語はプログラム中の命令 1 つ1 つを逐一翻訳しながら実行する。その簡便さから(a)言語は実行時間が問題とならない場合によく用いられている。

```
sub summation {
                                        int summation(int p, int q) {
     my (p, q) = @_;
                                          int sum;
     my \$sum = \$p + \$q;
                                          sum = p+q;
     return $sum;
                                          return sum;
                 図 1
                                                     図 2
                                 (c)インタープリタ
1 (a)コンパイラ
                    (b)Perl
                                                       (d)C
2 (a)コンパイラ
                    (b)C
                                 (c)インタープリタ
                                                       (d)Perl
3 (a)インタープリタ
                    (b)Perl
                                 (c)コンパイラ
                                                       (d)C
4 (a)インタープリタ
                    (b)C
                                 (c)コンパイラ
                                                       (d)Perl
```

問 26

次に示した説明文の中で、コンピュータやネットワークに関する用語の説明として不適切なものを選択 肢の中から一つ選べ。

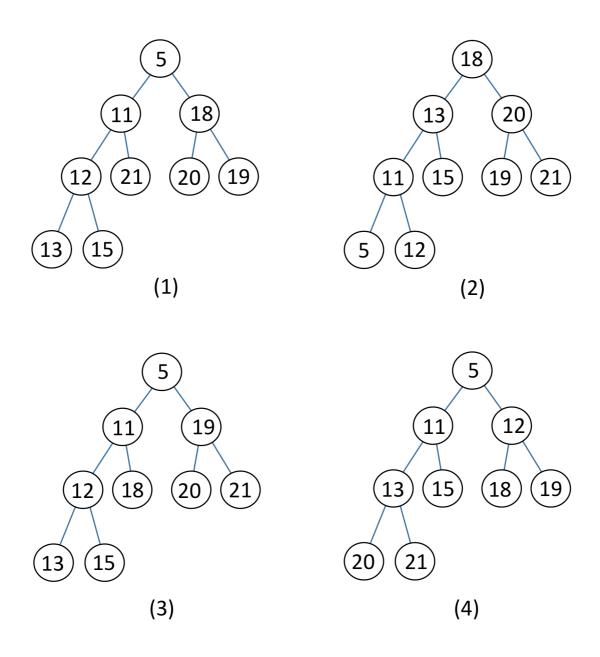
- 1 TCP/IP は USB 機器を接続するためのプロトコルである。
- 2 MAC アドレスは、ハードウェアに書き込まれており、全世界でユニークな ID である。
- 3 DHCP は接続するコンピュータに対して、IP アドレス等の必要な情報を自動的に割り当てるプロトコルである。
- 4 IP アドレスの数は近々枯渇する可能性がある。そのため IP のプロトコルである IPv4 から IPv6 への移行などが行われている。

逆ポーランド記法 (後置記法) は、演算子を非演算子の後ろに記述する記法である。

例えば、12+43-* は通常の中置記法の ((1+2)*(4-3)) と同じ意味である。逆ポーランド記法で書かれた式が文字列として与えられたときに、式の計算結果を求めるプログラムに用いるデータ構造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 キュー
- 2 スタック
- 3 二分探索木
- 4 ヒープ

以下の 4 個の木は、同じ数字集合を木構造の接点に割り当てたものである。4 個のうち 3 個がヒープ構造になっており、残りの一つが二分探索木になっている。二分探索木となっているものを 4 個の木から一つ選べ。



ハッシュと呼ばれるデータ構造において、ハッシュキーの衝突を起こした時の対処法に連鎖ハッシュ法 と開番地法がある。ハッシュについてもっとも不適切なものを選択肢の中から一つ選べ。

- 1 ハッシュでは、登録されている要素数がnのとき、最悪の場合、要素の探索が $O(2^n)$ かかることがある。
- 2 ハッシュでは、十分大きなハッシュテーブルと理想的なハッシュ関数 (定数時間で計算可能であるとする) を用いれば、平均的に定数時間で要素を探索することができる。
- 3 開番地法では、ハッシュキーが衝突した場合、次に空いている番地を探す方法である。
- 4 連鎖ハッシュ法では、ハッシュキーが衝突した場合、衝突を起こしたデータをポインタでつなぎ、 リストで管理する。

間 30

データベースのモデルの一つであるリレーショナルデータベースに関する以下の説明のうち、もっとも 不適切なものを選択肢の中から一つ選べ。

- 1 データの操作言語として SQL が知られている。
- 2 データ上の関係を木構造として表現する。
- 3 リレーショナルデータベースを管理するソフトウェアはRDBMSと呼ばれる。
- 4 データの操作は関係代数という演算の体系に基づいている。

問 31

以下の(a)~(c)に入れる語としてもっとも適切なものを選択肢の中から一つ選べ。

(a) は、複数のデータベースを互いに (b) でつなぎ、表の間の連携を取ることが可能なデータベースである。また、(a) では、検索を高速に行うために (c) を利用することができる。

(a) リレーショナルデータベース(RDB)
 (b) 関係(リレーション)
 (c) 従属関係
 (d) XML
 (e) 従属関係
 (e) インデックス
 (e) インデックス
 (e) インデックス
 (e) インデックス
 (f) 関係(リレーション)
 (i) 関係(リレーション)

XML についての記述としてもっとも不適切なものを選択肢の中から一つ選べ。

- 1 XML はウェブを記述するための言語であり、現在のほとんどのウェブページは XML によって書かれている。
- 2 XML は木構造となっており、各要素をタグで囲み階層的なデータ構造を入れ子で表現する。
- 3 HTML を XML にもとづいて修正した実装に XHTML がある。
- 4 DTD は XML において文書構造を定義することができるスキーマ言語のひとつである。

間 33

以下に示した 5 個の数値データが与えられたとする。このデータにおける 3 個の基本統計量(平均値、中央値、最頻値)の値として正しい組み合わせのものを選択肢の中から一つ選べ。ここで、中央値はメディアン、最頻値はモードとも呼ばれる。

データ: 3, 4, 3, 6, 5

選択肢	平均値	中央値	最頻値
1	4	4.5	3
2	4.2	4	3
3	4	4.5	4
4	4.2	4	4

問 34

 $X \ge Y$ を正の相関を持つ連続確率変数とする。E(X)および E(Y)をそれぞれ $X \ge Y$ の期待値とする。正しいものを選択肢の中から一つ選べ。

- $1 \quad E(X-Y) = E(X) E(Y)$
- $2 \quad E(X-Y) = E(X) + E(Y)$
- $3 \quad E(X-Y) > E(X) E(Y)$
- 4 E(X-Y) > E(X) + E(Y)

以下の(a)~(c)に入れる語として適切なものを選択肢の中から一つ選べ。

正規分布に従う母集団の平均がある値 μ_0 に等しいかどうかを検査する t 検定においては、帰無仮説を「母集団の平均が μ_0 に等しい」として(a)を行う。得られた p 値が 0.027 であったとする。このとき、有意水準 0.01 において、帰無仮説は(b)され、有意水準 0.05 において、帰無仮説は(c)される。

1	(a)両側検定	(b)受容	(c)棄却
2	(a)片側検定	(b)受容	(c)受容
3	(a)片側検定	(b)棄却	(c)棄却
4	(a)両側検定	(b)棄却	(c)受容

間 36

正規分布 N(m,v)の平均パラメータ mおよび分散パラメータ vを最尤推定によって求めることを考える。次の説明文として、不適切なものを選択肢の中から一つ選べ。

- 1 標本数を増やすと平均パラメータ mの最尤推定値が増加することがある。
- 2 標本数を増やしても分散パラメータ v の最尤推定値が減少するとは限らない。
- 3 標本数を増やしても分散パラメータ v の最尤推定値は母分散に近づくとは限らない。
- 4 母分散は正であっても、分散パラメータ vの最尤推定値は負の値をとることがある。

問37 全員正解としました

以下の(\mathbf{a})~(\mathbf{c})にそれぞれ入る式の組み合わせとして、もっとも適切なものを選択肢の中から一つ選べ。

結果に対する原因の確率を計算する公式としてベイズの定理が知られている。Dを得られた結果、 H_1 、 H_2 、…, H_k を原因とする。ただし、 H_i (i=1, 2, …, k)は互いに排反で、それらの和事象は全事象であると仮定する。ベイズの定理では、事後確率(a)を計算するために、事前確率(b)と確率(c)を利用する。

```
1 (a) P(H_i \mid D) (b) P(H_i) (c) P(D \mid H_i)

2 (a) P(D \mid H_i) (b) P(D) (c) P(H_i \mid D)

3 (a) P(H_i \mid D) (b) P(H_i) (c) P(D \mid H_i)

4 (a) P(D \mid H_i) (b) P(D) (c) P(H_i \mid D)
```

当初はベイズの定理に関する語句の組み合わせを含んだ問題でしたが、問題文の推敲課程で違いを生じる設問を削除してしまったため、結果的に正解をひとつに絞れない選択肢となってしまいました。そのためこの問題は全員正解とします。

以下の(a)~(c)に入れる語として適切なものを選択肢の中から一つ選べ。

ある特定の機能を持つタンパク質を判別する手法を評価したいとする。TP を真陽性数、FP を偽陽性数、FN を偽陰性数、TN を真陰性数とする。性能評価には感度(sensitivity)、特異度(<math>specificity)がよく用いられる。感度は(<math>a)で、特異度は(b)で定義され、感度もしくは特異度が 1.0 を(c)。

1	(a)TP/(TP + FP)	(b)TP/(TP + FN)	(c)超えることはない
2	(a)TP/(TP + FN)	(b)TP/(TP + FP)	(c)超えることがある
3	(a)TP/(TP + FP)	(b)TN/(TN + FP)	(c)超えることがある
4	(a)TP/(TP + FN)	(b)TN/(TN + FP)	(c)超えることはない

問 39

以下の(ア)~(イ)に入れる語として適切なものを選択肢の中から一つ選べ。

線形サポートベクトルマシンの識別関数は $f(\mathbf{x}; \mathbf{w}, b) = \langle \mathbf{w}, \mathbf{x} \rangle + b$ で与えられる。ただし、 $\langle \mathbf{w}, \mathbf{x} \rangle$ は 2 つのベクトル \mathbf{w} と \mathbf{x} の内積である。学習の結果、識別関数の 2 つのパラメータは $\mathbf{w} = (+1, -1)$, b = 0 のように得られたとする。識別関数の閾値を 0 とすると、未知のデータ $\mathbf{x} = (2,1)$ は (\mathbf{r}) と 予測され、 $\mathbf{x} = (1,2)$ は (\mathbf{r}) と 予測される。

- 1 (ア)陽性 (イ)陽性
- 2 (ア)陰性 (イ)陽性
- 3 (ア)陽性 (イ)陰性
- 4 (ア)陰性 (イ)陰性

問 40

次の選択肢の中からクラスタリングを行う手法として、もっとも不適切なものを選択肢の中から一つ選べ。

- 1 k-平均法 (k-means method)
- 2 UPGMA (Unweighted Pair Group Method with Arithmetic mean)
- 3 サポートベクトルマシン (Support Vector Machine)
- 4 自己組織化マップ (Self-Organizing Map)

次に示したデータベース名とその内容の組み合わせの中で、もっとも不適切なものを選択肢の中から一つ選べ。

1 HapMap 遺伝子発現データベース

2 UniProt タンパク質配列データベース

3 Pfam タンパク質モチーフ・ドメインデータベース

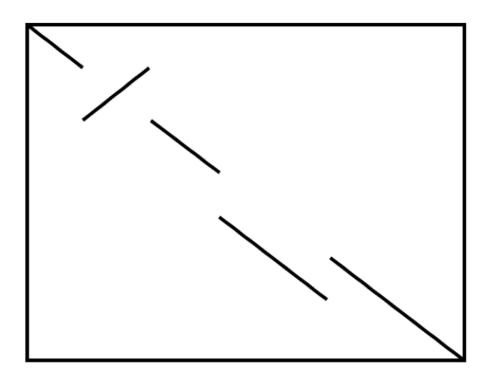
4 Ensembl 統合ゲノムデータベース

問 42

生物種 A、B のゲノムの相同な領域についてドットプロット解析を行ったところ、以下の図が得られた。 図の解釈として、もっとも不適切なものを選択肢の中から一つ選べ。

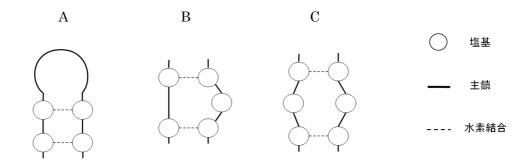
生物種Aのゲノム配列

生物種Bのゲノム配列



- 1 いずれかの生物種で逆位が起こっている。
- 2 生物種 A に欠失が起こっているか、または、生物種 B に挿入が起こっている。
- 3 生物種 A に重複が起こっているか、または、生物種 B に欠失が起こっている。
- 4 両生物種に同一のパリンドローム配列が存在する。

以下に示す RNA 二次構造のモチーフ A \sim C にもっとも適した用語の組み合わせを選択肢の中から一つ 選べ。



1 A:内部ループB:多重ループC:バルジループ2 A:バルジループB:ヘアピンループC:内部ループ3 A:ヘアピンループB:バルジループC:内部ループ4 A:ヘアピンループB:内部ループC:バルジループ

問 44

以下の文は、配列相同性検索ツール BLAST に含まれるプログラムにおいて、クエリ (問い合わせ配列) の種類と、それを用いて実行したい事柄の組み合わせについて述べたものである。もっとも不適切なものを選択肢の中から一つ選べ。

選択肢	プログラム名	クエリの種類と実行したい事柄
	11	ある生物種のタンパク質遺伝子の塩基配列をクエリにして、遠縁種の
1 blastx		ホモログを見つけたい
	1.1	ある生物種の非コード RNA 遺伝子の塩基配列をクエリにして、他の生
$\frac{2}{2}$	blastn	物種でホモログを見つけたい
	. 1.1	ある生物種のタンパク質遺伝子のアミノ酸配列をクエリにして、遠縁
3 psi-blast		種のホモログを見つけたい
4	blastp	ある生物種の非コード RNA 遺伝子の塩基配列をクエリにして、遠縁種
4		のホモログを見つけたい

Burrows-Wheeler 変換 (BWT) に関する以下の記述にあてはまる語の組み合わせで、もっとも適切なものを選択肢の中から一つ選べ。

BWT は、文字列の圧縮率を高めるための(a)変換方法のひとつである。次世代シークエンサの配列解析ツールの多くでは、BWT を用いることで、標的ゲノムの塩基配列データベースの高速な検索を可能にしている。長さ L の配列をゲノム中で検索するには、ゲノムの大きさによらず(b)時間で完了する。

- 1 (a) 不可逆
- (b) $O(L \log L)$
- 2 (a) 可逆
- (b) $O(L \log L)$
- 3 (a) 可逆
- (b) O(L)
- 4 (a) 不可逆
- (b) O(L2)

問 46

Sequence Read Archive (SRA) は、日米欧の各拠点で運営される、次世代シークエンサからの出力および関連データの国際公共データベースである。SRAのデータ登録においては、信頼性が確認できるようにクオリティスコアを付加した数種類のデータフォーマットが使用される。クオリティスコアを含まないために登録に不適切なデータフォーマットを選択肢の中から一つ選べ。

- 1 FASTQ
- 2 SFF
- 3 BAM
- 4 FASTA

高速な配列比較を行うために、長さkの短い部分配列(k-mer)が完全一致する箇所を検索し、その周辺でより正確なアラインメントを求める方法がある。以下の $(a)\sim (d)$ を埋める語句の組み合わせとして、もっとも適切なものを選択肢の中から一つ選べ。

k-mer は塩基配列では 4^k 種類、アミノ酸配列では(a)種類ある。k-mer を長くすると偽陽性のヒットが(b)高速化が期待できるが、検出感度は(c)する。また、完全一致する k-mer を高速に検索する技術の一つに(d)がある。

1	(a) 20^{k}	(b)増えて	(c)低下	(d)k-平均法
2	(a) 20^{k}	(b)減って	(c)低下	(d)ハッシュ表
3	(a) 3^{k}	(b)減って	(c)低下	(d)ハッシュ表
4	(a) 20^{k}	(b)増えて	(c)上昇	(d)k-平均法

配列 A,B,C 間の多重配列アラインメントから、下表に示すような 3 配列間の距離データが得られたとする。

表:3配列間の距離

	A	В	C
A	_	30	40
В	_	_	50
\mathbf{C}	_	_	_

近隣結合法などの距離行列に基づく系統樹作成法においては、これを下図に示すような無根系統樹上での距離に当てはめることにより、枝の長さを決定するのが基本的な考え方となっている。図の距離 a の値としてもっとも適切なものを選択肢の中から一つ選べ。

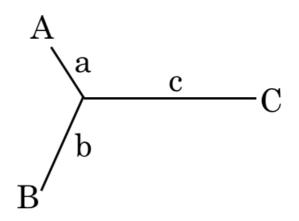
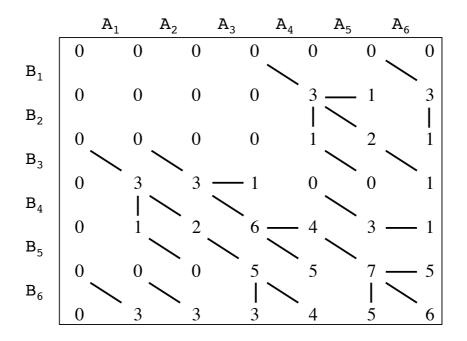


図:3配列の無根系統樹

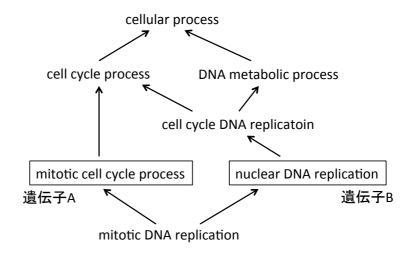
- 1 10
- 2 20
- 3 50
- 4 0

以下の図は、それぞれ長さ 6 塩基の 2 本の DNA 配列 A, B (ここでは各文字列を $A_1...A_6$, および $B_1...B_6$ として表している)を、スミス・ウォーターマン (Smith-Waterman) 法を用いてアラインメントした ときに、計算の過程で作成されるスコアおよびパスの行列を表している(計算は左上隅から右下隅に向かって行われている)。これに基づいて構築される局所的アラインメントとして、もっとも適切なものはどれか。選択肢の中から一つ選べ。



- $1 \quad A_2 \quad A_3 \\ B_3 \quad B_4$

以下の図は Gene Ontology で定義されている用語(GO ターム)間の概念階層の一部を抜き出したもので、矢印($X \rightarrow Y$)は is_a の関係(「X は Y の一種である」)を表している。いま、関心のある 2 つの遺伝子 A,B について、その機能としてそれぞれ "mitotic cell cycle process" および "nuclear DNA replication" という GO タームがつけられているとする。このとき、この 2 つの遺伝子に共通の機能をもっとも的確に表す GO タームはどれか。選択肢の中から一つ選べ。



- 1 cellular process
- 2 cell cycle process
- 3 cell cycle DNA replication
- 4 mitotic DNA replication

問 51

次世代シークエンサは様々な用途に用いられるが、そのデータ解析の多くは、読み出された大量の短い配列(リード read)をリファレンス配列上に位置付けるマッピング処理が基本となっている。以下のうち、リード配列のマッピング処理に基づく解析としては、もっとも不適切なものを選択肢の中から一つ選べ。

- 1 エキソームのリシークエンシングによる、エキソン配列中の一塩基変異の検出
- 2 RNA-Seq 解析による、ヒトの組織特異的な選択的スプライシングパターンの検出
- 3 ChIP-Seg 解析による、転写因子のゲノム上の結合部位の探索
- 4 de novo アセンブリによる、新規に分離された微生物の全ゲノム配列決定

問 52

4 塩基からなる DNA 配列モチーフが以下のような重み行列で表現されている。

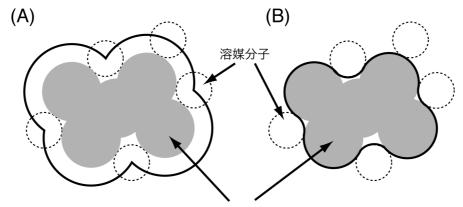
位置	1	2	3	4
A	-35	5	-35	11
С	8	-29	-29	-22
G	-8	9	15	-29
Т	4	-17	-8	-34

この重み行列を用いて、7 塩基の配列 CGTGTAG を検索したときに、もっとも高いスコアを示す部分 配列はどれか。選択肢の中から一つ選べ。

- 1 CGTG
- 2 GTGT
- 3 TGTA
- 4 GTAG

以下の記述の(\mathbf{a})~(\mathbf{d})に入る語句の組み合わせとして、もっとも適切なものを選択肢の中から一つ選べ。

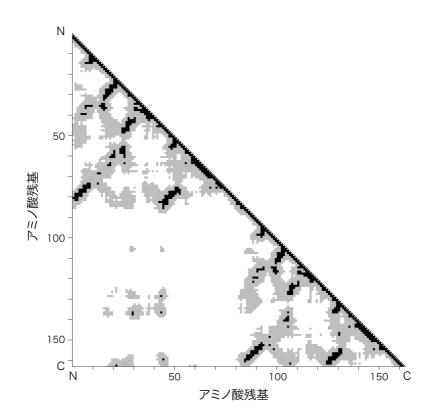
下に示した二次元の模式図Aの太線のように、球に模した溶媒分子を溶質分子のファンデルワールス球群に接するようにくまなく転がしたときの球の中心がなぞる面は、(a)表面といわれる。一方で、模式図Bの太線のように、溶媒分子の球の表面がなぞる面は、(b)表面といわれる。また、結合した二つの溶質分子が解離する過程においては、二つの溶質分子の(a)表面の総面積は(c)する。このことは、この過程の前後で解離する溶媒分子の数の方が結合する溶媒分子の数よりも(d)ことを示す。



溶質分子のファンデルワールス球群

1 (a)コノリー(b)溶媒接触(c)減少(d)多い2 (a)コノリー(b)溶媒接触(c)増加(d)少ない3 (a)溶媒接触(b)コノリー(c)減少(d)多い4 (a)溶媒接触(b)コノリー(c)増加(d)少ない

下図は、あるタンパク質のコンタクトマップであり、 C_{α} 原子間距離が 8Å 以下のアミノ酸残基ペアを黒で、16Å 以下のアミノ酸残基ペアを灰色で塗っている。このコンタクトマップから読み取れること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 このタンパク質は2つのドメインから構成されている。
- 2 このタンパク質のもっとも N 末端側にあるドメインともっとも C 末端側にあるドメイン同士には コンタクトがみられる。
- 3 このタンパク質を構成するすべてのドメインは立体構造が類似している。
- 4 このタンパク質のもっとも N 末端側にあるドメインにおいて、ドメイン内の N 末端領域と C 末端 領域は平行 β シート様の構造を形成している。

タンパク質主鎖の折り畳みパターンをフォールドという。天然タンパク質に頻繁にみられる以下のフォールドのうち、βストランドのみで構成されるものを選択肢の中から一つ選べ。

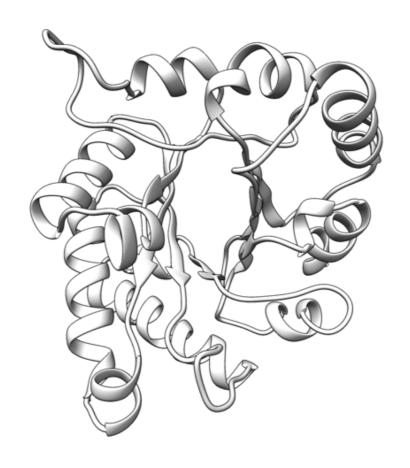
- 1 グロビンフォールド
- 2 免疫グロブリン (Ig) フォールド
- 3 TIM バレルフォールド
- 4 フェレドキシンフォールド

問 56

PDB (Protein Data Bank) が提供するファイルフォーマット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。

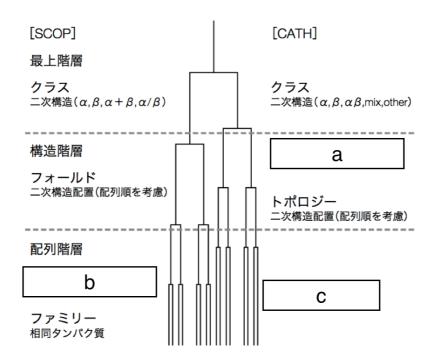
- 1 PDBML は従来の PDB フォーマットよりもファイルサイズが小さく抑えられており、また、内容 を人間が目で見て確認しやすくなっている。
- 2 PDBx/mmCIF は STAR 形式 (self-defining textarchive and retrieval) 文法で記述されている。
- 3 PDBx/mmCIF はデータの定義が明確なフォーマットであり、人間が読んでも理解しやすくコンピュータによる処理にも向いている。
- 4 PDBMLはmmCIFの内容をXMLフォーマットで記述したものである。実データとなる文字列に"<" と"/>"で囲まれた標識を埋め込むことで、表示ソフトに対して文書構造や書式、他のサイトへのリンク情報などを指定することができる。

タンパク質フォールドを分類したデータベースに SCOP (Structural Classification Of Proteins) がある。以下のリボンモデルで表されたタンパク質は SCOP でどのように分類されるか。クラスとフォールドの正しい組み合わせを選択肢の中から一つ選べ。



選択肢	クラス	フォールド
1	$\alpha + \beta$	ロスマンフォールド
2	$\alpha + \beta$	TIM バレルフォールド
3	α / β	ロスマンフォールド
4	α / β	TIM バレルフォールド

下図は SCOP と CATH のフォールド分類階層を図示したものである。a,b,c に入る語としてもっとも適切な組み合わせを選択肢の中から一つ選べ。



- 1 (a)スーパーファミリー
- 2 (a)アーキテクチャー
- 3 (a)フォールド
- 4 (a)フォールド

- (b)ホモロジー
- (b)スーパーファミリー
- (b)ホモロジー
- (b)スーパーファミリー
- (c)ファミリー
- (c)ホモロジー
- (c)アーキテクチャー
- (c)ファミリー

二つのタンパク質のアミノ酸配列の一致度、立体構造の違い(ここでは C_{α} に関する Root Mean Square Deviation (RMSD) と定義する)、および相同性についての記述のうち、もっとも不適切なものを選択肢の中から一つ選べ。

- 1 アミノ酸配列の一致度が 30%以上であれば、RMSD は数 Å 以下であることが多い。
- 2 アミノ酸配列全体の一致度が 30%程度であっても、機能を担う領域に限定すると、配列全体の一致度よりも一致度が高く、RMSD の値が小さい場合が多い。
- 3 アミノ酸配列の一致度が 100%であっても、実験条件の違いなどによって、RMSD の値が 0 ではない場合がある。
- 4 RMSD が数 Å 以下であっても、アミノ酸配列の一致度が 30%未満であれば、相同であるとはいえない。

間 60

立体構造を推定したいタンパク質のアミノ酸配列があり、PDB に登録された分子に対して BLAST による配列検索をしたところ、有意な類似性を示すタンパク質が複数見つかった。この後に行うタンパク質の立体構造の予測法としてもっとも適切なものを一つ選べ。

- 1 ホモロジーモデリング法
- 2 フォールド認識法
- 3 分子動力学法
- 4 フラグメントアセンブリー法

問 61

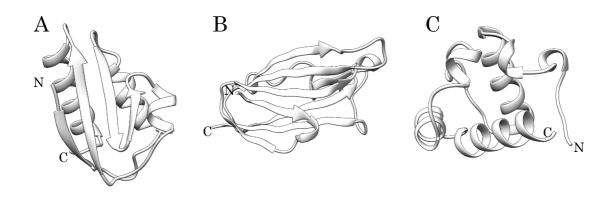
タンパク質の立体構造予測についての記述のうち、もっとも不適切なものを選択肢の中から一つ選べ。

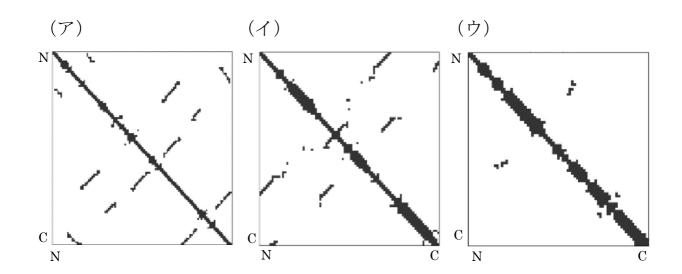
- 1 立体構造を推定したいタンパク質のアミノ酸配列があるが、PDB に登録された分子に対して BLAST による配列検索をしたところ、有意な類似性を示すタンパク質は見つからなかったので *ab initio* 構造予測法を使用した。
- 2 ホモロジーモデリング法では、標的と鋳型のアミノ酸配列の一致度が高いほど一般に予測精度は高く、100 残基をこえてアミノ酸配列一致度が 30%を上回る場合が良好なモデルを構築するための目 安の1つである。
- 3 フォールド認識法は、相同なタンパク質のフォールドが既知である場合のみに使用される。
- 4 タンパク質の立体構造予測は生体分子の理解と制御に重要であり、医薬品の設計 (ドラッグデザイン) にも利用されている。

以下に示す実験手法のうち、タンパク質の立体構造決定にもっとも不適切なものを選択肢の中から一つ 選べ。

- 1 X 線回折法
- 2 NMR(核磁気共鳴)分光法
- 3 MRI(核磁気共鳴画像法)
- 4 電子顕微鏡法

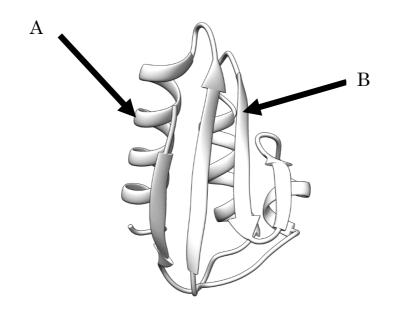
以下の3つのリボンモデルで表されたタンパク質A,B,Cについて、コンタクトマップ(P)、(イ)、(ウ)を計算した(順不同)。コンタクトマップでは、アミノ酸残基の C_{α} 原子間の距離が $6\mathring{A}$ 以下の場合に対 応するマス目が黒く塗られている。タンパク質とコンタクトマップの正しい組み合わせを選択肢の中か ら一つ選べ。

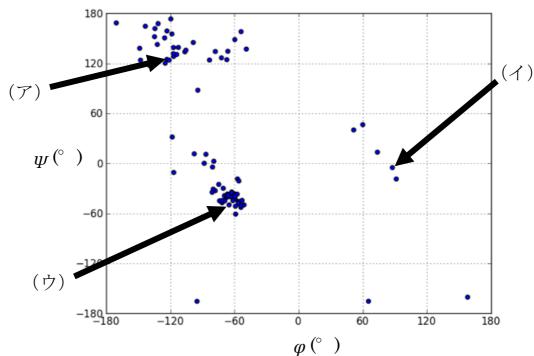




- B (イ) C (ウ) $1 \quad A - (\mathcal{T})$
- B (ウ) C (ア) 2 A − (≺)

タンパク質の主鎖の二面角 φ と ψ の二次元分布図はラマチャンドランマップ(Ramachandran map)と 呼ばれる。以下にタンパク質のリボンモデルとそのラマチャンドランマップを示す。矢印 A,B で示され たアミノ酸残基はラマチャンドランマップの矢印(ア)、(イ)、(ウ)のどれに対応するか。正しい組み 合わせを選択肢の中から一つ選べ。





- $1 \quad A (\mathcal{T}) \qquad B (\mathcal{T})$
- 2 A (イ) B (ウ)
- 3 A (ウ) B (ア)
- $4 \quad A (\mathcal{I}) \qquad B (\mathcal{T})$

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。

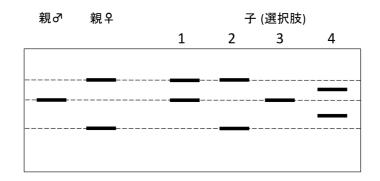
- 1 連鎖不平衡とは、2つの遺伝子座の対立遺伝子の特定の組み合わせの頻度が有意に高くなる現象である。
- 2 連鎖不平衡は、同一染色体上にある2つの遺伝子座の距離が遠いほど強くなる傾向がある。
- 3 2 つの遺伝子座が同一染色体に存在する場合のほか、交配がランダムでない場合や特定のハプロタイプが生存に有利な場合にも、連鎖不平衡が生じる。
- 4 LOD スコアを用いた家系分析を行うことにより、遺伝的疾患の原因遺伝子が連鎖地図上で遺伝子マーカーからどの程度離れた位置にあるかを推定できる。

問 66

遺伝的多型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 多型を利用して特定の形質や由来を持つ個体の検出に利用できる DNA 部位を、遺伝子マーカーまたは DNA マーカーとよぶ。
- 2 1 塩基多型 (SNP) には、表現型に影響を与えるものと与えないものがある。
- 3 1 塩基多型 (SNP) には、体細胞のゲノム中に生じた突然変異も含まれる。
- 4 制限酵素断片長多型や反復配列多型を示す電気泳動パターンは、指紋のように個人の特定に利用できることから、DNA フィンガープリントとよばれることがある。

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が疑われるマウス(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。この両親の子供としてもっとも妥当な子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。



問 68

分子系統樹に関する以下の記述について、(a) と(b) 内に入る語句の組み合わせとしてもっとも適切なものはどれか。選択肢の中から一つ選べ。

分子系統樹は、塩基配列やアミノ酸配列の違いをもとに、それぞれの (a) 関係を木構造で表現したものである。系統樹作成に用いた遺伝子やタンパク質同士よりも遠い関係にあることがあらかじめ分かっているものを (b) として加えることで、無根系統樹の根の位置を推定することができる。

- 1 (a)親子 (b)外群
- 2 (a)親子 (b)祖先型
- 3 (a)類縁 (b)外群
- 4 (a)類縁 (b)祖先型

オーソログの推定に関する以下の記述について、(a) と (b) 内に入る語句の組み合わせとしてもっとも適切なものはどれか。選択肢の中から一つ選べ。

遺伝子間の関係が、種分化の過程で生じたオーソログであることを判断する基準として(a) がよく採用されるが、これは(b) によりパラログが多数存在する場合には必ずしも有効な方法でないので、系統樹を作成し、そこから推定される進化過程にもとづいてオーソログを決めることが望ましい。

1 (a)アラインメントスコア (b)選択的スプライシング

2 (a)双方向ベストヒット (b)選択的スプライシング

3 (a)アラインメントスコア (b)遺伝子重複

4 (a)双方向ベストヒット (b)遺伝子重複

問 70

以下の文章は、いくつかの系統樹推定法についてその特徴を述べたものである。文章中の(a) と (b) 内に入る語句の組み合わせとして、もっとも適切なものを選択肢の中から一つ選べ。

系統樹推定法には大きく分けて、平均距離法(UPGMA 法)や(a)のように距離行列を用いる方法と、最尤法(ML 法)のように配列上の文字置換を直接利用する方法がある。UPGMA 法はすべての 枝で進化速度が一定であることを暗に仮定しているため有根系統樹が得られるが、(a)や ML 法は 進化速度の一定性を仮定しないため、得られるのは無根系統樹である。オーソログとパラログが混在する複合系統樹を推定する方法としては、(b)を用いるのが妥当である。

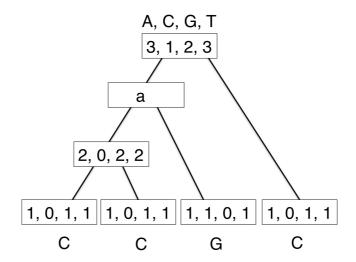
1 (a)近隣結合法 (NJ 法) (b)前者

2 (a)近隣結合法 (NJ 法) (b)後者

3 (a)最大節約法 (MP法) (b)前者

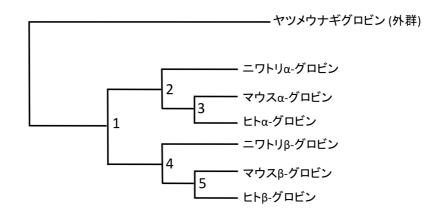
4 (a)最大節約法 (MP法) (b)後者

以下の図は、ある樹形の系統樹の特定の塩基サイトでの最小置換数を、最大節約法(MP 法)により求めた結果を示している。図中で最下段の C や G は現存遺伝子の塩基を、四角内の数字は対応するノードでの塩基がそれぞれ A, C, G, T であった場合の最小置換数を示している。例えば、空欄(a) に相当するノードの塩基が A であった場合は、少なくとも 2 回の置換を経なければ、このノードから分岐した現存遺伝子の塩基(C, C, G)のパターンを説明できないので、最初の数値は 2 である。空欄(a) に入る数値の組み合わせとして、もっとも適切なものを選択肢の中から一つ選べ。



- 1 2, 1, 1, 2
- 2 2, 1, 2, 3
- 3 2, 0, 1, 2
- 4 2, 1, 2, 2

以下の図はグロビン族の遺伝子を用いて作成した分子系統樹であり、数字はノードを示す。この系統樹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 ヤツメウナギグロビンを外群として使用した場合、 α -および β -グロビン族のルート(根)はノード 1 に相当する。
- 2 ノード2とノード5は同一の種分化イベントに相当する。
- 3 この系統樹上では、遺伝子重複は1回だけ起こっている。
- 4 α-および β -グロビン族遺伝子の出現は、ほ乳類の種分化に先立って起こった。

問 73

メタゲノム解析では、ある環境に生息する生物群の DNA を抽出し網羅的に配列決定することで、そこにどんな生物種がいてどのような機能を持っているかを推定することができる。メタゲノムを利用した細菌叢解析の手順について、もっとも不適切なものを選択肢の中から一つ選べ。

- 1 次世代シークエンサから得られた配列は精度が高いので、得られた配列全てを以降の解析ステップに用いることができる。
- 2 得られた配列を de novo でアセンブルしてコンティグを作成する。
- 3 BLAST 等でデータベース検索を行い、既知の遺伝子やゲノムとの対応を取ることで、どんな遺伝子を持つ細菌がどれだけいるのか推定する。
- 4 推定された遺伝子群を KEGG 等のパスウェイデータベースに当てはめることで、得られた環境全体でどのようなパスウェイが顕著に使われているか観察する。

次世代シークエンサの説明としてもっとも不適切なものを選択肢の中から一つ選べ。

- 1 Illumina 社の HiSeq シリーズは、1 ランあたりのリード数が多く、個人ゲノムの決定に使われている。
- 2 PacBio 社の RS シリーズは、リード長が長いので、ゲノム上のリピート配列をまたいで読むことができる可能性が高くなる。
- 3 Roche 社の GS シリーズは、1 ランあたりのリード数は少ないが、1 リードあたり数百塩基の長さを読むことができる。
- 4 100bp 程度の短い配列しか読めない機種を用いても、無限に配列数を増やしていくと、個人のゲノムを染色体ごとに一本ずつのゲノムとして推定できる。

問 75

オーミクス解析に使われる実験機器とその目的の組み合わせとして、もっとも不適切なものを選択肢の中から一つ選べ。

【実験機器】 【目的】

1 DNA マイクロアレイ 遺伝子発現量の解析

2 次世代シークエンサ ChIP-Seq 解析

3 質量分析装置 代謝産物の解析

4 キャピラリーシークエンサ メタボローム解析

問 76

転写制御解析やエピゲノム解析でよく行われる解析手法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一つ選べ。

- 1 ゲノム上の転写因子結合部位を推定するため、抗転写因子抗体を用いた ChIP-Seq 解析が行われている。
- 2 DNA 上のメチル化部位を推定するために、抗メチル化 DNA 抗体を用いた DIP-Seq 解析が用いられる。
- 3 ヒストンの修飾に応じた抗体を使って ChIP-Seq を行うことで、ゲノム上のヒストン修飾部位が推 定できる。
- 4 ヌクレオソームが DNA 上にどのように配置されているかを調べるために、バイサルファイトシー クエンス法が用いられる。

クラスター分析手法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。

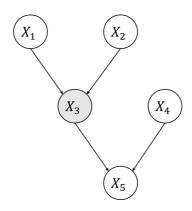
- 1 階層型クラスタリングでは事前にクラスター数を指定しなくても分類が可能である。
- 2 k-平均法は結果が初期値によって異なることがある。
- 3 サポートベクトルマシンによる分類は教師なし学習である。
- 4 混合ガウスモデルは潜在変数を含んだモデルとみなすことができる。

問 78

ベイジアンネットワーク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 ベイジアンネットワークは、確率変数間の条件付独立性を表すモデルである。
- 2 いずれのベイジアンネットワークでも、有向閉路をもつことができる。
- 3 ベイジアンネットワーク上の確率変数は、離散値だけでなく、連続値もとることができる。
- 4 ダイナミックベイジアンネットワークはベイジアンネットワークを時間方向に展開したものである。

下記の図は5変数 X_1 , X_2 , X_3 , X_4 , X_5 からなるベイジアンネットワークであり、 $X_1 \rightarrow X_2$ は、 X_2 の状態が X_1 の状態に依存していることを表す。変数 X_3 のみ観測されている条件で、他の変数は未観測である。



この時、変数間の条件付独立性の関係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 変数 X_1 と X_2 は確率的に独立である。すなわち $P(X_1, X_2) = P(X_1)P(X_2)$ である。
- 2 変数 X_1 と X_4 は確率的に独立である。すなわち $P(X_1, X_4) = P(X_1)P(X_4)$ である。
- 3 変数 X_1 と X_5 は確率的に独立である。すなわち $P(X_1, X_5) = P(X_1)P(X_5)$ である。
- 4 変数 X_2 と X_4 は確率的に独立である。すなわち $P(X_2, X_4) = P(X_2)P(X_4)$ である。

下記選択肢より、(ア)と(イ)に入る記述としてもっとも適当なものを一つ選べ。

時刻tにおける遺伝子発現量をy(t)として表した連立微分方程式を下記の式で表す。

$$\frac{d}{dt}\mathbf{y}(t) = \mathbf{b} - A\mathbf{y}(t)$$

y(t)の初期値をy(0)とし、Aは次のように対角化可能な行列であるとする。

$$A = P^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P$$

この時、定数 C_1 , C_2 を用いて、y(t)は下記の一般解が得られる。

$$\mathbf{y}(t) = A^{-1}\mathbf{b} + P^{-1} \begin{pmatrix} C_1 \exp(-\lambda_1 t) \\ C_2 \exp(-\lambda_2 t) \end{pmatrix}$$

上記一般解より、 λ_1 と λ_2 の(ア)である時、tの値が十分大きければ(イ)に収束し、定常状態となる。

1 (ア)値が共に正

 $(\ \ \) \ A^{-1}b$

2 (ア) 値が共に負

- $(\ \) \ A^{-1} b$
- 3 (ア) 値のどちらかが負
- $(\langle \ \rangle) y(0)$
- 4 (ア) 値のどちらかが負
- (イ) 0

平成 27 年度
日本バイオインフォマティクス学会(JSBi)
バイオインフォマティクス技術者認定試験
【解答】

問 1 問 2 問3 問 4 問 5 問6 問7 問8 問 9 問 10 2 4 2 2 4 2 2 2 1 1 問 18 問 19 問 11 問 12 問 13 問 14 問 15 問 16 問 17 問 20 4 3 2 4 3 3 3 1 3 3 問 21 問 22* 問 23 問 24 問 25 問 26 問 27 問 28 問 29 問 30 1 3 1 2 3 全員正解 3 2 2 1 問 37** 問 40 問 31 問 32 問 33 問 34 問 35 問 36 問 38 問 39 1 2 1 1 3 4 4 全員正解 4 3 問 44 問 42 問 43 問 45 問 46 問 50 問 41 問 47 問 48 問 49 1 4 3 2 1 2 4 3 4 3 問 51 問 53 問 55 問 60 問 52 問 54 問 56 問 57 問 58 問 59 2 1 2 4 3 4 4 4 4 1 問 61 問 62 問 63 問 64 問 65 問 66 問 67 問 68 問 69 問 70 2 2 3 3 4 3 3 1 3 4 問 71 問 72 問 73 問 74 問 75 問 76 問 77 問 78 問 79 問 80 1 2 1 4 4 3 2 4 1 1

試験問題に記載されている会社名または製品名は、それぞれ各社の商標または登録商標です。	
なお、試験問題では、®および TMを明記していません。	