

1 生命を情報として解析する

「進化思考の世界」(三中 信宏著 NHK ブックス) の 195 頁に塩基配列で種間の違いを分析する方法が説明・例題が以下の様に紹介されています。

いま、生物 P、Q、R、S のいずれもが持つある遺伝子の塩基配列が、次のようになったとしよう。

```
P AACCTGCTA
Q AAGCTCTAT
R ATGCTCAAC
Q ATGGTCAAC
```

この時、9 個の塩基からなる配列という「情報」を考えることができる。

(中略)

たとえば、P と Q の間で塩基配列を比較すれば、塩基の異なる位置は、第 3 番目と第 6 から 9 番目の計 5 個であるから両者の違いは、「5」ということになる。

同様にしてすべてのペアの間で塩基配列の違いを数値化してみると、次の様になる。

	P	Q	R	S
P	0	5	6	7
Q	5	0	3	4
R	6	3	0	1
S	7	4	1	0

「進化思考の世界」(三中 信宏著 NHK ブックス) の 195 頁より

1.1 塩基配列の違い bioinfo.c

各塩基配列から、上記の様な表を作成します。

```

/*
 塩基配列の違いを計算します。
  bioinfo.c
*/
#include <stdio.h>

int main()
{
  char base_name[4] = "PQRS";
  char base[][9]    = {"AACCTGCTA",
                      "AAGCTCTAT",
                      "ATGCTCAAC",
                      "ATGGTCAAC"};

  int gap;
  int table[4][4];

  int i;
  int j;
  int k;

  printf("塩基配列の違いを計算します。 \n");

  for(i = 0; i < 4; i++){
    for(j = 0; j < 4; j++){
      gap = 0;
      for(k = 0; k < 9; k++){
        if(base[i][k] != base[j][k]){
          gap++;
        }
      }
      table[i][j] = gap;
    }
  }

  printf(" ");
  for(i = 0; i < 4; i++){
    printf("%c ", base_name[i]);
  }
  printf("\n");

```

```
for(i = 0; i < 4; i++){
    printf("%c ", base_name[i]);
    for(j = 0; j < 4; j++){
        printf("%d ", table[i][j]);
    }
    printf("\n");
}

return 0;
}
```

1.2 計算結果

塩基配列の違いを計算します。

```
 P Q R S
P 0 5 6 7
Q 5 0 3 4
R 6 3 0 1
S 7 4 1 0
```