

受 験 番 号					

問題冊子にも受験番号を書きなさい。

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

平成16(2004)年度大学院入学試験問題

修士・博士後期課程

専 門 科 目

平成15年8月27日(水)

13:00~15:00(120分)

注意事項

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。
2. 本冊子の総ページ数は10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。
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。
4. 問題は7題出題されます。問題1~3の3問と問題4~7から選択した2問の合計5問に解答しなさい。
5. 解答用紙は計5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。
6. 解答は日本語または英語で記入しなさい。
7. 解答用紙の指定された箇所に、受験番号と選択問題の場合はその用紙で解答する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。
8. 草稿用紙は本冊子から切り離さないこと。
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。
10. 解答用紙・問題冊子は持ち帰ってはいけません。

(このページは草稿用紙として使用してよい)

問題 1 : 以下の問に答えよ。

$G$  を  $N$  個の頂点  $v_1, v_2, \dots, v_N$  から成る自己ループを含まない有向グラフとする。 $A_{ij}$  を  $v_i$  から  $v_j$  への有向辺の本数とすると、 $G$  の接続行列を  $A = \{ A_{ij} \}$  と定義する ( $1 \leq i, j \leq N$ )。頂点間の経路長をその経路が通過する有向辺の本数で定義する。

問 1 .  $v_1$  から  $v_3$  を通って  $v_2$  に至る長さ 2 の経路が存在するための必要十分条件は  $A_{13} \cdot A_{32} \geq 1$  であることを示せ。

問 2 .  $v_1$  から  $v_2$  に至る長さ 2 の経路の総数を求めよ。

問 3 . 任意の正整数  $r$  に対して接続行列  $A$  の  $r$  乗  $A^r$  を計算するアルゴリズムを記述せよ。

問 4 . 任意の正整数  $r$  に対して、 $(A^r)_{ij}$  を接続行列  $A$  の  $r$  乗  $A^r$  の第  $ij$  成分とする。 $(A^r)_{ij}$  は  $v_i$  から  $v_j$  に至る長さ  $r$  の経路の総数であることを証明せよ。

問題 2 : 以下の問に答えよ。

問 1 . 以下の実験方法について、原理を簡潔に 5 行以内で述べよ。

- a . dideoxy 法による DNA の塩基配列決定
- b . 蛋白質の二次元ゲル電気泳動法
- c . 酵母 two-hybrid 法

問 2 . 真核生物の蛋白質をコードする核内遺伝子は、多くの場合、エキソンとイントロンからなる。このような構造を持つ遺伝子からの初期 RNA 転写産物は、種々の修飾を受けた後に核外に輸送され、蛋白質合成の鋳型に供される。初期 RNA 転写産物は核内でどのような修飾を受けるか、受ける順番に従って記せ。

問題3：以下の問に答えよ。

問1．表3-1は塩基置換モデルの2を底とした対数オッズ行列である。transitionとtransversionの起こりやすさについて、このモデルでは何が仮定されているか。

注) transitionとはAとGの間およびTとCの間の塩基置換、transversionとはそれ以外の塩基置換を指す。

表3-1

	A	G	T	C
A	0	-7	-9	-9
G	-7	0	-9	-9
T	-9	-9	0	-7
C	-9	-9	-7	0

問2．配列データベース検索において、遠縁の遺伝子を探すために、通常はアミノ酸配列が使われる。なぜDNA配列が使われないのか、理由を述べよ。

問3．次ページ表3-2のようなアミノ酸の置換行列が与えられている。図3-1のアラインメントのスコアを求めよ。

```
LLLSAANVCLDP  
: . . . : . . . : . :  
LFIAATNICMDP
```

図3-1

表3 - 2

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
J	C	S	T	P	A	G	N	D	E	Q	H	R	K	M	I	L	V	F	Y	W
1	C	9	-1	-1	3	0	-3	-3	-3	-4	-3	-3	-3	-1	-1	-1	-1	-2	-2	-2
2	S	-1	4	1	-1	1	0	1	0	0	0	-1	-1	0	-1	-2	-2	-2	-2	-3
3	T	-1	1	5	-1	0	-2	0	-1	-1	-1	-2	-1	-1	-1	-1	0	-2	-2	-2
4	P	3	-1	-1	7	-1	-2	-2	-1	-1	-1	-2	-2	-1	-2	-3	-3	-2	-4	-3
5	A	0	1	0	-1	4	0	-2	-2	-1	-1	-2	-1	-1	-1	-1	0	-2	-2	-3
6	G	-3	0	-2	-2	0	6	0	-1	-2	-2	-2	-2	-3	-4	-4	-3	-3	-3	-2
7	N	-3	1	0	-2	-2	0	6	1	0	0	1	0	0	-2	-3	-3	-3	-2	-4
8	D	-3	0	-1	-1	-2	-1	1	6	2	0	-1	-2	-1	-3	-3	-4	-3	-3	-4
9	E	-4	0	-1	-1	-1	-2	0	2	5	2	0	0	1	-2	-3	-3	-2	-3	-3
10	Q	-3	0	-1	-1	-1	-2	0	0	2	5	0	1	1	0	-3	-2	-2	-3	-1
11	H	-3	-1	-2	-2	-2	-2	1	-1	0	0	8	0	-1	-2	-3	-3	-3	-1	2
12	R	-3	-1	-1	-2	-1	-2	0	-2	0	1	0	5	2	-1	-3	-2	-3	-3	-2
13	K	-3	0	-1	-1	-1	-2	0	-1	1	1	-1	2	5	-1	-3	-2	-2	-3	-2
14	M	-1	-1	-1	-2	-1	-3	-2	-3	-2	0	-2	-1	-1	5	1	2	1	0	-1
15	I	-1	-2	-1	-3	-1	-4	-3	-3	-3	-3	-3	-3	1	4	2	3	0	-1	-3
16	L	-1	-2	-1	-3	-1	-4	-3	-4	-3	-2	-3	-2	2	2	4	1	0	-1	-2
17	V	-1	-2	0	-2	0	-3	-3	-3	-2	-2	-3	-3	-2	1	3	1	4	-1	-3
18	F	-2	-2	-2	-4	-2	-3	-3	-3	-3	-3	-1	-3	-3	0	0	0	-1	6	3
19	Y	-2	-2	-2	-3	-2	-3	-2	-3	-2	-1	2	-2	-2	-1	-1	-1	-1	3	7
20	W	-2	-3	-2	-4	-3	-2	-4	-4	-3	-2	-2	-3	-3	-1	-3	-2	-3	1	2

問4 . 図3 - 1のように、2本の配列の第*i*要素 ( $i=1, \dots, n$ )同士が同じ位置に整列するアラインメントを「ギャップなしアラインメント」と呼ぶことにする。2本の長さ  $n$  のアミノ酸配列  $x(1), \dots, x(n)$ 、 $y(1), \dots, y(n)$  と表3 - 2のような置換行列  $S(j, k)$  ( $j, k=1, \dots, 20$ ) が与えられたとき、その2本の配列のギャップなしアラインメントのスコアを求めるプログラムを書け。

問題 4 : 下記の問題に答えよ。

$A$  を  $n$  行  $m$  列の実数値行列とし、 $i$  ( $1 \leq i \leq n$ ) 行目  $j$  ( $1 \leq j \leq m$ ) 列目の値を  $A_{ij}$  と表記する。また  $i$  行の最大値の列番号を  $\alpha(i)$  と記述する。ただし最大値が複数存在する場合は、最も左の列番号を選択する。また 実数  $x$  に対して  $x$  以上で最小の整数を  $\lceil x \rceil$  と記述する。

問 1 . 行  $i$  において行列要素の比較を  $(m-1)$  回実行して、 $\alpha(i)$  を計算するアルゴリズムを述べよ。

問 2 . 行  $i$  に関して  $\alpha(i)$  が単調に増加する、すなわち  $1 \leq i_1 \leq i_2 \leq n$  に対して  $\alpha(i_1) \leq \alpha(i_2)$  のとき  $A$  を単調な行列と呼ぶ。単調な行列  $A$  では、 $\alpha(\lceil n/2 \rceil)$  の値を利用して  $\lceil n/2 \rceil$  以外の行  $i$  の  $\alpha(i)$  を計算するときに、 $A$  の調べる範囲を限定できる。この方法について述べよ。

問 3 . 単調な行列  $A$  の全ての行  $i$  における  $\alpha(i)$  を、行列要素の比較を高々  $m \lceil \log_2(n+1) \rceil$  回実行して計算するアルゴリズムを述べよ。

問題 5 : 実験 1 ~ 4 に関する以下の問いに答えよ。

- (実験 1) サイトカイン X で刺激する前後のヒトリンパ球から RNA を単離して、DNA マイクロアレイ解析を行なった。その結果、サイトカイン X による刺激によってハイブリダイゼーションシグナルの増強を示す遺伝子が 50 個見いだされた。
- (実験 2) 実験 1 と同様の解析を転写阻害剤の存在下で行なったところ、上記の 50 個の遺伝子のうち 10 個では依然としてシグナルの増強を認めたが、他の 40 個ではシグナルの増強が認められなくなった。
- (実験 3) その 40 個の遺伝子の上流配列を計算機で比較解析した結果、共通配列として TTCTGGGA が同定された。この配列を含む 2 本鎖 DNA プローブの末端を  $^{32}\text{P}$  で標識して、ヒトリンパ球核抽出物とインキュベートした後、非変性ポリアクリルアミドゲルに展開したところ、2 本の移動度が減弱したバンド (B1 と B2) を見いだした。
- (実験 4) 実験 3 の反応に未標識の 2 本鎖 DNA プローブを過剰量添加すると、B1 の消失を認めたが、B2 は消失しなかった。

- 問 1 . 実験 1 において、逆転写反応によって総 RNA 中の mRNA (メッセンジャー RNA) を選択的に標識するには、どのようなプライマーを用いるとよいだろうか? 理由とともに簡潔に答えよ。
- 問 2 . 実験 2 で同定された 10 個の遺伝子が、サイトカイン X 処理によってハイブリダイゼーションシグナルの増強を示した機構としてはどのようなものが考えられるか。
- 問 3 . 実験 3 のインキュベーション反応で用いた緩衝液の組成は、20 mM Tris-HCl (pH X), Y mM KCl, 0.5 mM EDTA, 5% Glycerol である。細胞内の生理的条件にもっとも近い条件と考えられる X と Y の値を下記から選択せよ。
- $$X = \{6.0, 6.5, 7.0, 7.5, 8.0\}$$
- $$Y = \{0, 50, 100, 150, 200\}$$
- 問 4 . 実験 4 を行なった目的は何か。またその結果から何が示唆されたか。さらに行なうべき対照実験は何か。
- 問 5 . 共通配列 TTCTGGGA がサイトカイン X による遺伝子発現誘導に重要であることを直接的に示すには、どのような実験を行なえばよいだろうか。簡潔に述べよ。



問題6：以下の設問はそれぞれ一つの表題と括弧内の三つのキーワードから構成される。各々に含まれる三つのキーワードを用い、表題について解説した文章を記述せよ。説明文中では三つのキーワードに下線を記し、各キーワードに簡潔な解説を加えよ。

- 問1． 疾患遺伝子座の同定  
(マイクロサテライト、ロッド値、連鎖解析)
- 問2． 発現解析からの遺伝子ネットワークの推定  
(DNAマイクロアレイ、シスエレメント、自己組織化マップ)
- 問3． ゲノム塩基配列のアノテーション  
(EST、HMM、オントロジー)

問題7：蛋白質の立体構造に関する以下の問いに答えよ。

問1． 蛋白質の2次構造のうち、右巻きヘリックスと反平行シートの2つを主鎖原子間の化学結合と水素結合のパターンがわかるように図示せよ。なお、水素結合は破線で示すこと。

問2． 1次構造（アミノ酸配列）が判明しているが高次構造が知られていない蛋白質を考える。ある2次構造予測法を使って、この蛋白質の各残基がヘリックス（H）ストランド（E）、それ以外（-）であることを予測したところ、図7-1のような結果を得た。ここでは予測された2次構造を左から右へアミノ酸配列に沿って示してある。この予測結果から、この蛋白質にはどのような立体構造上の特徴があると考えられるか簡潔に説明せよ。

---E-----HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHE-----E-HHHHHHHHHHHHHHHHHHHHHHHHHHHHHH-E--E---H---E--E-

図7-1

問3． アラニン、グルタミン、イソロイシンの3種類のアミノ酸残基を疎水性の強い順に並べ、そのように並べた物理化学的理由を述べよ。

問4． 図7-2は、問2の蛋白質の各アミノ酸残基の疎水性をある疎水性指標を用いて計算し、左から右へアミノ酸配列に沿って示したものである。ただし、縦軸は疎水性の強さを表し正の値が大きいほど疎水性が強いものとする。図7-1及び図7-2から、この蛋白質は膜貫通型であると予想される。その理由を説明せよ。

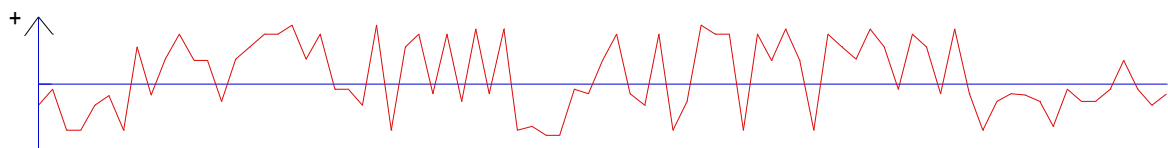


図7-2

問5． 蛋白質の1次構造が与えられたとき、その蛋白質が膜貫通型であるかどうかを判定して膜貫通部位を自動的に予測する方法を考案し、その計算の手順を説明せよ。