

数理生命情報学

（10 枚中 1 枚目）

問題紙が 10 枚、解答紙が 5 枚配布されていることを確認すること。

1 問につき 1 枚の解答紙を使用すること。

[1] は必答問題である。必ず解答すること。(60 点)

[2]～[9] は選択問題である。[2]～[9] の中から 4 題を選択して答えよ。(各 35 点)

数理生命情報学

（10 枚中 2 枚目）

[1]

次の 20 の語の中から 4 つを選び、それぞれの語の定義と特徴を記せ。

- (1) コドン
- (2) 真核生物の特徴
- (3) SNPs
- (4) エピジェネティクス
- (5) 酵素
- (6) スクレオチド
- (7) 転写調節因子
- (8) RNA プロセッシング
- (9) 全単射
- (10) 相関係数
- (11) 行列のランク
- (12) 累次 (逐次) 積分
- (13) ポアソン分布
- (14) ヤコビ行列
- (15) 再帰関数
- (16) ニューラルネットワーク
- (17) 仮引数と実引数
- (18) 二部グラフ
- (19) 隣接行列
- (20) 深さ優先探索

数理生命情報学

(10 枚中 3 枚目)

[2]

関数

$$f(x) = \frac{1}{1 + e^{-x}},$$

$$g(x) = \sin(\arctan x)$$

に関して、以下の問いに答えよ。ただし、 $\arctan x$ は $\tan x$ の逆関数を表し、値域を主値 $(-\frac{\pi}{2}, \frac{\pi}{2})$ に制限する。

(1) $f(x)$ の概形を描け。

(2) $\tan(\arctan x) = x$ の関係に注意して、 $g(x) = \frac{x}{\sqrt{1+x^2}}$ であることを示せ。また $g(x)$ の概形を描け。

(3) 広義積分 $\int_{-\infty}^0 f(x)dx$ の値を求めよ。

数理生命情報学

（10 枚中 4 枚目）

[3]

n を正の整数、 p, q, r を 0 以上 1 以下の実数とし、さらに $q + r \leq 1$ とする。また A を下記の 5 次正方行列とする。

$$A = \begin{pmatrix} 1-p & 0 & 0 & 0 & 0 \\ p & 1-p & q & 0 & 0 \\ 0 & p & 1-q-r & p & 0 \\ 0 & 0 & r & 1-p & p \\ 0 & 0 & 0 & 0 & 1-p \end{pmatrix}$$

- (1) A の固有値をすべて求めよ。
- (2) 上記 (1) の解のひとつである固有値 1 に対応する固有ベクトルのうち、各成分の和が 1 となるものを求めよ。
- (3) 表裏がある 4 枚のコインに対して下記の操作を繰り返す。

表向き、裏向きになっているコインの枚数を調べ、場合に応じて以下いずれかを行う。

- 表向きのほうが多い場合は、表向きになっているコインの 1 枚を裏向きにする。
- 裏向きのほうが多い場合は、裏向きになっているコインの 1 枚を表向きにする。
- 表向きと裏向きが同数（2 枚ずつ）の場合は、無作為に 4 枚のうち 1 枚を選ぶ。それが表向きなら何もせず、裏向きなら表向きにする。

操作を n 回繰り返した後に表の枚数が i 枚である確率を $s_{i,n}$ とする（ $n \in \{1, 2, \dots\}$; $i \in \{0, 1, 2, 3, 4\}$ ）。このとき p, q, r の値を適切に選べば下記等式が成立する。

$$\begin{pmatrix} s_{0,n} \\ s_{1,n} \\ s_{2,n} \\ s_{3,n} \\ s_{4,n} \end{pmatrix} = A \begin{pmatrix} s_{0,n-1} \\ s_{1,n-1} \\ s_{2,n-1} \\ s_{3,n-1} \\ s_{4,n-1} \end{pmatrix} = \dots = A^n \begin{pmatrix} s_{0,0} \\ s_{1,0} \\ s_{2,0} \\ s_{3,0} \\ s_{4,0} \end{pmatrix}$$

ここでベクトル $(s_{0,0} \ s_{1,0} \ s_{2,0} \ s_{3,0} \ s_{4,0})^T$ は 4 枚のコインの初期状態をあらわす。はじめに表が m 枚ならば $s_{m,0} = 1$ 、それ以外の成分は 0 である。上記等式が成立する p, q, r の値を求めよ。

- (4) 上記 (3) において、繰り返し数 n を大きくしていくと $s_{i,n}$ はいずれも一定値に収束する。それぞれの確率の収束値を求めよ。

数理生命情報学

(10 枚中 5 枚目)

[4]

以下の生物の特徴に関して書かれた文章を読んで、各問に答えよ。

生物と非生物をわける特徴として、「外界と区分けされていること」「自己複製すること」「代謝をおこなうこと」の3つがよくあげられる。

「外界から区分けされていること」は細胞の構造によって実現されている。(i) によって生物は内側と外側にしきられている。

(i) は (ii) 二重膜によって構成される。この区分けにより細胞内外はイオン濃度がしばしば異なる。この結果として生じる細胞内外の電位差を (iii) という。動物がもつ神経細胞は (iii) を変化させることで、情報処理と情報伝達を行っている。

「自己複製」は、細胞レベルでは、1つの細胞が2つの (iv) 細胞に分裂する細胞分裂によって実現されている。細胞分裂期(M期)に先行してDNA複製期((v) 期)では、(A) DNA複製が行われ、その後 (iv) 細胞に染色体が等量分配される。DNA複製の際には、もとの分子は維持されない。代わりにもとのDNA分子を解体し、そのパーツをもとに半保存的複製が行われる。真核生物においては、(B) 減数分裂とよばれる染色体数が減少する分裂が生殖細胞で行われる。

「代謝」は、生命活動を維持するための一連の化学反応の集合である。代謝の一つの目的はエネルギーを得ることであり、化合物をより単純な物質に変える異化作用が行われる。そうして得たエネルギーをもとに生体材料となる分子を合成する (vi) 作用が行われる。異化作用の一つとして好気呼吸がしられ、解糖系、(vii)、電子伝達系にこれらは分かれる。好気呼吸は (viii) という細胞小器官でよく行われる。代謝で行われる化学反応の多くは (C) 酵素反応である。生命維持に関わる基礎的な代謝経路は一次代謝、それ以外を二次代謝とよぶ。

- (1) 文章中 (i) から (viii) に入る最も適切な語を記せ。
- (2) 下線部(A)のDNA複製の仕組みをDNAの構造を踏まえて説明せよ。
- (3) 下線部(B)の減数分裂に関連して、生物種内の個体の多様性に減数分裂は寄与する。このメカニズムを述べよ。
- (4) 下線部(C)の酵素反応に関連して、酵素はどのような役割をもっているか説明せよ。
- (5) 進化系統樹を作成する際には、一次代謝または二次代謝どちらに関わる遺伝子に着目するのが良いか?理由を添えて答えよ。
- (6) R. ドーキンスは自己複製による情報伝達はDNAによる方法以外にも文化的レベルで生じると指摘した。このアイデアを説明せよ。

数理生命情報学

（10 枚中 6 枚目）

[5]

以下の塩基配列は、シロイヌナズナの葉緑体 DNA のうち、*rbcL* という遺伝子の一部を含む領域である。この配列に関する以下の設問に答えなさい。

```

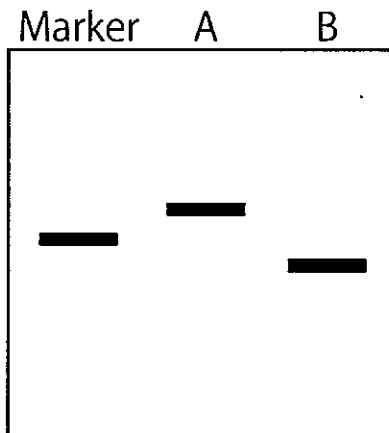
5' - ATGTCGAGTA GACCTTGTG TTTTGTTTTA TTGCAAGAAT TCTAAATTCA TGACTTGTAG
      GGAGGGACTT ATGTCACCAC AAACAGAGAC TAAAGCAAGT GTTGGGTTCA AAGCTGGTGT
      TAAAGAGTAT AAATTGACTT ACTATACTCC TGAATATGAA ACCAAGGATA CTGATATCTT
      GGCAGCATTG CGAGTAACTC CTCAACCTGG AGTTCCACCT GAAGAAGCAG GGGCTGCGGT
      AGCTGCTGAA TCTTCTACTG GTACATGGAC AACTGTGTGG ACCGATGGGC TTACCAGCCT
      TGATCGTTAC AAAGGACGAT GCTACCACAT CGAGCCCGTT CCAGGAGAAG AAACCTCAATT
      TATTGCGTAT GTAGCTTATC CCTTAGACCT TTTTGAAGAA GGTTCGGTTA CTAACATGTT
      TACCTCGATT GTGGGTAATG TATTGGGTTT CAAAGCCCTG GCTGCTCTAC GTCTAGAGGA
      TCTGCGAATC CCTCCTGCTT ATACTAAAC TTTCCAAGGA CCACCTCATG GTATCCAAGT
      TGAAAGAGAT AAATTGAACA AGTATGGACG TCCCCTATTA GGATGTACTA TTAAACCAAA
      ATTGGGGTTA TCCGCTAAAA ACTATGGTAG AGCAGTTTAT GAATGTCTAC GTGGTGGACT
      TGATTTTACC AAAGATGATG AGAATGTGAA CTCCCAACCA TTTATGCGTT GGAGAGACCG - 3'

```

- (1) 以下のプライマーセットを用いて、この領域の一部を PCR 法で増幅した。期待される PCR 産物の長さを答えなさい。

Forward: 5' -ATGTCACCACAAACAGAGACTAAAGC- 3'
Reverse: 5' -GTAAATCAAGTCCACCACG- 3'

- (2) (1) の PCR 増幅は成功し、もう一つ別の PCR 産物と、長さが 500bp であることがわかっているマーカー DNA とあわせて PCR 産物をアガロースゲルを用いて電気泳動したところ、以下のような泳動像が得られた。黒いバンド部分は DNA が存在することを示す。レーン A と B のうち、(1) で増幅した PCR 産物のレーンはどちらか。また、その理由も答えなさい。なお、泳動は、PCR 産物名が示されているゲルの上側がマイナス極、下側がプラス極で行った。



- (3) 以下は 4 種類の植物の *rbcL* 翻訳産物をアライメントしたアミノ酸配列の一部を示す。各位置のアミノ酸は対応しておりギャップはない。これらの 4 つのアミノ酸配列を用いて系統樹を作成する手順を示し、系統樹を図示しなさい。ただし、線の長さは正確な進化距離を反映する必要はなく、線は直線でなくてよい。

シロイヌナズナ: HIEPVPGEETQFIA
トマト: RIERVVGEKDQYIA
ワタ: DIEPVPGEEDQYIC
イチヨウ: DIEPVPGEENQFIA

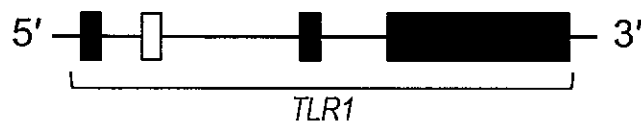
数理生命情報学

（10 枚中 7 枚目）

[6]

古代 DNA の解析から、数万年前までネアンデルタール人やデニソワ人などの現生人類とは別のヒト族の系統が生きていたことがわかってきている。人類進化研究に関する以下の設問に答えなさい。

- (1) 古代 DNA を用いた系統解析にはミトコンドリア DNA と核 DNA の両者が使われてきた。解析におけるそれぞれの利点と欠点を説明しなさい。
- (2) ユーラシア大陸やメラネシアに進出した現生人類集団のゲノムには、数%のネアンデルタール人やデニソワ人のゲノム情報が含まれていることがわかっている。現生人類とネアンデルタール人やデニソワ人との関係について推測できることを述べなさい。
- (3) 以下の模式図は、自然免疫において細菌やウイルスなどの特徴的な構造を見分けるセンサーとして働く Toll 様受容体をコードする *TLR1* という遺伝子のゲノム上の遺伝子構造を示している。



ボックス部分はエクソンを表している。白いボックスで表されたエクソンは一塩基多型の違いにより、成熟 mRNA に含まれる場合と含まれない場合がある。ある研究によると、白いボックスで表された部分のある塩基がネアンデルタール人が持っていたと思われる C である場合はこのエクソンは成熟 mRNA に含まれやすいが、現生人類で一般的な G であれば含まれにくい傾向があることが示されている。また、この部分が mRNA に含まれやすいかどうかは、発現する組織によっても異なる。

- (a) エクソンに挟まれ成熟 mRNA に反映されない遺伝子領域を何と呼ぶか。
 - (b) ゲノム上では同じ DNA 配列から異なる成熟 mRNA が生まれる仕組みの利点を答えなさい。
- (4) 主にユーラシア大陸北部から発掘された古代人骨から抽出された DNA を用いて人類進化の過程が明らかになってきたが、人類進化の全体像を把握する上で残されている課題はどのようなことだと考えられるか。

数理生命情報学

（10 枚中 8 枚目）

[7]

以下の各問に答えよ。なお採点時には、論理展開の正しさを評価する。

- (1) 異種個体間の共生について、共生以外の共存と区別しながら 100 字以内で説明せよ。
- (2) ある樹木個体について、地上に落ちた種子の量は、その樹木からの距離が近いほど多かった（図 1）。その後、ある成長段階まで育った子の数は、親樹木からやや離れたところにピークのある分布になっていた（図 2）。
- (a) 上記の観察をもとに、ひとつの種子がその成長段階まで育つ確率と、親樹木からの距離との関係を記述せよ。グラフを用いてもよい。
- (b) a のようになる原因について考察を書け。この樹木種と、植食者や病原体との相互作用に注目した考察を書いてもよい。
- (c) 自然林における樹木の種多様性が保持されている理由について考察を書け。

図 1

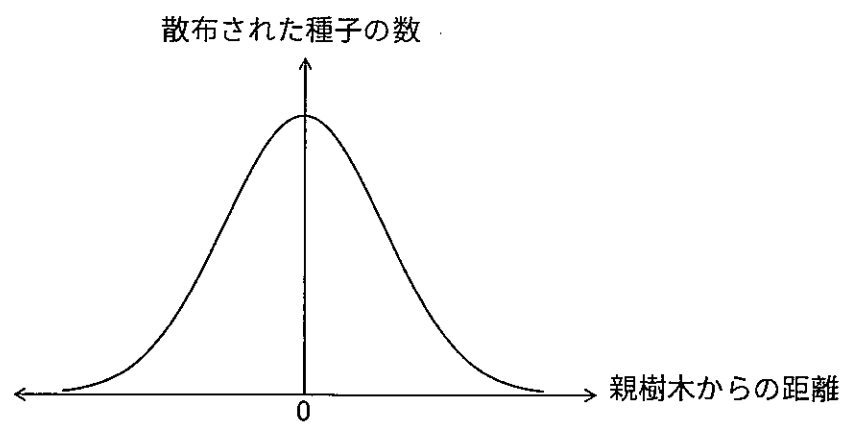
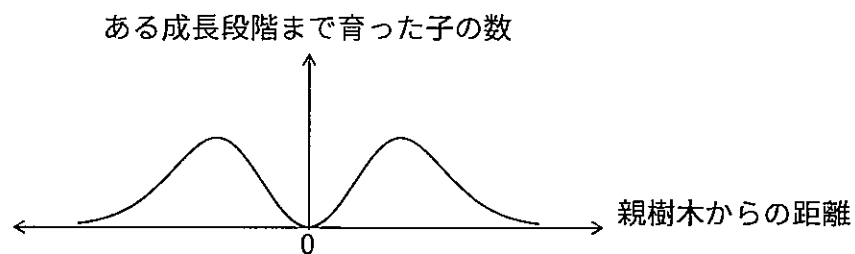


図 2



数理生命情報学

(10 枚中 9 枚目)

[8]

長さ L cm の 1 本のロープをいくつかのロープに切り分け、長さ ℓ cm のロープの価格 v_ℓ 円を

ℓ	1	2	3	4	5	6	7	8	9	≥ 10
v_ℓ	1	4	9	10	12	13	14	15	16	17

に従い販売するとする。ロープの長さ (cm) は常に整数とする。このとき、長さ L のロープの販売価格の最大値を P_L とする。

- (1) $L = 4$ のロープのすべての切り分け方を列挙し、それぞれの販売価格の合計を示しなさい。
- (2) $\ell = 1, \dots, L$ に対して、 P_ℓ を次を用いた漸化式で表しなさい： $v_\ell, P_m (1 \leq m \leq \ell - 1)$, 与えられた複数の引数の最大値を返す関数 $\max$.
- (3) 前問の漸化式を用いて、 P_L を求める計算プログラムを記述しなさい。なお、計算プログラムの表現には、自然言語、数式、疑似プログラミング言語など分かりやすいものを用いてよい。

数理生命情報学

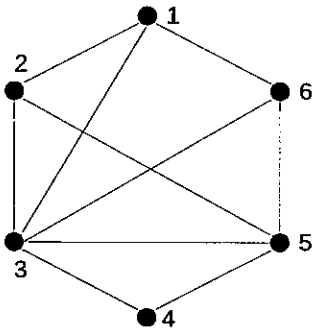
(10 枚中 10 枚目)

[9]

グラフ (graph) $G = (V, E)$ は, n 個の頂点 (node) からなる集合 $V = \{v_1, v_2, \dots, v_n\}$ と, 2 頂点からなる辺 (edge) の集合 E から構成される無向グラフ (undirected graph) である. ただし, 同じ頂点をつなぐ辺 (自己ループ) と多重辺を含まない (つまり, 任意の 2 頂点間の辺は高々 1 個) とする.

頂点 v_i と v_j の間に辺が存在するとき, v_i と v_j は隣接している (adjacent) という. v_i と v_j が隣接しているとき $a_{ij} = 1$ であり, それ以外は $a_{ij} = 0$ である変数 a_{ij} を (i, j) 成分とする $n \times n$ 行列 A_G をグラフ G の隣接行列 (adjacency matrix) とよぶ.

また, 隣接する頂点からなる列をパス (path) といい, それが m 個の頂点を含むとき, その長さを $m - 1$ とする. 例えば, 下図のグラフ H の頂点列 $1, 2, 5, 6, 1, 2, 5$ は長さ 6 のパスである.



H

以下の各問いに答えよ.

- (1) 任意の 2 頂点間に無向辺があるグラフを完全グラフ (complete graph) といい, n 頂点の完全グラフを K_n と記す. K_n の辺の総数を n を用いて表せ.
- (2) 図のグラフ H の隣接行列 A_H を示しなさい.
- (3) グラフ G において, 頂点 i から j への長さ k のパスの総数を $P_G(i, j, k)$ で表すこととする.
 - (a) $P_H(3, 3, 3)$ の値を示せ.
 - (b) 行列 A の k 乗を A^k と表し, k 乗行列 A_G^k の (i, j) 成分を $a_{ij}^{(k)}$ と表すとき, $a_{ij}^{(k)} = P_G(i, j, k)$ が成立することを示しなさい.