

| | | | | | |
|--|--|--|--|--|--|
| | | | | | |
|--|--|--|--|--|--|

東京大学 大学院新領域創成科学研究科
Graduate School of Frontier Sciences, The University of Tokyo
メディカル情報生命専攻
Department of Computational Biology and Medical Sciences
令和 8(2026)年度
2026 School Year
大学院入学試験問題
Graduate School Entrance Examination Question Booklet

専 門 科 目
Specialties

令和 8 年 1 月 13 日(火)
Tuesday, January 13, 2026

10:00~12:00

注意事項 Instructions

- 試験開始の合図があるまで、この冊子を開いてはいけません。
Do not open this booklet until the start of examination is announced.
- 本冊子の総ページ数は 25 ページです。落丁、乱丁、印刷不鮮明な箇所などがあつた場合には申し出ること。
This booklet consists of 25 pages. If you find missing, misplaced, and/or unclearly printed pages, notify it to the staff.
- 解答には必ず黒色鉛筆(または黒色シャープペンシル)を使用しなさい。
Only black pencils (or mechanical pencils) are allowed to answer the questions.
- 問題は 11 題出題されます。問題 1~11 から選択した合計 3 問に解答しなさい。ただし、問題 1~11 は同配点です。
There are 11 exam questions (Question 1 to 11). Answer 3 questions out of the 11 questions. Note that Question 1 to 11 are equally weighted.
- 解答用紙は計 3 枚配られます。各問題に必ず 1 枚の解答用紙を使用しなさい。解答用紙に書ききれない場合は、裏面にわたってもよい。
You are given 3 answer sheets. You must use one answer sheet for each question. You may continue to write your answer on the back of the answer sheet if you cannot conclude it on the front. But you must not proceed to write on the second sheet.
- 解答は日本語または英語で記入しなさい。
Answers should be given in Japanese or in English.
- 解答用紙の指定された箇所に、受験番号と選択した問題番号を記入しなさい。問題冊子にも受験番号を記入しなさい。
Fill the designated blanks at the top of each answer sheet with your examinee number and the question number you are to answer. Fill the designated blanks at the top of this page with your examinee number.
- 草稿用紙は本冊子から切り離さないこと。
The blank pages are provided for making draft. Do not detach them from this booklet.
- 解答に関係ない記号、符号などを記入した答案は無効とします。
An answer sheet is regarded as invalid if you write marks and/or symbols unrelated to the answer on it.
- 解答できない場合でも、解答用紙すべてに受験番号を記入して提出しなさい。
Turn in the answer sheet with your examinee number, even if you cannot solve the question.
- 解答用紙・問題冊子は持ち帰ってはいけません。
Do not take the answer sheets and this booklet out of the examination room.

(このページは草稿用紙として使用してよい)

(Blank page for draft)

(このページは草稿用紙として使用してよい)

(Blank page for draft)

(このページは草稿用紙として使用してよい)

(Blank page for draft)

問題 1

A. 以下の文章を読み、(1)–(10) の問いに答えなさい。

(1) 以下の説明の空欄 (a) ~ (c) に当てはまる語句を答えなさい。

TCA 回路は、真核生物の (a) 内で行われる重要な代謝経路である。グルコースの解糖系で生成された (b) が、アセチル CoA に変換されて回路に入る。回路の最初の反応では、アセチル CoA とオキサロ酢酸が縮合して (c) が生成される。

(2) ある遺伝子座が 6 種類の対立遺伝子を持つとき、取りうるヘテロ接合体の遺伝型の種類の数はいくらか。数字で答えなさい。ただし父親由来と母親由来のアレルは区別できないものとする。

(3) ヒトゲノムについての下記の説明のうち、誤っているものを 1 つ選び、記号で答えなさい。

- A. ヒトゲノムは約 30 億塩基対からなる。
- B. ヒトゲノムには約 10 万のタンパク質コード遺伝子が存在する。
- C. ヒトゲノムの GC 含有率は 50%に満たない。
- D. ヒトゲノムのうち LINE などの反復配列が占める割合は 50%を超える。

(4) 文中の (a) – (d) に入る言葉を以下の語群から選び、a: XXX、b: YYY のように答えなさい。

(a) ワクチンは、標的ウイルスのタンパク質をコードする (a) を合成、投与するものである。細胞内でウイルスのタンパク質を作らせることで免疫記憶を成立させ、感染時に (b) が速やかに (c) 産生細胞に分化し、(c) を産生できるようになる。新型コロナウイルス感染症に対して最初に開発・承認された (a) ワクチンは、ウイルスの (d) タンパク質を標的としている。

語群：siRNA、mRNA、DNA、T 細胞、B 細胞、NK 細胞、抗体、グランザイム、クロマチン、ACE2、スパイク、エンベロープ

(5) 文中の (a)、(b) に入る言葉を以下の語群から選び、a: XXX、b: YYY のように答えなさい。

真核細胞は特定のタンパク質を選択的に分解する仕組みを備えている。まず、標的タンパク質に (a) と呼ばれるタンパク質がいくつも結合し、巨大タンパク質複合体である (b) によって分解される。

語群：アクチン、ビオチン、エクスポーチン、ユビキチン、デスモソーム、ペルオキシソーム、プロテアソーム、スプライソソーム

(問題 1：次ページに続く)

- (6) 1つの白血病細胞が1日に1回分裂して増殖すると、骨髄全体を占拠する(約 2×10^8 個まで増える)のに28日かかる。この時、骨髄の25%を占めるのにかかる日数について、最も適切なものを1つ記号で選べ。ただしこの問題では、白血病細胞は同じ速度で単調に分裂を続けるものとする。また正常細胞との競合、白血病細胞のアポトーシス、腫瘍不均一性は考慮しない。
- A. 3日 B. 7日 C. 14日 D. 26日
- (7) ある遺伝子 X は2量体を形成して機能するタンパク質 Y をコードする。2量体は形成できるが正常な活性を妨害するような変異が遺伝子 X の片方のアレルに生じた。この時、タンパク質 Y の活性は変異がない場合と比較して何%になるか答えなさい。
- (8) 通常の体細胞（分化した細胞）に少数の遺伝子を導入して人工的に作られた多能性の幹細胞を何と言うか。最も適切なものを1つ選びなさい。
- A. ES細胞 C. iPS細胞
B. オルガノイド D. 間葉系幹細胞
- (9) 細胞壁を持つものはどれか、以下の選択肢から全て選び、記号で答えなさい。
- A. ヒトニューロン C. マイコプラズマ E. サルモネラ菌
B. ヒト卵細胞（卵子） D. ノロウイルス F. 酵母菌
- (10) 細胞分裂直後の動物細胞に注目したとき、次の細胞分裂において、以下の4つの過程はどのような順番で起こるか。反応が起こる順番に並べ、 $X \rightarrow Y \rightarrow Z$ のように記号で答えなさい。
- A. 染色体の整列 C. 染色体DNAの複製
B. 細胞質の分裂 D. 姉妹染色分体の分離

B. 以下の文章を読み、(1)–(10) の問いに答えなさい。

- (1) 典型的なコンピュータの浮動小数点数表現形式を用いたとき、正確に表現できる数を以下から全て選べ。
- A. 0.1 B. 0.25 C. 1/3 D. 79
- (2) A, C, G, T の文字からなる長さ 16 の任意の DNA 配列を表現できる、最小ビット数は以下のうちどれか？
- A. 8 B. 16 C. 32 D. 64
- (3) ファイル A とファイル B の内容を連結して結果をファイル C に書き込む UNIX コマンドラインを答えよ。
- (4) インターネットについて、以下の文の中で正しいものを全て選べ。
- A. インターネット上で 2 者が通信を開始すると、両者が専有する通信回線が確保される。
 - B. メールサーバーには到着したメールが保管されている。
 - C. HTTPS による通信は暗号化されておらず、盗聴者が通信の内容を見ることができる。
 - D. グローバル IP アドレスを持たない機器も NAT (Network Address Translation) を用いてインターネットに接続することができる。
- (5) データの散らばり具合を表す指標のうち、「分散」とは何か説明せよ。
- (6) 生物情報の統計解析において「質的変数」を以下の選択肢から全て選び、記号で答えなさい。
- A. 患者の性別 (男性/女性)
 - B. 年齢
 - C. 体重 (kg)
 - D. 疾病分類コード (ICD-11)
 - E. 組織サンプルの種類 (肝臓/腎臓/心臓)

(問題 1 : 次ページに続く)

- (7) 相関関係と統計学的因果関係の違いに関する説明として正しいものを以下の選択肢から全て選び、記号で答えなさい。
- A. 相関があれば必ず因果関係が存在する。
 - B. 相関関係は2つの変数の間に一定の規則的な関係があることを示すが、それが原因と結果を示すとは限らない。
 - C. 因果関係があるとは、原因とされる変数に対して介入を行った場合に、結果変数の確率分布が変化することを意味する。
 - D. 交絡因子があると、相関が見られても因果関係がない場合がある。
 - E. 因果関係があっても、観測データでは相関が現れないことがある。
- (8) 仮説検定における「P値」に関する説明として正しいものを以下の選択肢から1つ、または2つ選び、記号で答えなさい。
- A. P値は、帰無仮説が正しい確率のことである。
 - B. P値は、偶然のみで観測データが出現した確率である。
 - C. P値とは、帰無仮説が正しいと仮定したときに、観測された統計量以上に極端な結果が得られる確率である。
 - D. P値や統計的有意性は、効果の大きさを意味しない。
 - E. 科学的な結論や、ビジネスや政策における決定は、P値がある値を超えたかどうかのみで判定できる。
- (9) 表計算ソフトに以下のような表を作成した。D2, E4 にセル入力すべき式を答えよ。E列には、何れかの回で80点以上を取った受験者については「合格」、それ以外の場合には「不合格」と表示したい。

| | A | B | C | D | E |
|---|-------|-----|-----|-----|----|
| 1 | 氏名 | 第一回 | 第二回 | 最大値 | 合否 |
| 2 | Alice | 60 | 90 | | |
| 3 | Bob | 30 | 80 | | |
| 4 | Chris | 76 | 70 | | |

- (10) 臨床検査において、次のような表を得た。感度と特異度を求めよ。

| | | 真の状態 | |
|------|----|------|-----|
| | | 陽性 | 陰性 |
| 検査結果 | 陽性 | 80 | 10 |
| | 陰性 | 20 | 190 |

(問題1：終わり)

問題 2

A. 以下の (1) ~ (7) の問いに答えなさい。

- (1) タンパク質を構成するアミノ酸は、通常何種類存在するか答えなさい。
- (2) タンパク質を構成するアミノ酸のうち、塩基性アミノ酸を 3つ 答えなさい。
- (3) グリシンとアラニンがペプチド結合で結合したジペプチドの分子量を求めなさい。ただし、アラニンの側鎖はメチル基である。
- (4) アミノ酸やタンパク質は、pH が等電点 (pI) よりも低い溶液中では、全体としてどのような電荷を帯びるかを答えなさい。
- (5) アスパラギン酸は酸性アミノ酸であり、以下の pKa が与えられている： α -カルボキシ基： $pK_{a1}=2.10$ 、 α -アミノ基： $pK_{a2}=9.82$ 、側鎖カルボキシ基 (R 基)： $pK_{a3}=3.86$ 。このとき、アスパラギン酸の等電点 (pI) を求めなさい。
- (6) あるタンパク質のモル吸光係数 (ϵ) は $43,000 \text{ M}^{-1}\cdot\text{cm}^{-1}$ (280 nm) である。光路長 1 cm のセルを用いて測定したところ、吸光度 (A) は 0.86 であった。このとき、溶液中のタンパク質のモル濃度 (mol/L) を求めなさい。
- (7) 0.010 mol/L の水酸化ナトリウム (NaOH) 水溶液の pH を求めなさい。ただし、NaOH は完全に電離するとし、水の自己電離は無視できるものとする。

B. 以下の説明の空欄 (a) ~ (j) に当てはまる最も適切な語句を答えなさい。

RNA は DNA とは異なり、通常は (a) 鎖で存在し、部分的に (b) を形成することで、さまざまな二次構造や三次構造をとることができる。RNA に含まれる糖は (c) であり、これは DNA に含まれる糖である (d) とは異なる。また、RNA では塩基のひとつとしてウラシルが使われており、これは DNA の (e) に相当する。

tRNA は翻訳過程において (f) をリボソームに運搬する役割を持ち、約 70~90 ヌクレオチドから構成される。tRNA の二次構造は「クローバーリーフ構造」と呼ばれ、三次元構造では (g) 型に折りたたまれている。tRNA のアンチコドンは mRNA の (h) と塩基対を形成し、正確な翻訳を可能にしている。一方、RNA の中には酵素のような触媒活性をもつものもあり、これらは (i) と呼ばれる。特にリボソーム内の rRNA が、(j) 反応を担っていることは、RNA が触媒としても重要な役割を果たしていることを示している。

(問題 2 : 次ページに続く)

C. 以下の説明の空欄（ a ）～（ j ）に当てはまる最も適切な語句を答えなさい。

タンパク質の機能はその立体構造、特に高次構造に強く依存しており、構造は一次構造に内在する情報によって自発的に決定される。二次構造には（ a ）や（ b ）があり、これらは主に（ c ）結合によって安定化される。三次構造の形成には、水素結合、疎水性相互作用、イオン結合に加えて、（ d ）結合のような共有結合も関与している。

タンパク質の構造には、機能的に独立して折りたたまれた領域である（ e ）が存在することが多く、これらは特定の機能や結合活性に関与することが多い。細胞内では、タンパク質が誤って折りたたまれるのを防ぐために、（ f ）と呼ばれる分子群が正しい折りたたみを補助している。一部の（ f ）はATP 依存的に構造変化を誘導することで、誤った中間体の（ g ）を防いでいる。アミロイドは異常に折りたたまれたタンパク質の一種で、アルツハイマー病や（ h ）などの疾患の原因となることが知られている。

タンパク質の立体構造は、X線結晶構造解析や（ i ）などの物理的手法によって決定される。近年では、深層学習を用いた構造予測手法である（ j ）の登場により、アミノ酸配列からの三次構造予測が飛躍的に向上した。

D. 以下の説明の空欄（ a ）～（ h ）に当てはまる最も適切な語句または数値を答えなさい。

酵素反応の速度論を記述するモデルの一つに、ミカエリス・メンテン式がある。このモデルでは、酵素（E）と基質（S）が結合して酵素-基質複合体（ES）を形成し、基質（S）は生成物（P）へと変換される。

反応式は以下のように表される：



このモデルにおける反応初速度（v）は基質濃度 [S] に依存し、次の式で表される：

$$v = (a) \times [S] / ((b) + [S])$$

ここで、V_{max} は酵素が完全に基質と結合したときに達する最大反応速度であり、（ c ）のときに観察される。また、K_m は（ d ）定数と呼ばれ、酵素と基質との親和性を示す指標である。

ある酵素について、V_{max} = 100 μmol/min、K_m = 5.0 mM とする。基質濃度 [S] = 8.0 mM のとき、v = （ e ） μmol/min である。また、基質濃度 [S] = 1.0 mM のとき、v = （ f ） μmol/min である。このように、[S] が小さいほど反応速度は（ g ）し、K_m の値が小さいほど、酵素と基質の親和性が（ h ）ことを意味する。

（問題 2：終わり）

問題 3

A. 以下の文章を読み、(1)–(8) の問いに答えなさい。

大腸菌の DNA を複製する DNA ポリメラーゼは 10^{10} 塩基に 1 塩基の誤りを起こすと言われており、この正確性は 10^4 塩基に 1 塩基の誤りを起こすと言われていた RNA ポリメラーゼと比べると驚くほど高い。⁽¹⁾ しかし細胞内の DNA には非常に多くの損傷が自然に生じている。例えば、水溶液中で生じる加水分解により、シトシン塩基の (a) 基が除去され、ウラシル塩基に変化する。あるいは DNA への UV の照射により、DNA 上の隣り合う 2 個のシトシンまたはチミン塩基の間に共有結合が生じ、(b)・ダイマーと総称される構造が形成される。さらに、細胞内の代謝により生じる (c) は、塩基だけでなく DNA の糖-リン酸骨格の損傷をもたらす。これらに加えて染色体中に存在する トランスポゾン⁽²⁾ もその活性により変異を誘発することがある。細胞はこのように自然に生じる DNA の損傷を感知して修復する機構を備えている。その結果、細菌では 1 世代につき 10^{10} 塩基あたり 3 塩基の変異が生じる⁽³⁾ と言われている。

変異の解析は遺伝子機能の解明にも重要な役割を果たしてきた。私たちはトリプトファン代謝に関わる複数の遺伝子で構成された オペロン⁽⁴⁾ に属し、トリプトファンの生合成に必須な酵素 A に注目した。大腸菌株 X は、この酵素 A のアミノ酸コード領域に酵素活性を弱める点変異が生じており、トリプトファンを含まない培地では生育できない。そこで大腸菌株 X から自然に生じる変異を利用し、トリプトファンを含まない培地で生育できる変異体を多数分離した。一部の變異体では大腸菌株 X が持つ酵素 A の元のアミノ酸変異部位に再び点変異が生じていた。次に、元のアミノ酸変異部位のほか単一の点変異が生じた変異体のみを抽出して解析することにした。すると、元の變異部位とは異なるアミノ酸部位に生じた酵素 A の 2 つ目の点變異⁽⁵⁾ に加え、酵素 A と同一オペロンに属するタンパク質 B のアミノ酸を変化させる点変異や、酵素 A が属するオペロンとは別の遺伝子座から転写されるタンパク質 C のアミノ酸を変化させる点變異⁽⁶⁾ が検出された。さらには、酵素 A が属するオペロンの近傍でアミノ酸コード領域の外に生じた点變異も検出された。⁽⁷⁾

- (1) 文中の (a)、(b)、(c) に入る言葉を、a: XXX、b: YYY のように答えなさい。
- (2) 下線 1 について、DNA ポリメラーゼと RNA ポリメラーゼのどのような分子機能の違いが、これらの酵素の正確性の違いを生んでいるか 100 字程度で説明しなさい。
- (3) 下線 2 について、トランスポゾンとは何か、50 字程度で簡潔に説明しなさい。
- (4) 下線 3 について、15 時間培養した大腸菌コロニーには、理論上、何種類の変異が大腸菌ゲノムに生じることになるか。途中の式と共に答えなさい。ただし、大腸菌ゲノムは 460 万塩基対であり、大腸菌は 30 分ごとに分裂したとする。また、生じた突然変異は大腸菌の生育に中立であり、全く同じ変異は 2 度とは生じないと仮定する。 $2^{10} = 10^3$ と近似する。

(問題 3 : 次ページに続く)

- (5) 下線4について、オペロンとは何か、50字程度で簡潔に説明しなさい。
- (6) 下線5について、元の変異と2つ目の変異はどのような関係にあることが想定されるか、考えられることを50字程度で説明しなさい。
- (7) 下線6について、酵素Aとタンパク質Cにはどのような関係があることが想定されるか、考えられることを50字程度で説明しなさい。
- (8) 下線7について、どのような変異が想定されるか、考えられることを50字程度で説明しなさい。

B. 以下の文章を読み、(1)–(6)の問いに答えなさい。

ヒトにおいて親から子に遺伝情報が受け継がれる時、(a) と呼ばれる新規変異⁽¹⁾が100程度は起こる。また、親世代が持つ遺伝的多様性は減数分裂の際に(b)の影響でシャッフルされながら子世代に受け継がれる。遺伝的伝達の際に2つの遺伝的座位の間に(b)が起きずに子世代に遺伝的情報が伝達される、という現象を(c)という。(b)が起きるか起きないかは、2つの遺伝的座位の間の距離に依存する⁽²⁾と知られている。集団において、ある2つの遺伝的座位の染色体上の距離が近い結果、歴史的に(b)があまり起きておらず、結果としてそれらの遺伝的座位のバリエーションの組み合わせ頻度が集団内で非ランダムとなっていることがある。これを連鎖不平衡という。

実際に、同じ染色体上の近傍に位置する5つの2アレル座位の遺伝型を測定し、アレル頻度(表1)とハプロタイプ頻度(表2)を表にまとめた。ここでハプロタイプとは、同じ染色体上のアレルの並びを意味する。

表1 5座位のアレル頻度

| 座位 | A | B | C | D | E |
|--------|-------|-------|-------|-------|-------|
| アレル1 | T | T | C | C | A |
| アレル2 | A | G | T | T | T |
| アレル2頻度 | 0.470 | 0.479 | 0.326 | 0.326 | 0.329 |

表2 5座位からなるハプロタイプ頻度

| A-B-C-D-E 座位のハプロタイプ | 頻度 |
|---------------------|-------|
| T-T-C-C-A | 0.515 |
| A-G-T-T-T | 0.308 |
| A-G-C-C-A | 0.158 |
| A-G-C-C-T | 0.011 |
| T-G-T-T-A | 0.009 |

- (1) 文中の (a)、(b)、(c) に入る言葉を以下の語群から選び、a: XXX、b: YYY、c: ZZZ のように答えなさい。

語群：刷り込み効果、水平伝播、デノボ突然変異、機能喪失変異、遺伝的転写、遺伝的組換え、遺伝的浮動、遺伝的距離、遺伝的相関、遺伝的連鎖

- (2) 下線 1 について、このような新規変異は父親由来が母親由来よりも多いことが知られているが、そのような変異発生の父母差が生じるタイミングとして、最もふさわしいものを以下の選択肢から 1 つ選び記号で答えなさい。

- A. 精子、または卵子の形成時
B. 受精時
C. 胚発生初期
D. 胚細胞の分化期

- (3) 下線 2 について、一般に 2 つの遺伝的座位が別々の染色体にある場合は、バリエントアレルの組み合わせはランダムに子に伝達される。この法則をなんというか。

- (4) もし 5 座位間にまったく連鎖不平衡がなく、集団構造化などその他のアレル頻度の組み合わせに与える影響もないとした場合、表 1 を元にとすると、ハプロタイプ A-G-T-T の頻度はどれくらいであると推定されるか。計算式だけでも良い。

- (5) 表 2 のハプロタイプ構成とその頻度を元にとすると、この 5 座位の間には比較的組換えが起りやすい場所があると考えられる。どの座位とどの座位の間にそれがあると考えられるか。また、ゲノム上で組換えがとりわけ起りやすい場所のことをなんと呼ぶか。

- (6) Y 染色体上の二座位の間の連鎖不平衡は一般にどうであると考えられるか。25 字程度で理由とともに述べなさい。

(問題 3 : 終わり)

問題 4

A. 以下の文章を読み、(1)–(7) の問いに答えなさい。

大学院生のエムさんは、遺伝子 A への突然変異によりタンパク質 A の機能が低下したヒト培養細胞株 Y に遺伝子 A を導入することで、細胞機能を回復させる実験に取り組んでいます。まず、正常なヒト細胞株 Z から mRNA を抽出して、開始コドンから終止コドンまで 0.9 kb の遺伝子 A の cDNA (1) を合成し、これをプラスミド DNA にクローニングして、遺伝子 A の発現ベクターを作成しました。作成したプラスミド DNA を細胞に導入すると細胞内で遺伝子 A が発現してタンパク質 A が合成 (2) されます。

遺伝子導入した細胞内で実際にタンパク質 A が発現しているか確認するため、エムさんはウェスタンブロッティング法 (3) で検証しました。作成したプラスミド DNA を細胞株 Y に導入し、2 日間培養した後、細胞溶解液を作成しました。対照実験として、細胞株 Y および Z に加えて、遺伝子 A を含まないプラスミド DNA を導入した細胞株 Y からも細胞溶解液を作成しました。4 種類の細胞溶解液に SDS と還元剤を含む緩衝液 (4) を加えて加熱し、それらを SDS ポリアクリルアミドゲル電気泳動 (SDS PAGE) により分離しました。電気泳動後、分離したタンパク質をゲルからニトロセルロース膜に転写し、膜上のタンパク質 A を特異的に検出 (5) しました。(図 1) 興味深いことに、細胞株 Y では細胞株 Z とは異なり、タンパク質 A のバンドが 2 本 (6) 検出されました。

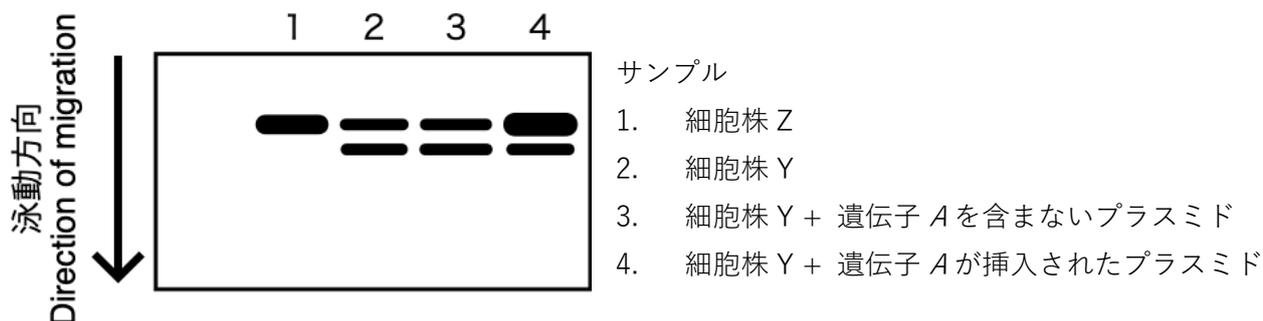


図 1. ウェスタンブロッティングの検出結果

- (1) 下線部 1 について、遺伝子 A の cDNA と染色体 DNA 上の遺伝子 A との違いを 60 字程度 で簡潔に説明しなさい。
- (2) 下線部 1 について、タンパク質 A の大きさとして最も適切なものを以下の選択肢から選び記号で答えなさい。
- | | | |
|------------|-----------|------------|
| A. 1 kDa | C. 10 kDa | E. 100 kDa |
| B. 3.3 kDa | D. 33 kDa | F. 330 kDa |
- (3) 下線部 2 について、細胞内でプラスミド DNA からタンパク質 A が合成されるまでの過程を、各反応が細胞内のどの区画で起こるかを含めて 80 字程度 で簡潔に説明しなさい。

B. 以下の文章を読み、(1)–(7) の問いに答えなさい。

シングルセル解析技術 ⁽¹⁾ は個々の細胞における分子状態を解析する上で用いられる。博士研究員の S さんは、治療抵抗性の一因である細胞の多様性の実態を明らかにしようと、ヒトがん組織検体を用いて シングルセル RNA-seq 解析 ⁽²⁾ を実施した。まず、細胞を分離するために、結合組織の主成分を分解する (a) 等により組織を酵素処理した。次に、各細胞の mRNA を、(b) と細胞および RNA 分子ごとに固有のバーコードが含まれたオリゴヌクレオチドで捕捉し、(c) 反応によって、バーコードが付加された cDNA を合成した。複数検体から得られたシングルセル RNA-seq データを統合して解析したところ、バッチ効果 ⁽³⁾ が見られたため、バッチ効果を考慮して以後の解析を実施した。解析結果から浸潤に関わる遺伝子群の発現が高いがん細胞の細胞集団クラスター X を同定した。また、クラスター X では、他のがん細胞クラスターに比べて、遺伝子 A の発現が上昇しており、遺伝子 B を含むいくつかの遺伝子の発現が低下していた。S さんが文献を調べたところ、遺伝子 A は ヒストン修飾 ⁽⁴⁾ に関わるタンパク質をコードしており、遺伝子 B は細胞周期の制御因子として報告されていた。

がんの治療抵抗性には、がん細胞だけでなく、周囲の間質細胞の状態も深くかかわっている。S さんは、他の細胞種の遺伝子発現状態にも着目した。その結果、クラスター X のがん細胞は、ある線維芽細胞集団と相互作用していることが推定された ⁽⁵⁾。

(1) 下線 1 について、最も 不適切な記述 を 1 つ選べ。

- A. 各 1 細胞のゲノム変異や遺伝子発現状態を計測することができ、Nature Methods 誌の”Method of the Year”に選出されたことがある。
- B. シングルセルゲノム解析では、ヒト細胞 1 個あたり 30 億コピーの核ゲノム DNA を解析する必要があり、膨大なシーケンス量が必要である。
- C. シングルセルトランスクリプトーム解析では、マイクロフリューディクスやマイクロドロップレット、マイクロウェルなどを活用する手法が開発されており、プラットフォームによっては、数万以上の細胞を 1 細胞に単離して解析することができる。
- D. シングルセルエピゲノム解析手法の一つとして、Tn5 トランスポザラーゼを用いて各 1 細胞のオープンクロマチン状態を計測するシングルセル ATAC-seq がある。

(2) 下線 2 に関連して、文中の (a) – (c) に入る最も適切な言葉を以下の語群から選び、a: XXX、b: YYY のように答えなさい。

語群：ヘリカーゼ、セルラーゼ、コラゲナーゼ、ポリ A、ポリ C、ポリ G、ポリ T、PCR、逆転写、転写、翻訳、アミノ酸、rRNA、cDNA、gDNA

(3) 下線 3 について、以下の語群の用語を全て用いて 100 字程度 で説明せよ。

語群：生物学的、実験条件

- (4) 下線4について、以下の文中の (a) - (h) に入る言葉を以下の語群から選び、a: XXX、b: YYY のように答えなさい。

ヒストンは、(a) 電荷のタンパク質であり、(b) 電荷の (c) と相互作用する。クロマチンの基本単位である (d) は、ヒストン (e) に (c) が巻き付いた複合体であり、(f) 種類のコアヒストンからなる。ヒストンの末端部分は (g) と呼ばれ、メチル化やアセチル化などさまざまな修飾を受ける。これらヒストン修飾は、クロマチン構造を制御し、転写の活性化、抑制に寄与している。例えば、ヒストン H3 における9番目のリジンのトリメチル化 (H3K9me3) は、(h) を形成し、転写を抑制する。

語群：正、負、DNA、RNA、アクチンフィラメント、細胞、ミトコンドリア、リソソーム、ヌクレオソーム、2量体、4量体、6量体、8量体、2、4、6、8、10、ヒストンキャップ、ヒストンテール、ユークロマチン、ヘテロクロマチン、オープンクロマチン

- (5) 下線4のパターンを計測する手法として、ChIP-seq (クロマチン免疫沈降シーケンス) 法がある。以下の語群の用語を全て用いて ChIP-seq の原理を 100字程度 で説明せよ。

語群：免疫沈降、架橋 (クロスリンク)、DNA、タンパク質

- (6) Sさんがさらに先行研究を調べると、一部のがん種では、遺伝子Aの過剰発現や機能獲得型変異が見出されており、遺伝子Aにおけるこれらの異常は、遺伝子Bの転写制御領域におけるヒストン修飾異常、および、遺伝子Bの遺伝子発現異常に寄与していることが報告されていた。このことを受けて、遺伝子Aがどのような機能を持ち、クラスターXでどのように作用して遺伝子Bの転写抑制を引き起こしているかと推察できるか、「ヒストン修飾酵素」という用語を用いて、100字程度 説明せよ。

- (7) 下線5について、シングルセル RNA-seq のデータから細胞間相互作用を推定する方法を、「リガンド」という用語を用いて、100字程度 で説明せよ。

(問題4：終わり)

問題 5

A. 以下の (1) ~ (10) の問いに答えなさい。

(1) 以下の A~D の細胞を、多能性が高い順に並べなさい。ただし、同程度の多能性を持つ場合は「=」で表記すること。(例：A=B>C>D)

- A. 造血幹細胞
B. ES 細胞
C. 受精直後の受精卵
D. 神経幹細胞

(2) 哺乳類胚盤胞で、内部細胞塊から直接分化するものを全て選びなさい。

- A. 外胚葉 B. 中胚葉 C. 内胚葉 D. 栄養外胚葉

(3) 脊椎動物初期胚で、外胚葉が神経板へと運命転換（神経誘導）するきっかけとして最も重要なものを1つ選びなさい。

- A. 高濃度 BMP シグナルを受け続けること
B. 背側中胚葉からの BMP シグナルの阻害
C. 卵黄中の栄養分が枯渇すること
D. 精子由来ミトコンドリアの分解

(4) 造血幹細胞移植後の免疫再構築に関する記述で正しいものを1つ選びなさい。

- A. NK 細胞など自然免疫の回復は遅い。
B. 腸内フローラの多様性が増し、免疫調整機能が改善する。
C. 胸腺機能が低下しているため、ナイーブ T 細胞の再構築が遅延する。
D. 移植前の記憶を持つメモリー T 細胞が最も早く回復する。

(5) 第一減数分裂において染色体の非分離が起こった場合、生じ得る配偶子の組み合わせとして正しいものを次の選択肢から選びなさい。ここで、 n は正常な染色体数を表します。

- A. $n, n, n+1, n-1$ C. $n+1, n+1, n, n-2$
B. $n+1, n-1, n+1, n-1$ D. $n+2, n, n, n-2$

(6) 主に内因性抗原を細胞傷害性 T 細胞に提示する主要分子として最も適切なものを1つ選びなさい。

- A. 主要組織適合性複合体クラス I (MHC class I)
B. 主要組織適合性複合体クラス II (MHC class II)
C. CD4
D. CD8
E. B 細胞受容体 (BCR)
F. T 細胞受容体 (TCR)

- (7) マクロファージの主な機能として正しいものをすべて選びなさい。
- A. パーフォリン/グランザイムによる細胞傷害
 - B. オプソニン化された標的の貪食
 - C. IgE による脱顆粒とヒスタミン放出
 - D. 抗原提示
- (8) 免疫チェックポイント機能の阻害を目的に使用される抗 PD-1 抗体治療について、直接的な効果として期待されるものを1つ選びなさい。
- A. 樹状細胞の抗原提示能の抑制
 - B. T 細胞疲弊の解除と機能の回復
 - C. B 細胞のクラススイッチの抑制
 - D. NK 細胞による抗体依存性細胞傷害 (ADCC) の阻害
- (9) 同種造血幹細胞移植後に末梢性の制御性 T 細胞 (Treg) が著減した場合、最もリスクが高い合併症として考えられるものを1つ選びなさい。
- A. 生着不全
 - B. 微小残存病変 (MRD) の増悪
 - C. 急性 GVHD (移植片対宿主病)
 - D. 真菌感染症
- (10) 季節性インフルエンザに対して、毎年ワクチン接種が推奨される理由を 20 字程度 で答えなさい。

B. 次の文章を読み、(1) ~ (6) の問いに答えなさい。

山中君と高橋君は、「ES 細胞で高発現している遺伝子を導入すれば、分化した体細胞を ES 細胞様の性質を持つ幹細胞にリプログラミングできるのではないか」という仮説を立て、以下の実験を行いました。まず、マウス胎児線維芽細胞 (MEF) の Fbx15 遺伝子 ⁽¹⁾ の下流にネオマイシン耐性遺伝子をノックインしました。この細胞は、*Fbx15* が高発現した時のみネオマイシンに耐性を示します。次に、ES 細胞で高発現している遺伝子を 24 個選んで、それらを一つずつ MEF に導入し、ネオマイシンを添加した ES 細胞用の培地で培養してみました。しかし、ネオマイシン耐性細胞を得ることはできませんでした。そこで、思い切って 24 遺伝子を全部 MEF に導入してみました。すると今度は、ネオマイシン耐性を示すコロニーをいくつか得ることができました。それらのコロニーは丸みを帯びた形をしており、ES 細胞とそっくりでした。遺伝子導入後の細胞から RNA を数日おきに採取して *Egf4* 遺伝子と *Coll1a1* 遺伝子の発現量の変化を qPCR で調べたところ、以下の結果が得られました (表 1) ⁽²⁾。また、*Fbx15* 遺伝子のプロモーター領域を調べたところ、(a) していました ⁽³⁾。そして、ES 細胞の多能性を評価する実験 ⁽⁴⁾ を行い、24 遺伝子の導入により MEF を ES 細胞様にリプログラミングできることを確認しました。

(問題 5 : 次ページに続く)

さらに、山中君と高橋君はいくつかの実験を追加して (5)、24 遺伝子のうち *Oct4*、*Klf4*、*Sox2*、*Myc* の4つが重要であることを突き止めました。このうち (b) と (c) は、ES 細胞の幹細胞性維持に必須の役割を果たす遺伝子です。*Myc* はがん遺伝子としても知られており、ヒストンアセチル化酵素との相互作用を介して遺伝子発現を調節しています。(d) は *TP53* の発現を抑制することが知られており、この機能が細胞のリプログラミングに重要なのかもしれません。

表 1: *Fgf4* および *Coll1a1* の発現量の変化。遺伝子導入前の発現量を 1.0 とする。

| | day 0 | day 7 | day 14 | day 21 |
|----------------|-------|-------|--------|--------|
| <i>Fgf4</i> | 1.0 | 2.7 | 6.5 | 9.8 |
| <i>Coll1a1</i> | 1.0 | 0.8 | 0.3 | 0.1 |

- (1) 下線 1 について、*Fbx15* の代わりになると考えられる遺伝子はどれか。以下から 1 つ選びなさい。
- A. *Dnmt3a* B. *Bcl2* C. *Kras* D. *Nanog*
- (2) 下線 2 について、*Fgf4* と *Coll1a1* はどのような遺伝子と考えられるか。発現量の変化を踏まえて合わせて 50 字程度で説明しなさい。
- (3) 下線 3 について、*Fbx15* 遺伝子のプロモーター領域を調べたところ、以下の変化が確認されました。(a) に当てはまる文章として最も適切なものを選びなさい。
- A. DNA のメチル化が低下 C. FBX15 結合領域が欠失
 B. ヒストンのアセチル化が低下 D. ES 細胞に特異的な配列が存在
- (4) 下線 4 について、ES 細胞の多能性評価には使われない方法はどれか。以下から 1 つ選びなさい。
- A. 三胚葉への分化マーカー発現解析 C. キメラマウス作製
 B. 移植後のテラトーマ形成試験 D. 細胞質の ATP 濃度測定
- (5) 下線 5 について、どのような実験を行えば 24 遺伝子の中から重要な 4 遺伝子を突き止めることができるか。実験計画を考えて、200 字程度で説明しなさい。
- (6) (b) (c) (d) に当てはまる遺伝子名を答えなさい。

C. 免疫グロブリン分子の多様性について、仮定 i ~ iv のもとで (1) ~ (4) の問いに答えなさい。また、T 細胞受容体の多様性に関する (5)、(6) の問いに答えなさい。

[仮定]

- i. IgH (重鎖) 遺伝子の V、D、J セグメント数はそれぞれ $V=40$ 、 $D=23$ 、 $J=6$ とする。
- ii. 軽鎖 κ 遺伝子のセグメント数は $V_{\kappa}=34$ 、 $J_{\kappa}=5$ 、軽鎖 λ 遺伝子のセグメント数は $V_{\lambda}=30$ 、 $J_{\lambda}=4$ とする。
- iii. 計算は組合せ多様性のみとし、接合多様性、体細胞超変異、非生産的再構成、対立遺伝子排除の失敗、等は考慮しない。
- iv. 軽鎖は κ と λ のいずれか一方が発現する (相互排他的発現) とし、重鎖と軽鎖のペアリングに制約はないものとする。

- (1) IgH の遺伝子再構成による組合せの総数を求めなさい。
- (2) 軽鎖 κ の遺伝子再構成による組合せの総数を求めなさい。
- (3) 軽鎖 λ の遺伝子再構成による組合せの総数を求めなさい。
- (4) 重鎖と軽鎖によって生じる抗体の組合せの総数を求めなさい。
- (5) 免疫グロブリン分子と同様に、T 細胞受容体 (TCR) も極めて多様である。未知の病原体に対し、この初期多様性がもたらす利点を 100 字程度 で説明しなさい。
- (6) T 細胞の多様性を損なわず自己に対する免疫応答を制御する仕組みについて、T 細胞の成熟過程と末梢におけるメカニズムをそれぞれ記載し 合わせて 100 字程度 でまとめなさい。

(問題 5 : 終わり)

問題 6

A. 以下の文章を読み、(1)–(10) の問いに答えなさい。

- (1) がんは、発生母地となる組織・細胞によって分類され、最も多くを占めるのは上皮組織から発生する Carcinoma である。Carcinoma 以外の悪性腫瘍の分類と発生母地となる組織名称の組み合わせを1つ挙げなさい。
- (2) アポトーシスと異なり、プログラムされない細胞死として定義され、細胞内容物が細胞外に漏出し顕著な炎症反応を引き起こす細胞死の形態の名称を答えなさい。
- (3) DNA をバイサルファイトで処理すると、メチル化されていないシトシンがウラシルに変換される。バイサルファイト処理後の DNA をもちいてシークエンス解析したところ、5'-TCTCG-3' という配列に対して処理後は 5'-TTTCG-3' となっていた。メチル化は 5'側から何番目の塩基でおきていたと考えられるか答えなさい。
- (4) がんで高頻度にゲノム異常を認めるシグナル伝達分子と、その機能に関する説明として、誤っている組合せを2つ選びなさい。
- | | |
|-----------------------------|-------------------------|
| A. ABL - 非受容体型チロシンキナーゼ | D. KRAS - GTP 結合タンパク質 |
| B. HER2 - 受容体型チロシンキナーゼ | E. CBL - ユビキチンリガーゼ |
| C. BRAF - 受容体型セリン/スレオニンキナーゼ | F. PI3K - セリン/スレオニンキナーゼ |
| | G. PTEN - ホスファターゼ |
- (5) 感染症により発がんのリスクが上がる事が知られている。そのような例について、正しい組み合わせを2つ選びなさい。
- | | |
|----------------------|------------------------|
| A. EB ウイルス - 肺がん | C. ヒトパピローマウイルス - 子宮頸がん |
| B. C 型肝炎ウイルス - 肝細胞がん | D. ヘリコバクター・ピロリ菌 - 大腸がん |
- (6) 欧米系集団に比べてアジア系集団において頻度が高いがんを2つ選びなさい。
- | | | |
|----------|----------|---------|
| A. 肝細胞がん | C. 前立腺がん | E. 大腸がん |
| B. 食道がん | D. 乳がん | |
- (7) 主に遺伝子増幅により活性化される原がん遺伝子を1つ選びなさい。
- | | | | |
|---------|--------|--------|--------|
| A. KRAS | B. MYC | C. CBL | D. ALK |
|---------|--------|--------|--------|
- (8) がんの微小環境の特徴の1つが、がん組織内に侵入してきた血管である。低酸素の状態において活性化され、血管新生の促進にはたらく転写因子を答えなさい。

(9) 消化管などに生じたがんは深部に浸潤すると腹腔内に広がることもある。これを何と呼ぶか、下記から最も適切なものを1つ選びなさい。

- | | |
|-----------|---------|
| A. 播種 | C. 形質転換 |
| B. 上皮間葉転換 | D. 脱分化 |

(10) がん化を促進する代謝物をオンコメタボライトという。脳腫瘍や軟骨肉腫で高頻度に認められる変異型 IDH1 によって生成されるオンコメタボライトの名称を答えなさい。

- | | |
|--|------------------------------|
| A. α -Ketoglutarate (α -KG) | C. Acetaldehyde |
| B. Retinoic acid | D. 2-Hydroxyglutarate (2-HG) |

B. (1)–(5) の問いに答えなさい。

(1) 大腸がんのゲノム上の変異総数は、肺がんや悪性黒色腫と比較しておよそ半分程度であるが、あるタイプの遺伝性大腸がんでは例外的に、肺がんや悪性黒色腫よりも数倍程度多くの変異が認められる。このようなタイプのがんではどのような機能を持った遺伝子の働きが損なわれていると考えられるか、50 字程度で説明しなさい。

(2) 大部分のヒトがん細胞ではテロメラーゼ活性が検出される。この酵素が担う反応と、がん細胞に対する影響について、合わせて 50 字程度で説明しなさい。

(3) がん遺伝子パネル検査では、がんの組織を用いて特定の遺伝子セットを対象にシーケンスが実施され、同定された変異に対して効果を示す薬を選択することで個別化医療を実現する。一方で、特定の遺伝子だけではなく、全ゲノムを調べる取り組みもいくつかの国で実施されている。全ゲノムを調べることのメリットとデメリットについて合わせて 50 字程度で答えなさい。

(4) ある女性の乳がん患者で、がん組織および正常組織のペアでゲノム検査を行った。その結果、*BRCA1* 遺伝子において、正常組織では一方のアレルに変異を認め、腫瘍組織では両アレルに変異を認めた。この患者の娘が *BRCA1* 遺伝子変異を持つ確率を答えなさい。また患者の娘に対しどのような検査や治療を勧めるべきか、合わせて 150 字程度で答えなさい。

(5) 家族性腫瘍の原因遺伝子の多くはがん遺伝子ではなく、がん抑制遺伝子である。その理由について考えられることを 200 字程度で答えなさい。

(問題 6 : 終わり)

問題 7

以下の問いに答えよ。

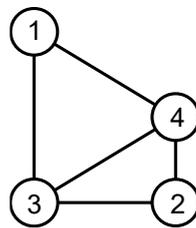
- (1) 行列 $A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$ の固有値を求めよ。
- (2) 行列 A の固有空間の基底を求めよ。
- (3) $P^T A P$ が対角行列になるような直交行列 P を求めよ。
- (4) $x = (x_1, x_2, x_3)^T$ を三次元実ベクトルとする。 $f(x) = x^T A x$ の最小値を与える最適解の集合は、三次元空間における直線になっている。この直線を表す方程式を示せ。
- (5) 実数の成分からなる対称正方行列に対して、すべての固有値が異なるならば、すべての固有ベクトルは相互に直交することを証明せよ。

問題 8

$G = (V, E)$ を点集合 V 、辺集合 E の単純連結無向グラフとする。 V の要素数を n とする。 V を 2 分割したものを $C = (S, T)$ を G のカットと呼ぶ。

ここで S と T は $S \subset V, T \subset V, S \neq \emptyset, T \neq \emptyset, S \cap T = \emptyset, S \cup T = V$ を満たし、 S と T を入れ替えたペア (S, T) 、 (T, S) は区別しないものとする。カット C のカットサイズを S と T をまたぐ辺の個数として定義する。また、カットサイズ s をもつカットの個数を s の重複度 $m(s)$ と定義する。以下の問に導出も含めて答えよ。

- (1) 以下のグラフ G_0 について、全ての可能なカットサイズ s とその重複度 $m(s)$ を求めよ。



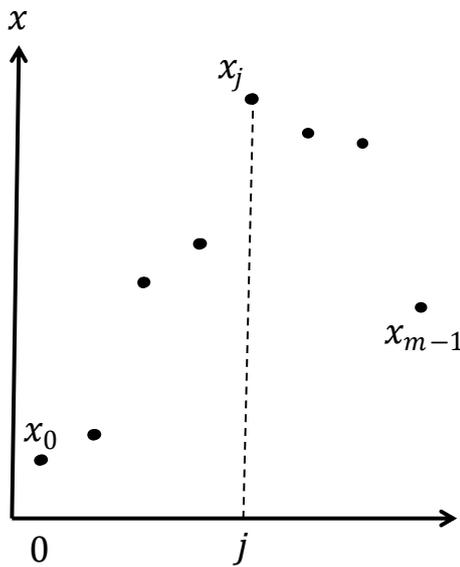
- (2) G が完全グラフ、すなわち $V = \{1, \dots, n\}$, $E = \{(i, j) | i, j = 1, \dots, n, i < j\}$ のとき、可能なカットサイズ s とその重複度 $m(s)$ を全て求めよ。
- (3) G が環状グラフ、すなわち $V = \{1, \dots, n\}$, $E = \{(i, i + 1) | i = 1, \dots, n - 1\} \cup \{(n, 1)\}$ のとき、カットサイズの最小値 s_{min} とその重複度 $m(s_{min})$ を求めよ。
- (4) G の最小カットサイズ s_{min} をもつカットの一つを $C_{min} = (S_{min}, T_{min})$ とする。このとき E の中で、 S_{min} と T_{min} をまたぐ辺の割合は $2/n$ 以下であることを示せ。必要であれば、性質 $|E| = \sum_{v \in V} \deg(v) / 2$ を用いてもよい。ただし、 $|E|$ は E の要素数、 $\deg(v)$ は点 v に接続する辺の数を表す。

問題 9

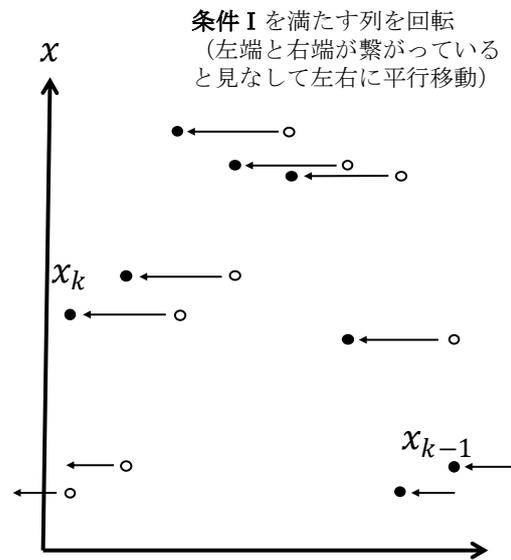
相異なる 2^n 個の要素を持つ整数配列 $A = \{a_0, a_1, \dots, a_{2^n-1}\}$ を整列したい。また、配列 $X = \{x_0, x_1, \dots, x_{m-1}\}$ は以下の条件の何れかを満たすときバイトニックであると定義する。

条件 I: ある j ($0 \leq j \leq m-1$) が存在して $0 \leq i < j$ を満たす全ての i について $x_i \leq x_{i+1}$ であり、かつ $j < i \leq m-1$ を満たす全ての i について $x_{i-1} \geq x_i$ である。言い換えると、 x_0, x_1, \dots, x_j と $x_j, x_{j+1}, \dots, x_{m-1}$ はそれぞれ広義単調増加列と広義単調減少列になっている。

条件 II: 配列 X を回転した配列 X' であって**条件 I** を満たすものが存在する。すなわち、ある k ($0 < k \leq m-1$) について配列 $X' = \{x_k, x_{k+1}, \dots, x_{m-2}, x_{m-1}, x_0, x_1, \dots, x_{k-1}\}$ が**条件 I** を満たす。



条件 I



条件 II

(1) 以下の配列がそれぞれバイトニックであるかどうかを答えよ。

- i. $\{1\}$
- ii. $\{1, 2, 3, 4\}$
- iii. $\{1, 3, 2, 4\}$
- iv. $\{3, 1, 2, 5, 6, 8, 7, 4\}$

(2) 長さが 2 の配列 $X = \{x_0, x_1\}$ は常にバイトニックであることを示せ。 $x_0 \neq x_1$ を仮定してよい。

(3) 相異なる 2^p 個の整数を要素に持つバイトニックな配列 $X = \{x_0, x_1, \dots, x_{2^p-1}\}$ に対して

配列 $X_{min} = \{\min(x_0, x_{2^p-1}), \min(x_1, x_{2^p-2}), \dots, \min(x_{2^p-1-1}, x_{2^p-1})\}$ および

配列 $X_{max} = \{\max(x_0, x_{2^p-1}), \max(x_1, x_{2^p-2}), \dots, \max(x_{2^p-1-1}, x_{2^p-1})\}$ を考える。

配列 X_{min}, X_{max} はバイトニックであり、 X_{min} の要素の最大値は X_{max} の要素の最小値より小さいことを仮定せよ。バイトニック配列 X を効率的に整列するアルゴリズムを示せ。また、 p についての最悪時間計算量を示せ。

- (4) (3) において、仮定を置かなくても実際に配列 X_{min}, X_{max} はバイトニックであり、 X_{min} の要素の最大値は X_{max} の要素の最小値より小さくなることを示せ。
- (5) (3)の操作を利用して、本大問の冒頭で定義した配列 A (バイトニックであるとは限らない) を整列するアルゴリズムを示せ。また、示したアルゴリズムを並列度無限の理想的な並列計算機、すなわち依存関係が無い独立した計算操作は全て単位時間で計算を終えることができる計算機で実行した場合の最悪時間計算量を n に関して示せ。

ヒント：昇順に整列された2つの配列 Y, Z が与えられたとき、 Z を反転して Y と連結することでバイトニックな配列を得ることができる。

(問題9：おわり)

問題 1 0

$d + 1$ 次元のベクトル $\mathbf{x} = (1, x_1, x_2, \dots, x_d)^\top$ に対し、母数ベクトル $\mathbf{w} = (w_0, w_1, w_2, \dots, w_d)^\top$ との内積による線形予測子 $\eta = \mathbf{w}^\top \mathbf{x} = w_0 + w_1 x_1 + \dots + w_d x_d$ を考える。

また、 n 個の観測データを $(\mathbf{x}^{(1)}, y^{(1)}), (\mathbf{x}^{(2)}, y^{(2)}), \dots, (\mathbf{x}^{(n)}, y^{(n)})$ とし、 $\eta^{(i)} = \mathbf{w}^\top \mathbf{x}^{(i)}$ とする。

$\mathbf{x}^{(i)}$ の成分は $\mathbf{x}^{(i)} = (1, x_1^{(i)}, x_2^{(i)}, \dots, x_d^{(i)})^\top$ と表記する。

このとき、次の(A)と(B)の2通りの問題を考える。

(A) 分類問題：応答変数 y が2値変数 $y \in \{0, 1\}$ で、 $y = 1$ または $y = 0$ を取る確率が

$$P(y = 1 | \eta) = \sigma(\eta)$$

$$P(y = 0 | \eta) = 1 - \sigma(\eta)$$

に従うとする。すなわち、 y は確率質量関数 $P(y|\eta) = \sigma(\eta)^y (1 - \sigma(\eta))^{1-y}$ のベルヌーイ分布に従うとする。ただし、 $\sigma(\eta)$ はシグモイド関数 $\sigma(\eta) = 1/(1 + e^{-\eta})$ とする。

(B) 回帰問題：応答変数 y が連続変数で平均 η 、分散1の正規分布に従うとする。

このとき、以下の問いに答えよ。ただし、導出の過程を省略せずに示すこと。

(1) シグモイド関数の導関数が $\sigma'(\eta) = \sigma(\eta)(1 - \sigma(\eta))$ となることを導出せよ。

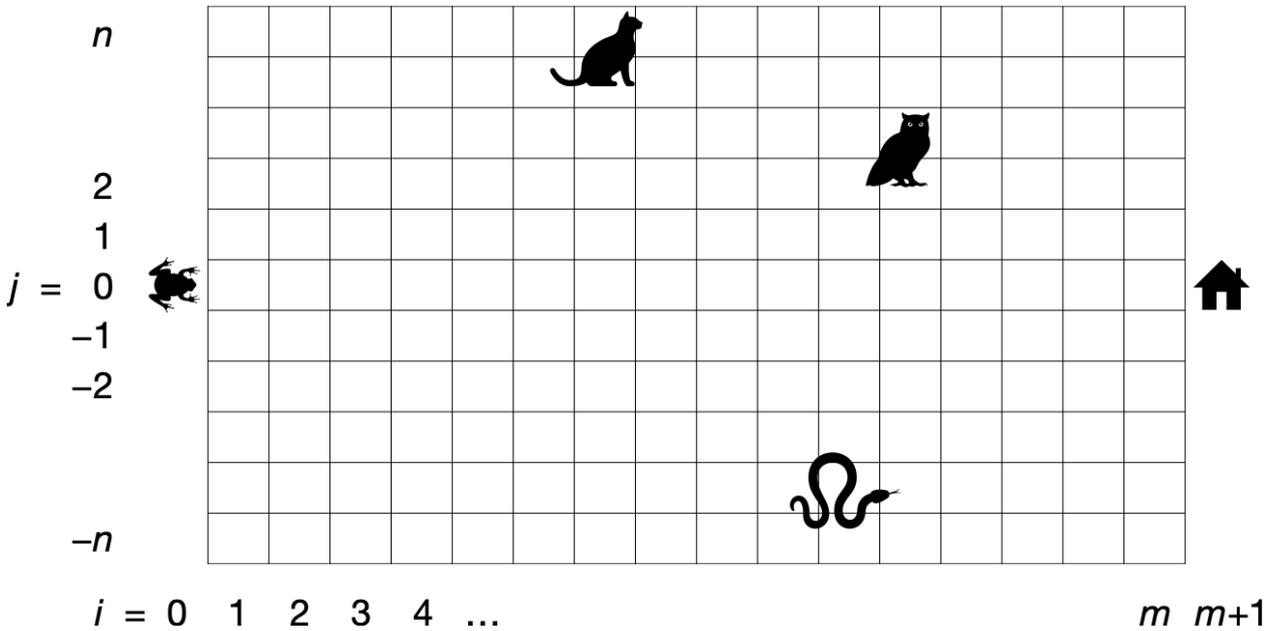
(2) (A)の問題に対する最尤推定を考え、勾配法を適用することを考える。対数尤度関数 $L(\mathbf{w}) = \log(\prod_{i=1}^n P(y^{(i)}|\eta^{(i)}))$ に対して、 $\frac{\partial L(\mathbf{w})}{\partial w_j} = \sum_{i=1}^n (y^{(i)} - \sigma(\eta^{(i)})) x_j^{(i)}$ を導出せよ。

(3) (A)の問題に対して最尤推定ではなく、単純に観測値 y とその予測値 $\hat{y} = \sigma(\eta)$ との二乗誤差を最小にするように、すなわち $L(\mathbf{w}) = -\frac{1}{2} \sum_{i=1}^n (y^{(i)} - \sigma(\eta^{(i)}))^2$ を最大にするようにパラメータを求める場合を考える。

このとき、 $\frac{\partial L(\mathbf{w})}{\partial w_j} = \sum_{i=1}^n (y^{(i)} - \sigma(\eta^{(i)})) \sigma(\eta^{(i)}) (1 - \sigma(\eta^{(i)})) x_j^{(i)}$ を導出せよ。

(4) (B)の問題に対する最尤推定を考え、勾配法を適用することを考える。対数尤度関数 $L(\mathbf{w})$ を定義し、 $\frac{\partial L(\mathbf{w})}{\partial w_j} = \sum_{i=1}^n (y^{(i)} - \eta^{(i)}) x_j^{(i)}$ を導出せよ。

問題 1 1



蛙が家に帰るには危険な野原を通らねばならない。その野原を $m \times (2n + 1)$ 個のマスタ目で表す。蛙は次の移動ができる：前（右側）に 1 マス進む、または斜め上に 1 マス進む、または斜め下に 1 マス進む。各マスでの死亡確率を $p_{i,j}$ ($1 \leq i \leq m$, $-n \leq j \leq n$) とする。蛙は死ぬと動かなくなる。蛙が出発する場所 ($i = 0, j = 0$) と家 ($i = m + 1, j = 0$) は安全な場所とする。

- (1) 蛙が以下のマスタ目を移動することを計画している： $(i = 1, j = 1), (i = 2, j = 1), (i = 3, j = 0)$ 。蛙が生き残る確率を答えよ。
- (2) 蛙が実際に上記の経路で移動し終えた。このとき蛙が生きている確率を答えよ。
- (3) $q(i, j)$ を蛙がマスタ目 (i, j) に到達する可能な経路の中での最大生存確率とする。ただし $1 \leq i \leq m$, $|j| \leq \min(n, i)$ とする。 $q(i, j)$ を計算するアルゴリズムの擬似コードを書け。計算時間は $n \times i$ に線形比例するか、それより短いものであること。
- (4) 蛙が家に到達する確率が最大になる経路を一つ出力するアルゴリズムの擬似コードを書け。計算時間は $n \times m$ に線形比例するか、それより短いものであること。

(このページは草稿用紙として使用してよい)

(Blank page for draft)

(このページは草稿用紙として使用してよい)

(Blank page for draft)

受験上の注意

試験開始の合図があるまで、この冊子を開いてはいけません。

1. 試験室においては、全て監督者の指示に従うこと。
2. 試験時間中は、受験票を常に机上におくこと。
3. 試験時間中に机においてよいものは、受験票、筆記用具（黒色鉛筆又は黒色シャープペンシル、消しゴム、必要な場合は電動ではない小型の鉛筆削り）、眼鏡（スマートグラス、ケースは不可）、ティッシュペーパー（中身のみ）だけである。これ以外の所持品を置いている場合には、不正行為となる可能性がある。
4. シャープペンシルの芯や芯ケースは机上に出してはならない。また、下敷、筆箱、電動式あるいは大型の鉛筆削り、ナイフ類、置き時計、キッチンタイマーは机に置いてはならない。
5. 試験時間中には、腕時計のほか、携帯電話、スマートフォン、スマートグラス、イヤホン、腕時計型端末、など電子機器類や電卓・分度器・コンパス等の補助具を使用してはならない。
6. 試験の際、不正行為を行った場合は、その場で受験の中止と退室が命じられ、それ以後の受験はできない。また、受験した科目全ての成績を無効とする。
 - i 次のことを行うと直ちに不正行為となる。

- | |
|---|
| <ol style="list-style-type: none">a. 受験票、解答用紙へ故意に虚偽の記入をすること（受験票の写真が本人と確認できない場合や、解答用紙に本人以外の名前・受験番号を記入するなど）b. カンニング（カンニングペーパー・参考書・他の受験者の答案等を見ること、他の人から答えを教わること）をすることc. 他の受験者に答えを教えること、カンニングの手助けをすることd. 試験時間中に、問題冊子を試験室から持ち出すことe. 解答用紙を試験室から持ち出すことf. 「解答はじめ。」の指示の前に、問題冊子を開くこと、解答を始めることg. 「解答やめ」の指示に従わず、筆記具や消しゴムを置かないこと |
|---|

- ii 上記 i 以外にも、次のことをすると不正行為となることがある。

- | |
|---|
| <ol style="list-style-type: none">a. 試験時間中に、腕時計および、携帯電話やスマートフォン、スマートグラス、イヤホン、腕時計型端末、電子辞書、ICレコーダー等の電子機器類や電卓・分度器・コンパス等の補助具をかばん等にしまわず、机等に置くこと、身に着けていること及び手に持っていることb. 試験時間中に携帯電話やスマートフォン、時計等の音（着信、アラーム、振動音など）を長時間鳴らすなど、試験の進行に影響を与えることc. 机等に何かを書きつけることd. 試験に関することについて、自身や他の受験者を利するような虚偽の申し出をすることe. 試験場において他の受験者の迷惑となる行為をすることf. 試験場において試験監督者等の指示に従わないことg. その他、試験の公平性を損なうおそれのある行為をすること |
|---|

7. コートは着たままでもよい。脱いだコートやかばん等を空いている席に置いてはならない。また、文字・記号や地図等の柄がついた衣類又は服を着てはならない。座布団、膝掛けの使用を希望する者は、監督者に申し出て許可を受けてから使用すること。
8. マスクは着用してよいが、マスクを着用している場合には、写真照合の際には、試験監督の指示に従ってマスクを一時的に外すこと。帽子や耳栓は特別に許可を得た場合を除き、使用を認めない。
9. 試験時間中に目薬・点鼻薬・吸入薬を使用したい場合や、水分補給をしたい場合、ハンカチ（無地の物に限る）を使用したい場合は、監督者に申し出て許可を受けてから行うこと。それ以外の飲食は認めない。
10. 途中で体調が悪くなった場合は、手を挙げて監督者に申し出ること。エチケット袋が必要な場合は手渡しする。状況によっては、別室での受験を指示することがある。体調が悪くなった場合は、キャンパス内の保健センターで診察を受けられるが、解答時間の延長は認められない。
11. 解答が終わった場合でも、また、受験を放棄する場合でも、許可が出るまで退室できない。
12. 解答は、必ず解答用紙の指定された箇所に記入すること。解答用紙の解答欄に、関係のない文字、記号、符号などを記入してはならない。また、解答用紙の欄外の余白には、何も書いてはならない。
13. 試験時間終了の際は、監督者の「解答やめ。」の指示と同時に筆記具と消しゴムを置くこと。監督者の指示があるまでは、退室してはならない。