

情報生物学 (1 / 2)

(注意) 問題 [1] [2] はそれぞれ別の答案用紙に解答すること。

[1] バイオインフォマティクスにおける遺伝子やタンパク質の配列解析について以下の問いに答えなさい。(計 50 点)

- (1) 配列アラインメントは、複数の塩基配列やアミノ酸配列を並べたもので、配列の類似性を知る手がかりになる。配列アラインメントは、ペアワイズアラインメントとマルチプルアラインメントに大きく分かれるが、その違いを説明しなさい。(10 点)
- (2) 塩基配列やアミノ酸配列では、機能を担う重要領域は進化的に保存されるため、マルチプルアラインメントで共通あるいは類似した部分配列を検出できる。このような特定の配列パターンをモチーフと呼ぶ。モチーフの表現法として、モチーフの各位置において最も出現頻度の高い文字だけを取り出すコンセンサス配列と呼ばれる方法がある。表 1 はプロモータ領域に見られる TATA ボックスの各塩基の出現頻度 (単位は%) を表すが、この場合のコンセンサス配列を示しなさい。(10 点)

表 1

| | 位置 1 | 位置 2 | 位置 3 | 位置 4 | 位置 5 | 位置 6 |
|---|------|------|-------|------|------|------|
| A | 3.7 | 91.1 | 0.0 | 94.5 | 67.3 | 97.3 |
| T | 83.6 | 8.9 | 100.0 | 5.5 | 32.7 | 0.0 |
| G | 2.9 | 0.0 | 0.0 | 0.0 | 0.0 | 2.7 |
| C | 9.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

- (3) モチーフの各位置において複数の文字を許容し、確率的なモチーフ表現法として、位置特異的スコア行列 PSSM (position specific scoring matrix) と呼ばれる行列で表現する方法がある。行列の各要素としては、位置 i における文字 a の出現頻度 $p_{i,a}$ をバックグラウンド (配列全体) における出現頻度 b_a で割った値の対数値 $\log(p_{i,a}/b_a)$ がよく使われる。コンセンサス配列の方法に比べて、対数値に基づく位置特異的スコア行列の方法の利点を説明しなさい。(10 点)
- (4) 対数値に基づく位置特異的スコア行列で任意の配列を検索したとき、各位置のスコアの和をとることによって、そのモチーフに似ている度合いを評価できる。TATA ボックスの位置特異的スコア行列が表 2 のように示されているとする。GAATACCGAGATATATGCGA という配列を検索した時、そのモチーフに最も似ている領域を同定しなさい。また、そのときのスコアを計算しなさい。(10 点)

表 2

| | 位置 1 | 位置 2 | 位置 3 | 位置 4 | 位置 5 | 位置 6 |
|---|-------|-------|-------|-------|-------|-------|
| A | -1.85 | 1.30 | -4.61 | 1.33 | 0.99 | 1.36 |
| T | 1.21 | -1.39 | 1.39 | -1.47 | 0.28 | -4.61 |
| G | -2.07 | -4.61 | -4.61 | -4.61 | -4.61 | -2.14 |
| C | -0.91 | -4.61 | -4.61 | -4.61 | -4.61 | -4.61 |

- (5) 位置特異的スコア行列によるモチーフ表現法の欠点について述べなさい。またその欠点を克服する可能性があるとするれば、どのような方法が考えられるか説明しなさい。(10 点)

情報生物学 (2 / 2)

(注意) 問題 [1] [2] はそれぞれ別の答案用紙に解答すること。

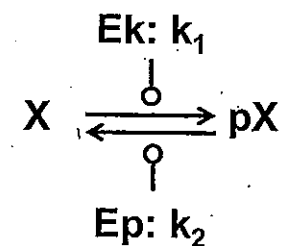
[2] 細胞内の酵素反応の特性について以下の文章を読んで、下記の問いに答えなさい。なお、(あ) ~ (き) および (a) ~ (c) には適切な語句や数式が入る。(計 50 点)

細胞内の酵素反応を考える時、(あ) や時定数は重要な指標である。そこで、細胞がどのようにこれらの指標を制御しているのかを考えてみる。酵素反応などが最大値の 50% を示す濃度のことを (あ) と呼ぶ。つまり、同じ効果を持つ酵素でも (あ) の値が (a) 分子の方が、入力が少なくとも同じ効果を持つことが可能である。一方で、時定数は応答速度の指標であり、簡単に言えばその分子を制御している分子にどの程度早く応答できるかの指標である。

ここで、次の酵素反応を考える。

X が Ek により k_1 という速度定数でリン酸化されて pX となる。

pX は Ep により k_2 という速度定数で脱リン酸化されて X となる。



この酵素反応は次の微分方程式で近似できる。

注: ただし [] は濃度を示す。

$$\frac{d[pX]}{dt} = k_1 \times [Ek] \times [X] - k_2 \times [Ep] \times [pX]$$

$X + pX = 1$ とした場合、式変形を行い

$$\frac{d[pX]}{dt} = -(\text{い}) \times [pX] + (\text{う})$$

となる。よって $t=0$ のとき pX が 0 の場合 ($pX(t=0)=0$),

$$pX(t) = (\text{え}) \times (1 - e^{-(\text{お}) \times t})$$

となる。ここで時定数は $1/(\text{お})$ となる。

次に pX の濃度に対する酵素 Ek の (あ) を求める。

定常状態を仮定した場合 ($t \rightarrow \infty$), pX(t) は

$$pX(t \rightarrow \infty) = [Ek] / (\text{か})$$

ここで (あ) は (き) となる。

細胞は (あ) や時定数をどのように変化させているかを考える。一つは酵素量を変える方法である。例えば、脱リン酸化酵素 Ep の量を多くすると上記の式から時定数は (b) なり、(あ) は (c) なる。

(1) (あ) ~ (き) に当てはまる語句や数式を入れなさい。違うひらがなに同じ語句や数式が入っても構わない。(各 6 点)

(2) (a) ~ (c) に当てはまる組み合わせを下から選びなさい (8 点)

- | | |
|----------------------------|----------------------------|
| ア) (a) 小さい、(b) 小さく、(c) 小さく | イ) (a) 小さい、(b) 小さく、(c) 大きく |
| ウ) (a) 小さい、(b) 大きく、(c) 小さく | エ) (a) 大きい、(b) 小さく、(c) 小さく |
| オ) (a) 小さい、(b) 大きく、(c) 大きく | カ) (a) 大きい、(b) 大きく、(c) 小さく |
| キ) (a) 大きい、(b) 小さく、(c) 大きく | ク) (a) 大きい、(b) 大きく、(c) 大きく |