

## Program PHASE oraz pakiet R – zadania do samodzielnego wykonania

1. Ze strony [www.theta.edu.pl](http://www.theta.edu.pl) pobrać plik o nazwie *phase\_R.prn* oraz zapisać go na dysku H.
2. Na podstawie pobranego pliku utworzyć plik wejściowy do programu PHASE i zapisać go pod nazwą *phase1.inp*.
3. Uruchomić program PHASE dla utworzonego wcześniej pliku tak, aby plik wynikowy miał nazwę *phase1\_res.out*.
4. Plik *phase\_R.prn* wczytać do programu R jako tablicę o nazwie *genotyp*.
5. Na podstawie tablicy *genotyp*, dla każdego markera obliczyć frekwencje homozygot recesywnych (11), homozygot dominujących (22) oraz heterozygot (12, 21). Wyniki mają się znaleźć w tabeli o nazwie *frekwencje* o wymiarach 12x3, gdzie w pierwszej kolumnie znajdują się frekwencje homozygot recesywnych, w drugiej kolumnie – frekwencje heterozygot, w trzeciej kolumnie – frekwencje homozygot dominujących.
6. Tablicę *frekwencje* zapisać w pliku o nazwie *frekwencje.txt*.
7. Otrzymane wyniki dotyczące frekwencji przedstawić na odpowiednim wykresie w Excelu (należy pamiętać o opisaniu wykresu).
8. Plik wynikowy *phase1\_res.out\_freqs* wczytać do programu R jako tablicę o nazwie *haplotyp*.
9. Obliczyć średnią, odchylenie standardowe oraz medianę frekwencji haplotypów.
10. Wyznaczyć haplotyp o największej i najmniejszej frekwencji.
11. Frekwencje haplotypów przedstawić na wykresie (funkcja *plot()* w pakiecie R).