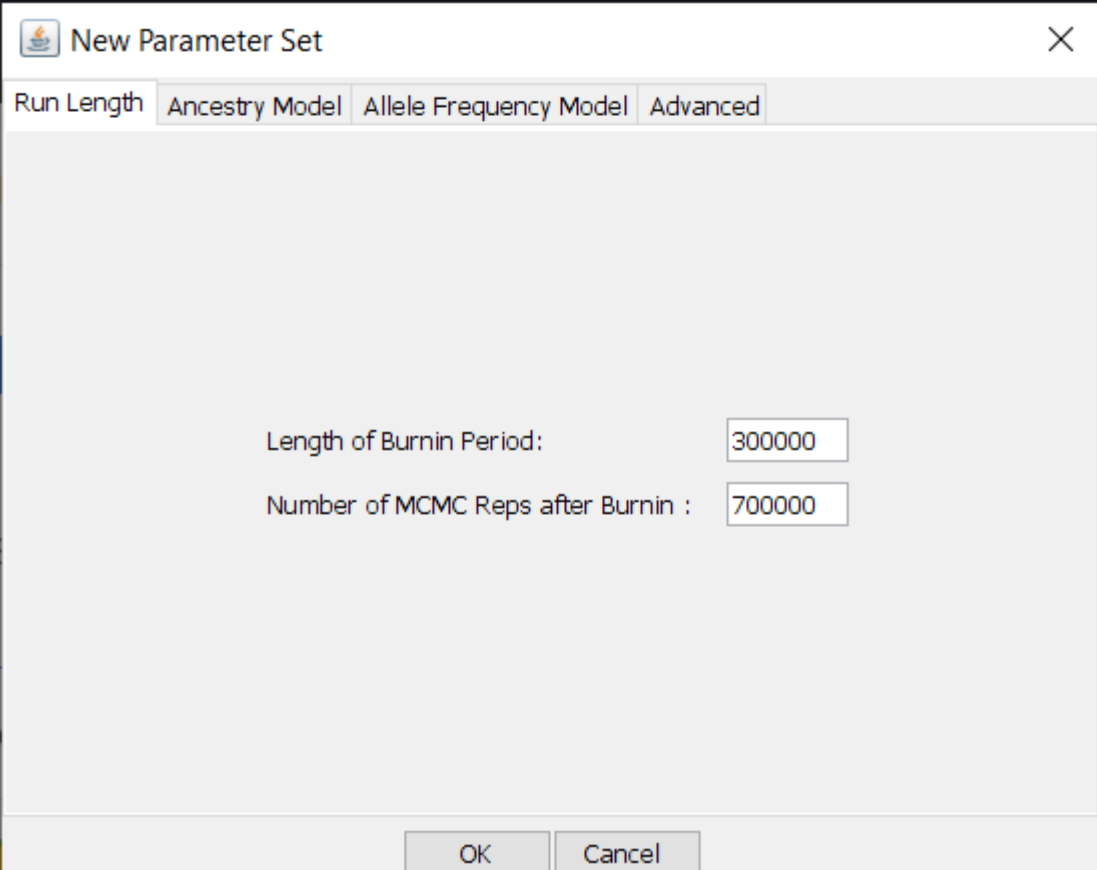


Paweł Grygielski (121678)  
Michał Somala (121725)

## Techniki przepływu genów - projekt zaliczeniowy

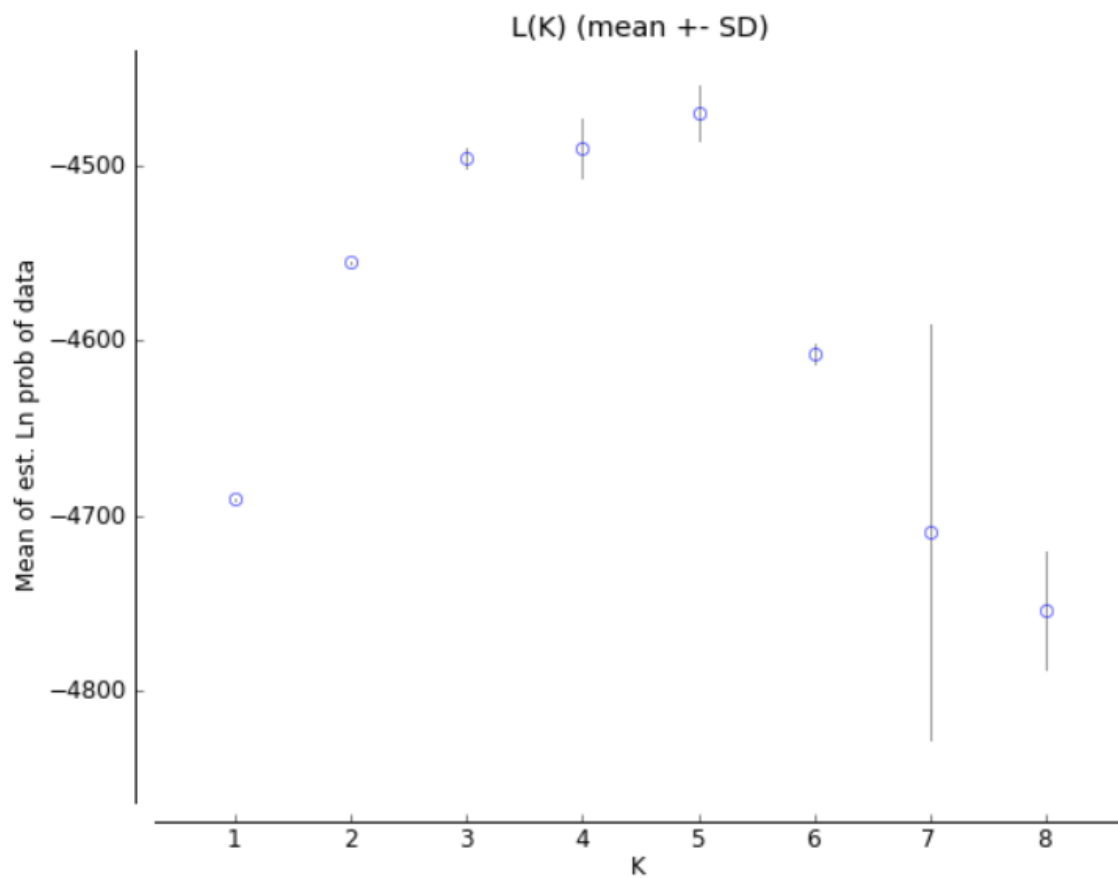
### a) ANALIZA STRUKTURY GENETYCZNEJ POPULACJI (KLASTROWANIE STRUCTURE)

Parametry, które zostały wprowadzone do programu structure:

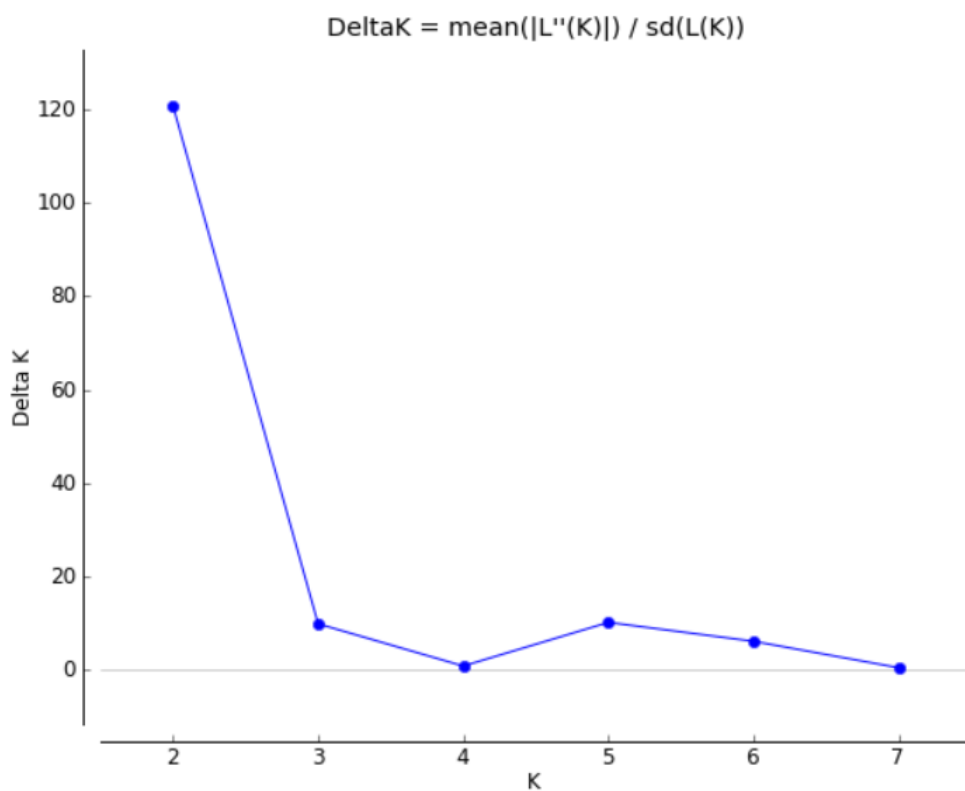


The screenshot shows a window titled "New Parameter Set" with a close button (X) in the top right corner. Below the title bar, there are four tabs: "Run Length", "Ancestry Model", "Allele Frequency Model", and "Advanced". The "Run Length" tab is currently selected. The main area of the window contains two input fields: "Length of Burnin Period:" with a value of 300000, and "Number of MCMC Reps after Burnin :" with a value of 700000. At the bottom of the window, there are two buttons: "OK" and "Cancel".

## L(K)



Na podstawie wykresu L(K), dwie (najniższe odchylenie) lub trzy (niskie odchylenie, wyższa wartość) odrębne genetycznie supbopulacje, to najbardziej prawdopodobna struktura genetyczna badanej populacji.

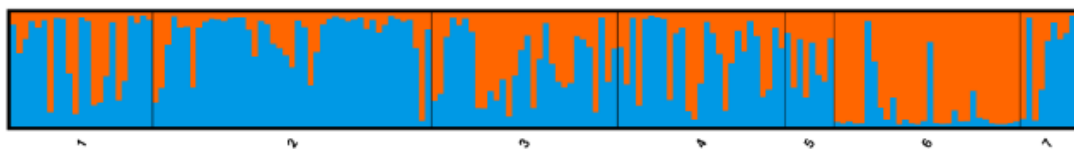


Na podstawie wykresu  $\Delta K$ , dwie odrębne genetycznie supbopulacje, to najbardziej prawdopodobna struktura genetyczna badanej populacji.

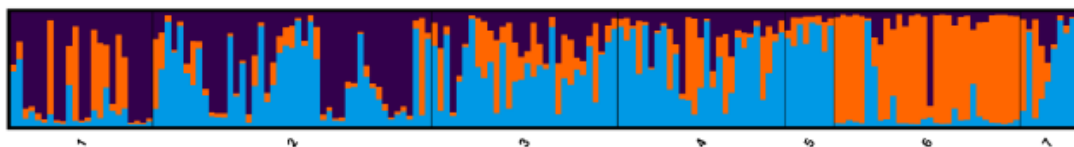
**K=1** 5/5, Mean(LnProb) = -4690.340, Mean(similarity score) = 1.000



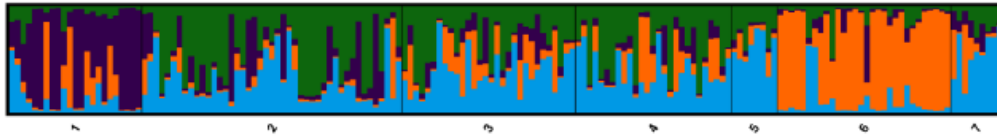
**K=2** 5/5, Mean(LnProb) = -4555.200, Mean(similarity score) = 0.996



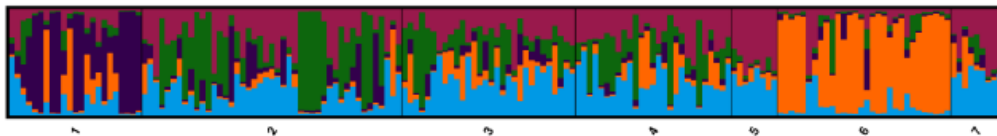
**K=3** 5/5, Mean(LnProb) = -4495.000, Mean(similarity score) = 0.994



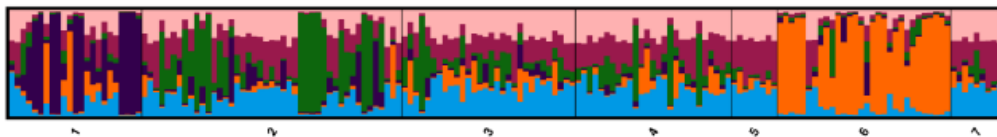
**K=4** 5/5, Mean(LnProb) = -4489.380, Mean(similarity score) = 0.888



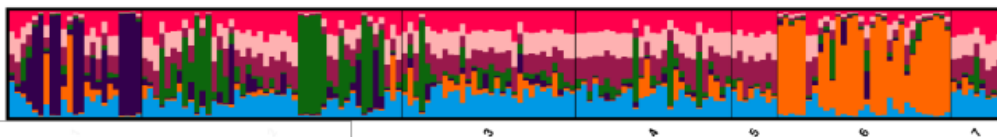
**K=5** 5/5, Mean(LnProb) = -4469.680, Mean(similarity score) = 0.919



**K=6** 5/5, Mean(LnProb) = -4606.820, Mean(similarity score) = 0.949



**K=7** 5/5, Mean(LnProb) = -4709.180, Mean(similarity score) = 0.965



/1686055277/K=4.MajorCluster.png

dla K=2 similarity score jest najwyższe, dodatkowo na wykresie DeltaK, 2 ma wartość zdecydowanie najwyższą. Na wykresie L(K) 2 ma najniższe odchylenie. W związku z powyższym dwie odrębne genetycznie subpopulacje, to najbardziej prawdopodobna struktura genetyczna badanej populacji.

## b) PODSTAWOWE PARAMETRY ZMIENNOŚCI GENETYCZNEJ DLA POSZCZEGÓLNYCH LOCI:



	N	He	Ho	F <sub>IS</sub>	F <sub>ST</sub>
Locus 1	9	0,83258669	0,7642276	0,046	0,047
Locus 2	7	0,76853643	0,6181818	0,118	0,106
Locus 3	13	0,82171513	0,6993007	0,084	0,084
Locus 4	5	0,70898145	0,6515152	0,036	0,054
Locus 5	8	0,70559701	0,702381	-0,062	0,076
Locus 6	18	0,85864913	0,8451613	-0,044	0,068
Locus 7	7	0,76711442	0,6167665	0,140	0,077
Locus 8	9	0,83567999	0,754491	0,061	0,045
Locus 9	7	0,71930829	0,6554054	0,048	0,052
Average	9,2	0,780	0,701	0,048	0,068

- N - W analizowanej populacji liczba alleli w locus waha się od 5 do 18, a średnia wartość to 9,2
- Ho i He - Średnia heterozygotyczność obserwowana(Ho) dla wszystkich loci jest mniejsza niż oczekiwana(He), co świadczy o niedoborze heterozygot. Różnica pomiędzy średnimi He i Ho wynosi 0,079, co nie jest wysoką wartością.
- Fis - średnie Fis = 0,048, świadczy to o bardzo niewielkim niedoborze heterozygot. Dla poszczególnych loci, te wartości wahają się od -0,062, co oznacza niewielki nadmiar heterozygot, do 0,118 niewielki deficyt heterozygot.
- Średnia wartość Fst i wynosi 0,068. Wartości bliskie 0 i oznaczają dobry przepływ genów między lokalizacjami, wartość 0,068 jest bliska 0. Można więc powiedzieć, że pomiędzy subpopulacjami jest dobry przepływ genów a co za tym idzie, subpopulacje te są do siebie podobne pod względem genetycznym.

### **c) SPRAWDZANIE CZY POPULACJA I SUBPOPULACJE SĄ W STANIE RÓWNOWAGI HARDY'EGO-WEINBERGA**

Locus	wartość p value w populacjach						
	1	2	3	4	5	6	7
1	1	0,4755	0,0378	0,8642	0,7282	0,7293	0,5344
2	0,1462	0,0131	0,8058	0,2041	0,5897	0,6949	0,0198
3	0,9629	0,0083	0,9718	0,1627	1	0,2332	0,2361
4	0,2349	0,8471	0,2566	0,088	0,4779	0,4145	0,3883
5	0,4199	0,9267	0,5628	0,5557	0,5887	0,3639	0,0687
6	0,3291	0,2188	0,5252	0,8997	0,5594	0,2373	0,2451
7	0,0532	0,2235	0,0595	0,665	0,4219	0,15	0,4384
8	0,2971	0,186	0,2478	0,2091	0,4871	0,0825	0,8157
9	0,3065	0,1681	0,6063	0,5001	0,5897	0,783	0,0775

W przypadku naszej populacji żadne locus po korekcie nie odbiega istotnie od równowagi Hardy'ego-Weinberga, można więc uznać, że równowaga jest zachowana.

#### **d) PRZEPŁYW GENÓW MIĘDZY POPULACJAMI GEOGRAFICZNYMI**

##### Pairwise FST

```
Estimates for all loci (diploid):
```

```
=====
```

```
pop      1      2      3      4      5      6
2      0.0740
3      0.0882  0.0266
4      0.1153  0.0458  0.0194
5      0.1514  0.0555  0.0610  0.0788
6      0.1182  0.0812  0.0521  0.0992  0.1098
7      0.0987  0.0342  0.0508  0.0777  0.0750  0.0823
```

Nm – efektywna liczba migrantów na pokolenie

$F_{st} = 1/(4Nm+1)$  |

$nm = 1/F_{ST}/4 - 1$

##### EFEKTYWNA LICZBA MIGRANTÓW

pop	1	2	3	4	5	6
2	2,378378					
3	1,834467	8,398496				
4	1,168257	4,458515	11,8866			
5	0,651255	3,504505	3,098361	2,172589		
6	1,115059	2,078818	3,798464	1,520161	1,276867	
7	1,532928	6,309942	3,92126	2,217503	2,333333	2,037667

Efektywna liczba migrantów jest najwyższa pomiędzy populacją 3 i 4, a najniższa pomiędzy populacją 1 i 5.