

Bioinformatyka ĆWICZENIE 6

1. Zasady stosowania pakietu SAS - prezentacja.
2. Omówienie programu SAS stosowanego na ćwiczeniach.
3. Samodzielne przygotowanie programu w SAS służącego do obliczania wsp. rekombinacji pomiędzy markerami z własnego zbioru danych:
 - przeliczyć wsp. rekombinacji na cM przy użyciu funkcji Haldane'a
 - przeliczyć wsp. rekombinacji na cM przy użyciu funkcji Kosambi

Wzór programu:

```
/* Wczytanie pliku */
data B;
    infile "C:/Documents and Settings/UP/Pulpit/dane/1.csv" missover dsd
    delimiter="," ;
    input ID SIRE $ DAM $ WEIGHT $ L1A1 L1A2 L2A1 L2A2 L3A1
    L3A2 L4A1 L4A2 L5A1 L5A2 L6A1 L6A2 L7A1 L7A2 L8A1 L8A2 L9A1 L9A2 L10A1 L10A2;
run ;

/*Tworzenie haplotypów dwumarkerowych*/
data REC1 (keep=H12 H23);
    set B ;
    H12 = L1A1*10+L2A1;
    H23 = L2A1*10+L3A1;
    output;
    H12 = L1A2*10+L2A2;
    H23 = L2A2*10+L3A2;
    output;
run;

/*Obliczanie liczebności poszczególnych haplotypów*/

    proc freq data=REC1;
        tables H12 / out=COUNTHAP12 ;
run;

    proc print data=COUNTHAP12 ;
run;

/*Sortowanie haplotypów wg liczebności*/
    proc sort data=COUNTHAP12 ; by COUNT ;
run;

    proc print data=COUNTHAP12 ;
run;

/*Obliczanie ws. rekombinacji i cM wg funkcji Haldane'a*/
data COUNTHAPa12 ;
    retain N 0 ;
    set COUNTHAP12 ;
    N=N+1 ;
run;

    proc print data=COUNTHAPa12 ;
run;

data FINAL ;
    retain REC 0 NREC 0 ;
    set COUNTHAPa12 ;
    if N=1 or N=2 then REC=REC+COUNT ;
    else if N=3 or N=4 then NREC=NREC+COUNT ;
    RECOMBINATION=REC/(REC+NREC) ;
run ;

proc print data=FINAL ;
run ;
```