

受験番号 Examinee number

--	--	--	--	--	--

東京大学 大学院新領域創成科学研究科 情報生命科学専攻

Department of Computational Biology, Graduate School of Frontier Sciences

The University of Tokyo

平成 21 (2009) 年度

2009 School Year

大学院入学試験問題 修士・博士後期課程

Graduate School Entrance Examination Problem Booklet, Master's and Doctoral Course

専 門 科 目

Specialties

平成 20 年 8 月 4 日 (月)

Monday, August 4, 2008

13:30~15:30

注意事項 Instructions

1. 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 problem booklet until the start of examination is announced.
2. 本冊子の総ページ数は 15 ページです。落丁、乱丁、印刷不鮮明な箇所などがあった場合には申し出ること。
This problem booklet consists of 15 pages. If you find missing, misplaced, and/or unclearly printed pages, ask the examiner.
3. 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。
Use black pencils (or mechanical pencils) to answer the problems.
4. 問題は 8 題出題されます。問題 1~8 から選択した合計 4 問に解答しなさい。ただし、問題 1~8 は同配点です。
There are eight problems (Problem 1 to 8). Answer four problems out of the eight problems. Note that Problem 1 to 8 are equally weighted.
5. 解答用紙は計 4 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 four answer sheets. You must use a separate answer sheet for each problem. You may continue to write your answer on the back of the answer sheet if you cannot conclude it on the front.
6. 解答は日本語または英語で記入しなさい。
Answers should be written in Japanese or English.
7. 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 problem number you are to answer. Fill the designated blanks at the top of this page with your examinee number.
8. 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 problem booklet.
9. 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.
10. 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 the problem 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

次の問題の解法を考えよう。

「 $n (\geq 5)$ 個の自然数から合計値が 3 の倍数になるように、重複なく 3 個を選択する」ただし、すべての組合せを調べずに、各自然数は高々 1 回だけ調べるようにしたい。以下の問いに答えよ。

- (1) $n = 4$ の場合、条件を満たす 3 個を選択できない場合があることを示せ。
- (2) $n \geq 7$ の場合、条件を満たす 3 個を選択できることを示せ。
- (3) $n = 5$ および $n = 6$ の場合、条件を満たす 3 個を選択できることを示せ。
- (4) 上の 2 つの性質を利用して、各自然数を高々 1 回だけ調べ、条件を満たす解を求めるアルゴリズムを設計せよ。

問題2

$n (\geq 1)$ 個の要素からなる配列を、ランダムに1つの要素を選択しながらソートするクイックソートについて、以下の問いに答えよ。

- (1) クイックソートのアルゴリズムを説明せよ。
- (2) クイックソートの速度が著しく劣化する場合を説明せよ。
- (3) 各要素を等確率で選択するとしたとき、 n 個の要素からなる配列を処理する計算コスト(データ比較の回数)の期待値 $C(n)$ は、次の式で表現できることを説明せよ。

$$C(n) = n - 1 + \frac{1}{n} \sum_{k=1}^n (C(k-1) + C(n-k)) \quad (n \geq 1)$$

ただし $C(0) = 0$ とする。

- (4) $C(n)$ は $2n \log_e n$ で近似できることを示せ。

問題3

マウスの2つの近交系 P1 と P2 の遺伝子座 X での遺伝子型をそれぞれ AA および BB で表すものとする。P1 と P2 を交配したところ、F1 世代の成熟個体の体重の平均は 50 g で標準偏差は 10 g、F2 世代の成熟個体の体重の平均は 50 g で標準偏差は 20 g であった。各世代における成熟個体の体重の分散は、遺伝的要因によって決まる成分と、遺伝的要因によって決まらない成分の和として表されるものとする。以下の問いに答えよ。

- (1) 遺伝子座 X が取り得る遺伝子型を F1 および F2 世代それぞれについて列挙せよ。
- (2) F1 世代での体重の分散を算出せよ。
- (3) F1 世代での体重の分散はどのような意味をもつか説明せよ。
- (4) F2 世代での体重の分散はどのような意味をもつか説明せよ。
- (5) 遺伝子座 X が体重に及ぼす効果について評価する方法を簡単に説明せよ。

問題4

2つの文字列 $x = x_1x_2 \cdots x_n$, $y = y_1y_2 \cdots y_m$ に対して、同じ文字が整列するカラム数が最大になるアラインメントを動的計画法で求める問題を考えるとき、最大整列カラム数を計算するためのスコア $F(s, t)$ (ただし $0 \leq s \leq n, 0 \leq t \leq m$) の漸化式は以下のように書くことができる。

$$F(i, j) = \max \begin{cases} F(i-1, j) \\ F(i, j-1) \\ F(i-1, j-1) + \delta(x_i, y_j) \end{cases} \quad \delta(a, b) = \begin{cases} 0 & \text{if } a \neq b \\ 1 & \text{if } a = b \end{cases}$$

以下の問いに答えよ。

- (1) $F(i, j)$ の初期条件を示せ。
- (2) $x = x_1x_2 \cdots x_n$ と $y = y_1y_2 \cdots y_m$ のアラインメントにおける最大整列カラム数を、スコア F を用いて書け。
- (3) ギャップ位置 1 か所につき d のペナルティを課すとき、つまり最大整列カラム数から (ギャップ数) $\times d$ を引いた値をアラインメントのスコアとすると、このスコアを求めるための漸化式を示せ。
- (4) 3つの文字列 $x = x_1x_2 \cdots x_n$, $y = y_1y_2 \cdots y_m$, $z = z_1z_2 \cdots z_r$ のマルチプルアラインメントを動的計画法で求める問題を考える。同じカラムに整列する文字のペアの数が最大になるアラインメントを、ギャップペナルティを考慮せずに求めることにする。つまり、

x : AC-CG

y : A-T-G

z : AATC-

では、AA が $(x_1, y_1), (x_1, z_1), (y_1, z_1)$ の 3 組、TT (y_2, z_3) の 1 組、CC (x_3, z_4) の 1 組、GG が (x_4, y_3) の 1 組でスコアが合計 6 となることにする。この場合のアラインメントのスコア (同じカラムに整列する文字のペアの最大数) を求めるための $F(i, j, k)$ の漸化式を書け。

問題5

下にあるのはコドン表である。ただしアミノ酸表記の中で、Stop は終止（ストップ）コドンを示している。また、事象 E の生起確率が $P(E)$ のとき、 E の情報量を $I(E) = -\log P(E)$ と定義する。

UUU Phe	UUA Leu	UAU Tyr	UAA Stop	GAA Glu	GAU Asp	GUG Val	GUU Val
UUC Phe	UUG Leu	UAC Tyr	UAG Stop	GAG Glu	GAC Asp	GUA Val	GUC Val
UCU Ser	UCA Ser	UGU Cys	UGA Stop	GGA Gly	GGU Gly	GCA Ala	GCU Ala
UCC Ser	UCG Ser	UGC Cys	UGG	GGG	GGC Gly	GCG Ala	GCC Ala
CCC Pro	CCG Pro	CGC Arg	CGG Arg	AGG Arg	AGC Ser	ACG Thr	ACC Thr
CCU Pro	CCA Pro	CGU Arg	CGA Arg	AGA Arg	AGU Ser	ACA Thr	ACU Thr
CUC Leu	CUA Leu	CAC His	CAG Gln	AAG Lys	AAC Asn	AUA Ile	AUC Ile
CUU Leu	CUG Leu	CAU His	CAA Gln	AAA Lys	AAU Asn	AUG Met	AUU Ile

以下の問いに答えよ。

- (1) 真核細胞において最も一般的な開始（スタート）コドンはどれか。
- (2) 表中で抜けているアミノ酸名をそれぞれ答えよ（同じとは限らない）。
- (3) 表中で薄い色と濃い色で網掛けされたアミノ酸グループそれぞれの物理化学的特徴を述べ、根拠を簡潔に示せ。
- (4) 全てのコドンが等確率で出現するモデルを想定したとき、もっとも情報量の少ないアミノ酸はどれか。複数ある場合は全て答えよ。
- (5) コドンの3文字目が認識できなくなっても情報量が変わらないアミノ酸はどれか。複数ある場合は全て答えよ。
- (6) 開始コドン以降の全てのコドンが等確率で出現するモデルを想定したとき、アミノ酸配列の長さの期待値を求めよ。

問題6

以下の問いに答えよ。

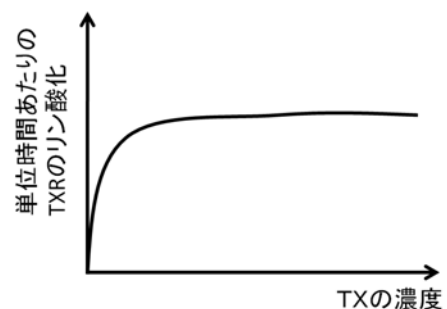
- (1) 真核生物クロマチンの基本単位となるヌクレオソームの構造について簡潔に説明せよ。
- (2) 真核生物の細胞骨格を3種類、それぞれの構成単位となっている蛋白質名とともにあげよ。更に、それぞれの細胞骨格の機能について簡潔に説明せよ。
- (3) グルコースの代謝に関する以下の問いに答えよ。
 - 1) 解糖系によって1分子のグルコースから、それぞれ何分子のピルビン酸、ATP、GTP、NADH、FADH₂、CO₂が生じるか。
 - 2) 強く収縮した筋細胞で、ピルビン酸は何に代謝されるか。また、この反応が進むのは何のためだと考えられるか。
 - 3) 好氣的条件下で、ピルビン酸は更にアセチル CoA に代謝される。この時、1分子のピルビン酸から、それぞれ何分子の ATP、GTP、NADH、FADH₂、CO₂が生じるか。
 - 4) 1分子のアセチル CoA が TCA サイクルで代謝されると、それぞれ何分子の ATP、GTP、NADH、FADH₂、CO₂が生じるか。
 - 5) NADH、FADH₂ から受け取った電子が電子伝達系を流れるのと共役して、プロトンはミトコンドリア内のどこからどこへと運ばれるか。その結果生じたプロトン勾配が果たす役割とは何か。

問題7

下記の文章を読んで下記の問いに答えよ。

Kw病は、患者が1万人に1人の頻度で見つかる常染色体性劣性の遺伝病である。連鎖解析の手法で、Kw病の原因遺伝子座位を1 Mbp ($=1 \times 10^6$ bp) の領域に追いつめることができた。この領域の塩基配列を決定したところ、N末端から順に、細胞外TX結合ドメイン、膜貫通領域、そして細胞内プロテインキナーゼドメインを持つTXレセプター(TXR)をコードする遺伝子TXRが見出された。

- (1) Kw病の保因者は、何人に1人の頻度で見つかるかと推定されるか。
- (2) 連鎖解析とは何か。DNA多型マーカー、家系、組み換えの3つのキーワードを用いて、簡潔に説明せよ。
- (3) TXRの膜貫通領域は α ヘリックスを形成していると予測された。 α ヘリックスの構造について、ポリペプチド鎖のどこどこが水素結合するのかを含めて、簡潔に説明せよ。さらにこの膜貫通ヘリックスを構成するアミノ酸として、相応しいものを3つ挙げてその理由を述べよ。
- (4) TXRは、細胞内のどこで合成され、どのような経路を辿って細胞膜へと運ばれるか。関与するオルガネラを順に挙げながら説明せよ。
- (5) Kw病の患者AとBのTXR遺伝子を調べたところ、それぞれ細胞外TX結合ドメインと細胞内プロテインキナーゼドメインに、ミスセンス変異が見つかった。健常者、患者AおよびBに由来する細胞に、様々な濃度のTXを加えて、単位時間あたりのTXRのリン酸化を調べると、三者でどのような差が見られると期待されるか。右図のようなグラフを描き、健常者と比較して、それぞれの患者で変化しない点を明確にしなしながら、説明せよ。



問題8

下記の塩基配列（塩基配列は5'→3'の方向で一方の鎖の配列だけを記している）は、ある仮想のほ乳類のゲノム配列の一部である。この配列に関して以下の問いに答えよ。

TCAAGCGGAAACCAGGTTGACCAATTTATAAAAATTTATTTGCTAAGCTGATCCATGGCC
AAAAACACACACACACACACACACACACACACACACACACATTTTCAGACGTATCA
TGA CTGATAGCATGACATGACGATCATGCCCRTTTGGCCCTTTTAACGCATGACTGAGG
GTAAAAGCCCTTGGACATAGAATTCGCGGGGGGCTTGCCGGGCGCGAATTCCCCGCCC
GCGCTTGCGCCATACGCGGTTTCGCCACGGACAGCCTCGCACCGGATCCTTTTTTTGAT
TAAACCCCTAGCTAGTCGATGCATATGGTCCAGTGCTAATTAGCTAGGGTT⁺¹GAAGTCCC
GACAAATTCCAGATCAGATCAGTAGTCCATGGCATTTTTTTGTCATCAGTCAGATCGAAG
ACAGTACAATTTCTGCGGGGCCCTXXTATAACACATGTATGACTGACATCAATGACCC
TCCCAAATTTTGCGGATARTTTTGATTGCCCTGATGATAGAAATTCACATAAACCC
TGAATCTCTCYCTCCTTCTCCTTCTTTGAAGATCCCCCAGGGCTCCTTTACCAAACTG
GTCCAACTCAGCAGGCCAGCTTCGCXXCGAATTCCAAACCCTTCACAGTTCATTTCA

- (1) データベースにある完全長 cDNA との相同性検索を行ったところ、灰色で塗りつぶした配列部分が一致した。また、+1 で示した G は完全長 cDNA の最初の塩基であった。灰色で塗りつぶした配列部分を何と言うか、一般名で答えよ。
- (2) 2 つの灰色で塗りつぶした配列の間にある部分（下線つき太字で示した XX から AG までの配列）は cDNA 配列中に存在しなかった。この中間部分の一般名を記せ。また、XX で表示した 2 塩基はさまざまな生物種で極めて高度に保存された配列である。もっとも頻度高く出現する XX の 2 塩基を記せ。
- (3) このほ乳類を数個体調べた結果、R および Y で記した個体間で異なった 2 種類の塩基をもつ多型部位があった。R および Y は、取りうる 2 つの塩基を同時に示す塩基記号である。R および Y に該当する 2 塩基をそれぞれ記せ。
- (4) この塩基配列の中には、R と Y 以外に個体間での多型性が極めて高い配列部分が他にも存在する。その配列部分と理由を記せ。

- (5) 2つの GAATTC 配列で挟まれた 400 塩基対の DNA をプラスミドベクターに組み込んだ。この組換えプラスミド DNA の大きさはちょうど 4000 塩基対となった。ついで、この組換えプラスミドをもつ大腸菌を十分量の培地で 10 時間培養した。培養開始時点で 1000 個の大腸菌があったと仮定して以下の質問に答えよ。ただし、 $2^{10} = 1000$ と近似して良い。また値は指数表現でも可とする。
- 1) GAATTC は、ある種の酵素で切断される配列である。このような特異的配列を認識して切断する酵素の一般名を答えよ。
 - 2) 今回用いた培養条件での大腸菌の分裂速度を 20 分としたとき、10 時間後の培養液中の大腸菌数を求めよ。
 - 3) 1 個の大腸菌に 1000 コピーのプラスミドが存在する場合、10 時間培養後のプラスミド分子の総数を求めよ。値は指数表現でも可とする。
 - 4) 10 時間培養後に大腸菌を溶菌してプラスミド DNA を回収した。回収率を理論値の 90%としたとき、回収されたプラスミド DNA の重量 (μg) を求めよ。ただし、DNA の 1 塩基対の分子量を 600、アボガドロ数を 6×10^{23} とする。
 - 5) 回収したプラスミドから 400 塩基対の DNA 部分だけを設問(5-1)の酵素を用いて切り出して回収した。この回収した 400 塩基対の DNA を 1 ml の水に溶解したときの濃度は $0.216 \mu\text{g}/\mu\text{l}$ となった。この操作における回収率 (%) を求めよ。

(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)