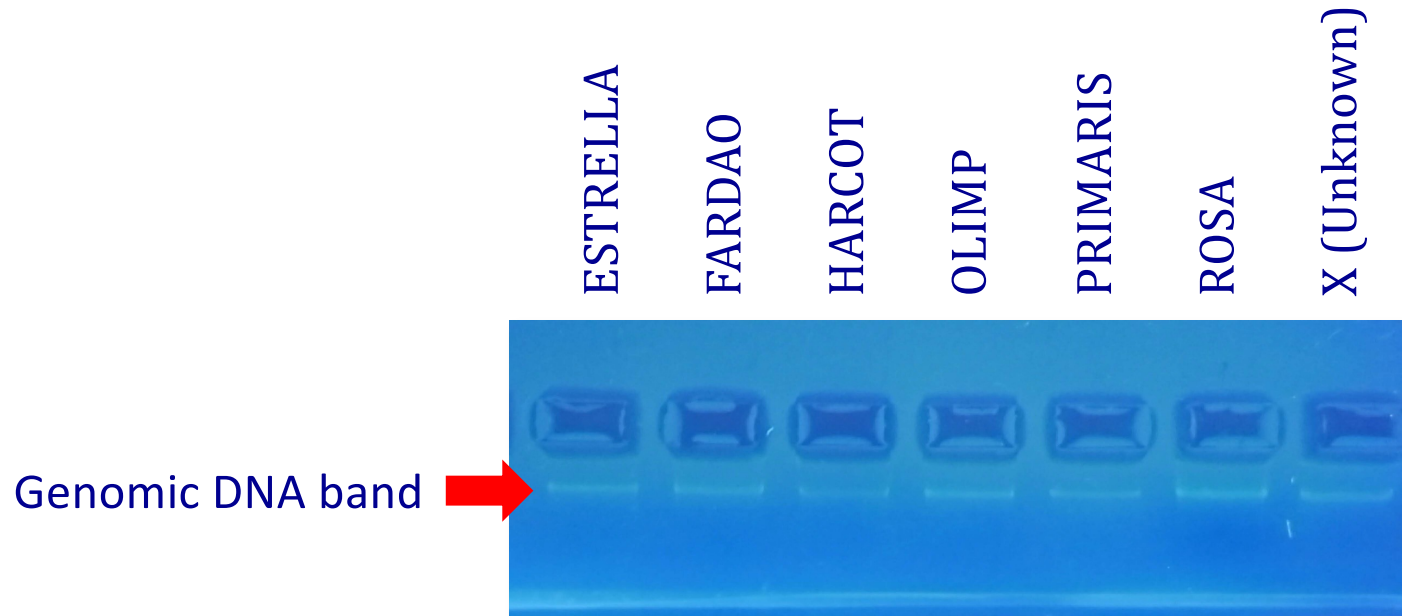


First lab activity

(DNA extraction from apricot leaves and agarose gel)

WE 26-05-2021



Second lab activity

(PCR amplification of six microsatellite loci)

WE 26-05-2021

SAMPLES

ESTRELLA
FARDAO
HARCOT
OLIMP
PRIMARIS
ROSA
X (Unknown)



SSR name	LG	T°C ann	T°C PCR	Allele size LxS	PAGE run HxR	PAGE buffer HxR	Provider	Primer Forward	Primer Reverse
CD195SSR	L1	57	50	190	150 min 30 mA	SB	Sicard et al.	CCTCTGATGTATTATCTTTCTGGC	TTCAACAGCTCCAAATTCAC
CD211SSR	L1	57	50	141	40+140 min 30mA	SB	Sicard et al.	GAAGATCATGTTAGAGAATAGTGG	CAGGGGTGACTTGGAAACC
Gol001	L1		55	200	150 min 30 mA	SB	Badenes	TGCAATCGATGAAGATTTGAC	TGGTCACCTTCTTCATATCC
N86B11SSR2a	L1	57	50	271	140 min 58 W	TBE	Decroocq (Marandel et al.)	GCCAATTTAACGCCGAAT	CTGGTTCCTATACCTGCATCA
PacC3	L1	56	56		120 min 30 mA	SB	Decroocq et al.	TGACTTGATCAGACTCGACA	TTGCATTTCATTTACAATAGA
UDA-025	L1	56	55		30+120 min 30 mA	SB	Testolin et al.	TCGAGAAAGCTGCACTGGTA	AAAGCTGCTTATTCGTGTGTG

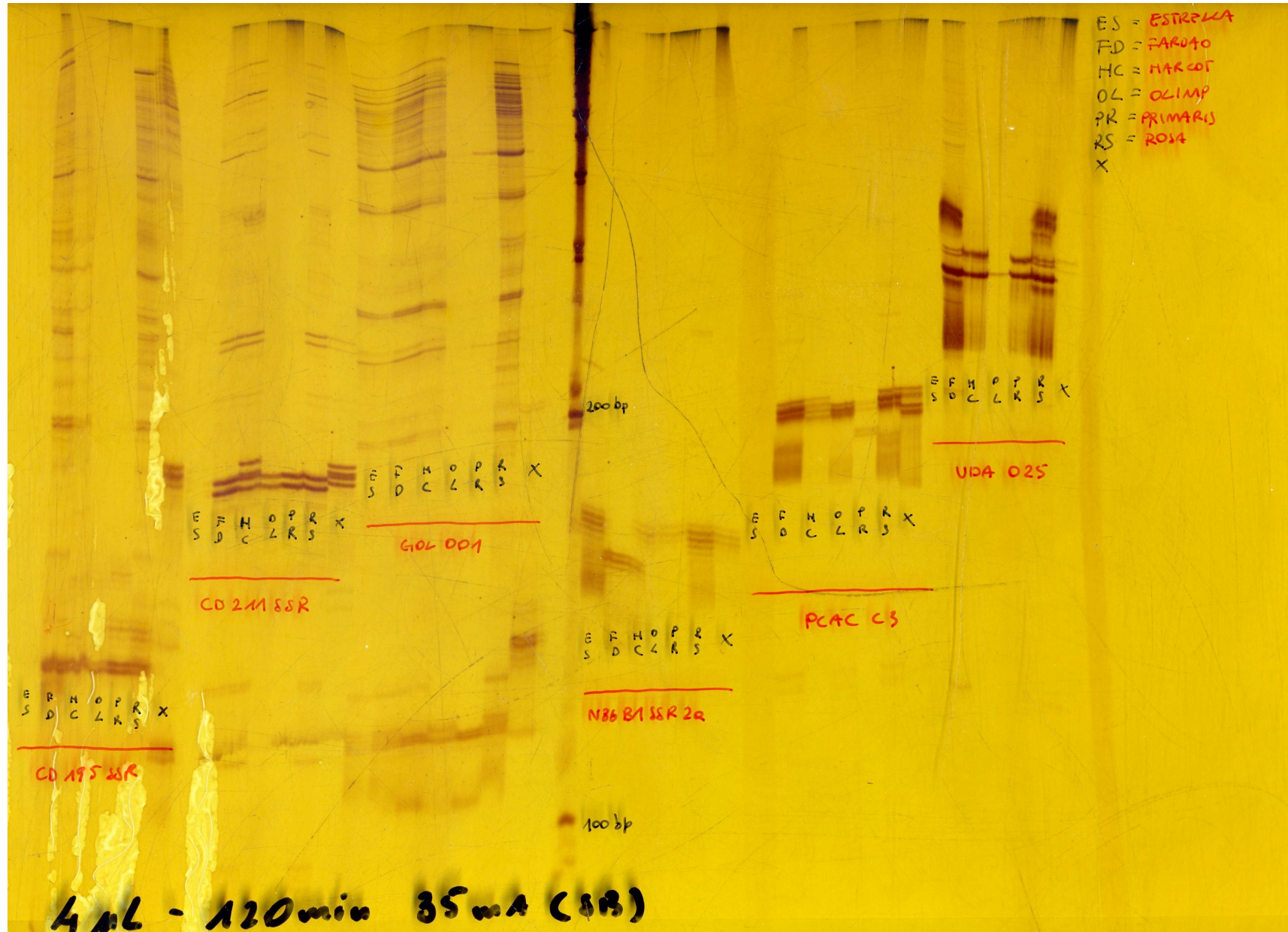
Apricot PCR	Stock	Units	Final concentration	ul for each sample	ul for nr. of samples
DNA	10	ng/uL	10	1	
10X buffer A	10	X	1	2	
MgCl2	25	mM	2,5	2	
dNTPs	2	mM	0,1	1	
Primer for	5	uM	0,25	1	
Primer Rev	5	uM	0,25	1	
Taq polymerase	5	U/uL	0,025	0,1	
H2O	-	-	-	11,9	

Program	Time	Temp	Cycles
1	05:00	94°	1x
2	00:30	94°	
2	01:30	57°	35x
2	00:30	72°	
4	05:00	72°	1x
5	∞	10°	/

Third lab activity

WE 26-05-2021

(polyacrylamide gel electrophoresis of PCR products)



Fourth lab activity

WE 26-05-2021

(some questions that could be addressed by this experiment)

Calculate the genetic distance matrix and the *Heterozygosity* (He) for the six SSR loci

Can you draw any hypothesis on the possible parental relationships between individuals?

Does the “X” individual show any similarity with the known accessions?

What is the degree of genetic similarity between the six accessions?

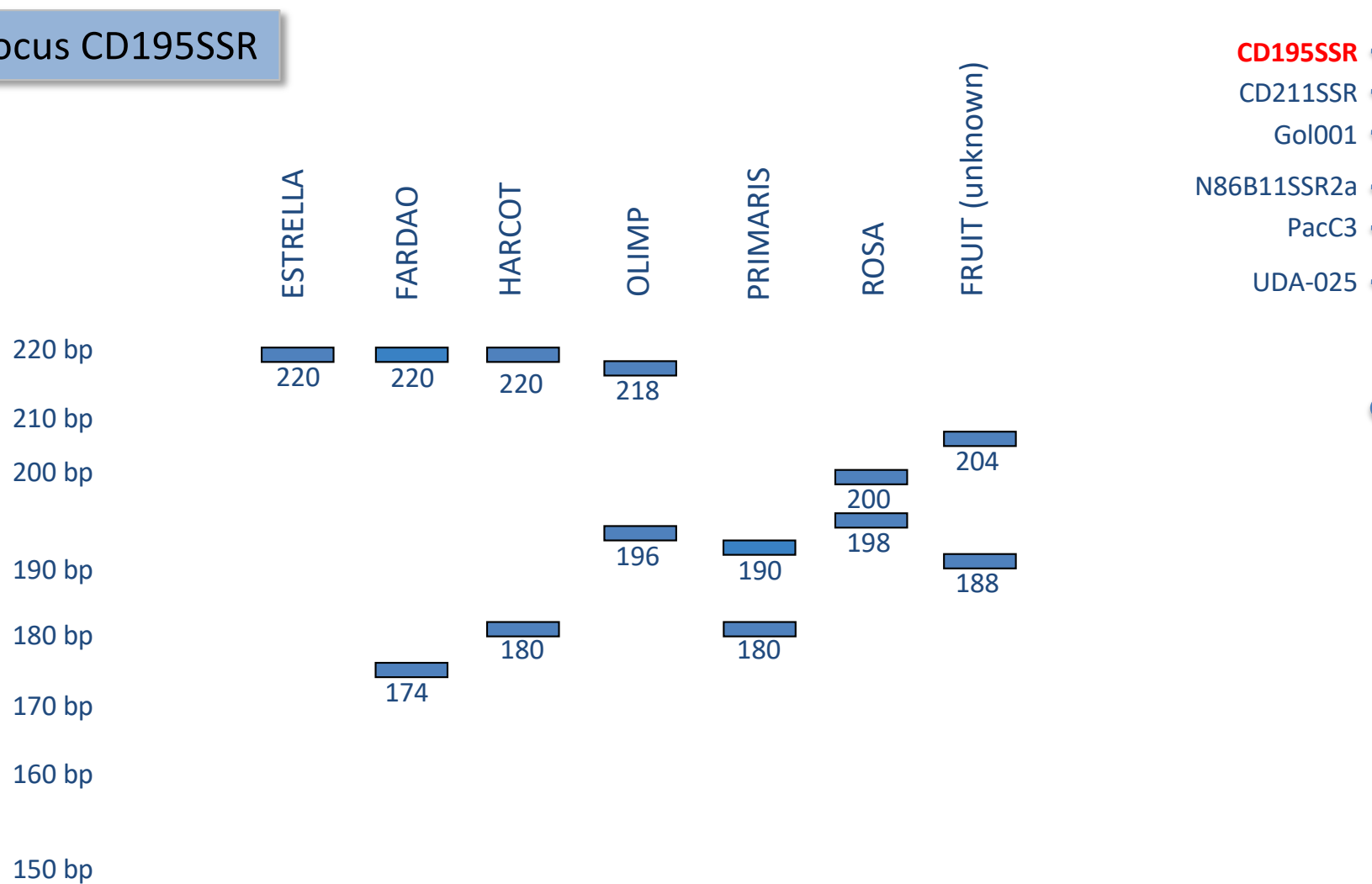
What can you tell about the possible resistance to PPV virus for these accessions?

Which accession(s) would you employ in a marker-assisted breeding (MAS) program targeting Sharka resistance?

Fourth lab activity (data collection and analysis)

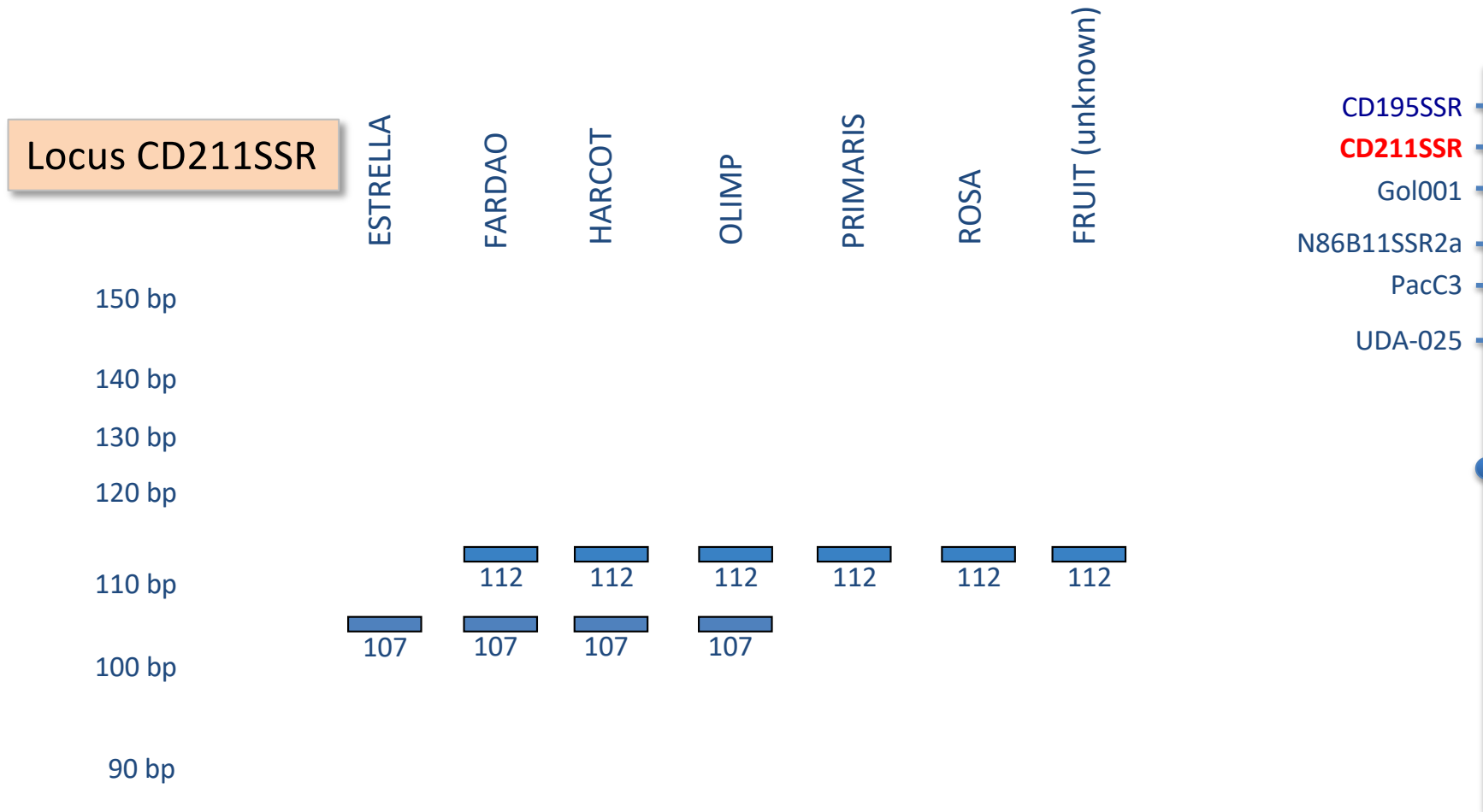
WE 26-05-2021

Locus CD195SSR



Fourth lab activity (data collection and analysis)

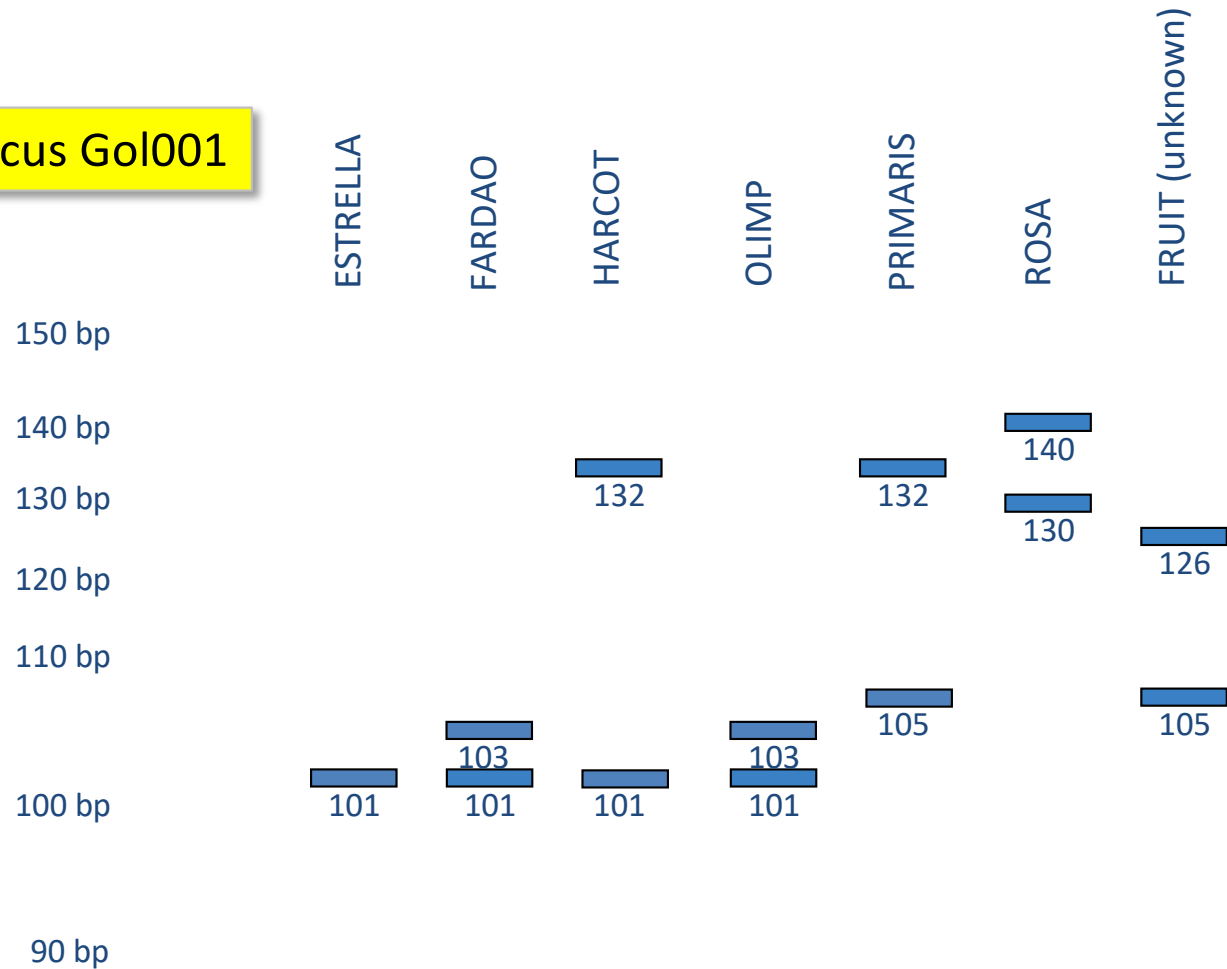
WE 26-05-2021



Fourth lab activity (data collection and analysis)

WE 26-05-2021

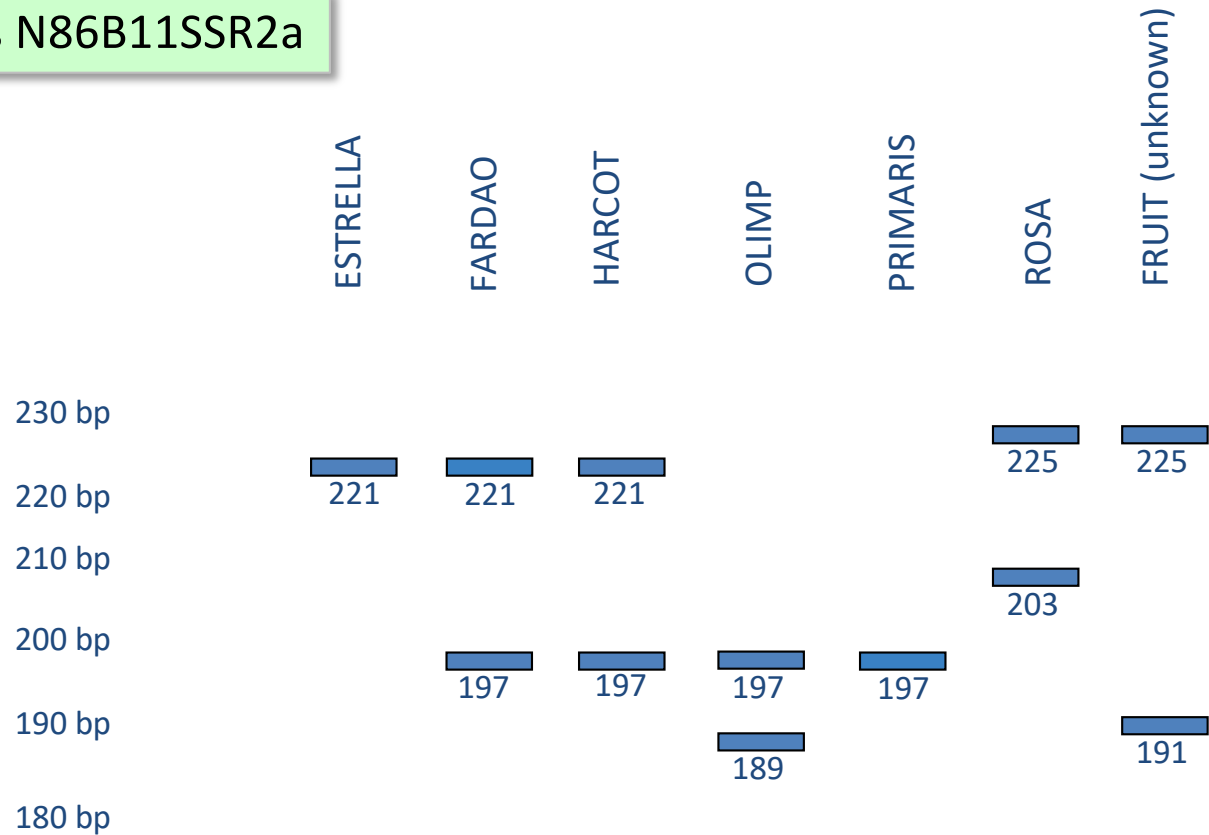
Locus Gol001



Fourth lab activity (data collection and analysis)

WE 26-05-2021

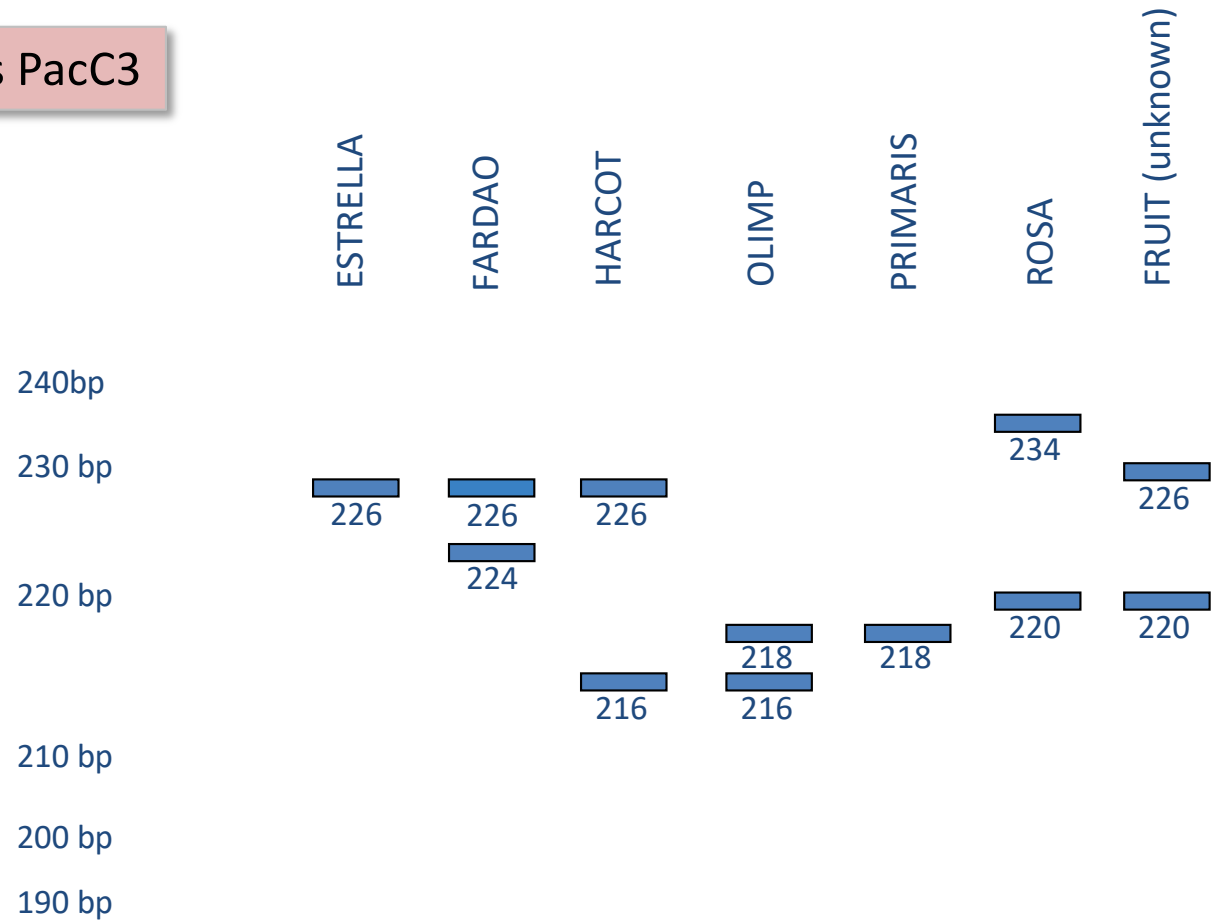
Locus N86B11SSR2a



Fourth lab activity (data collection and analysis)

WE 26-05-2021

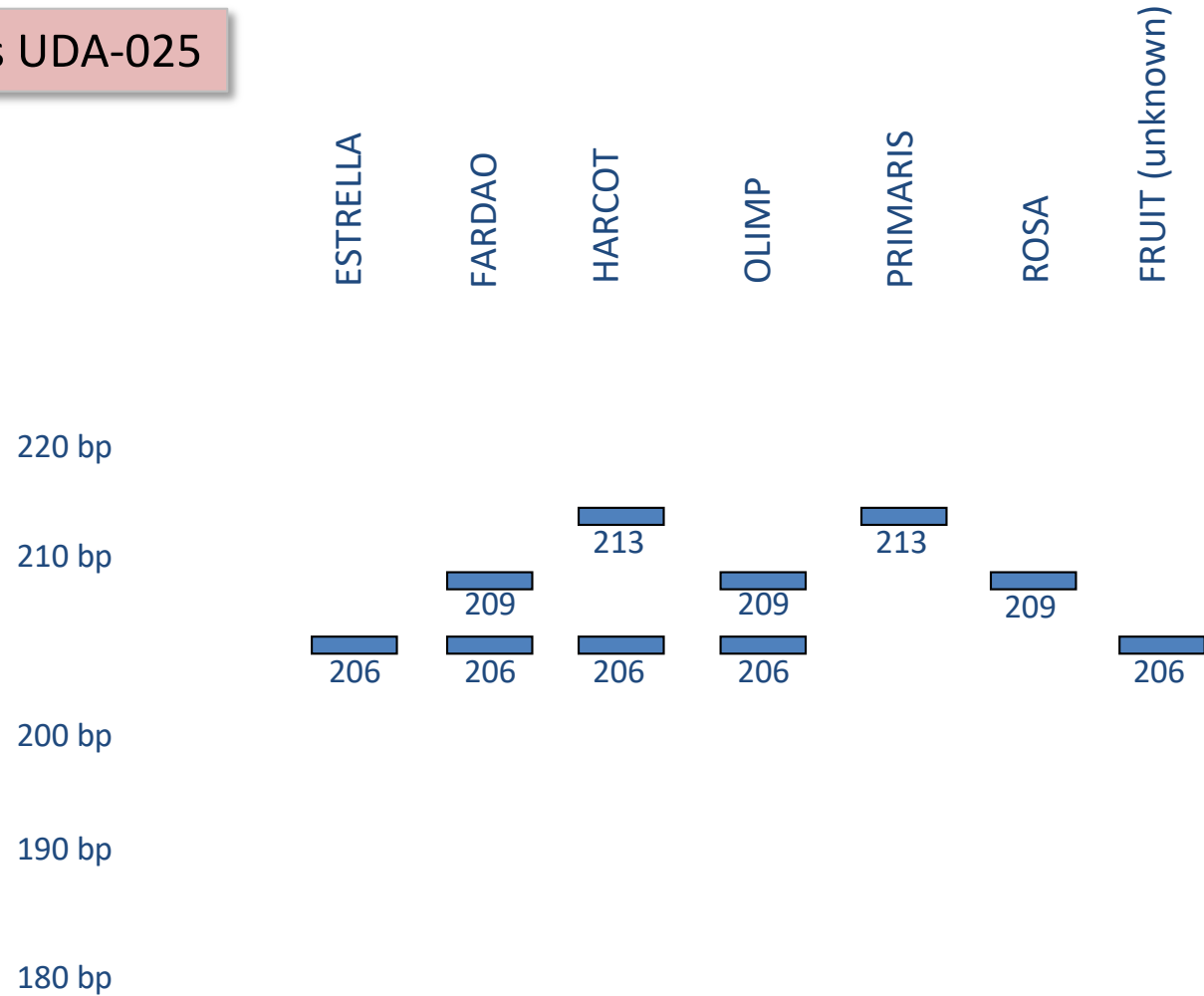
Locus PacC3



Fourth lab activity (data collection and analysis)

WE 26-05-2021

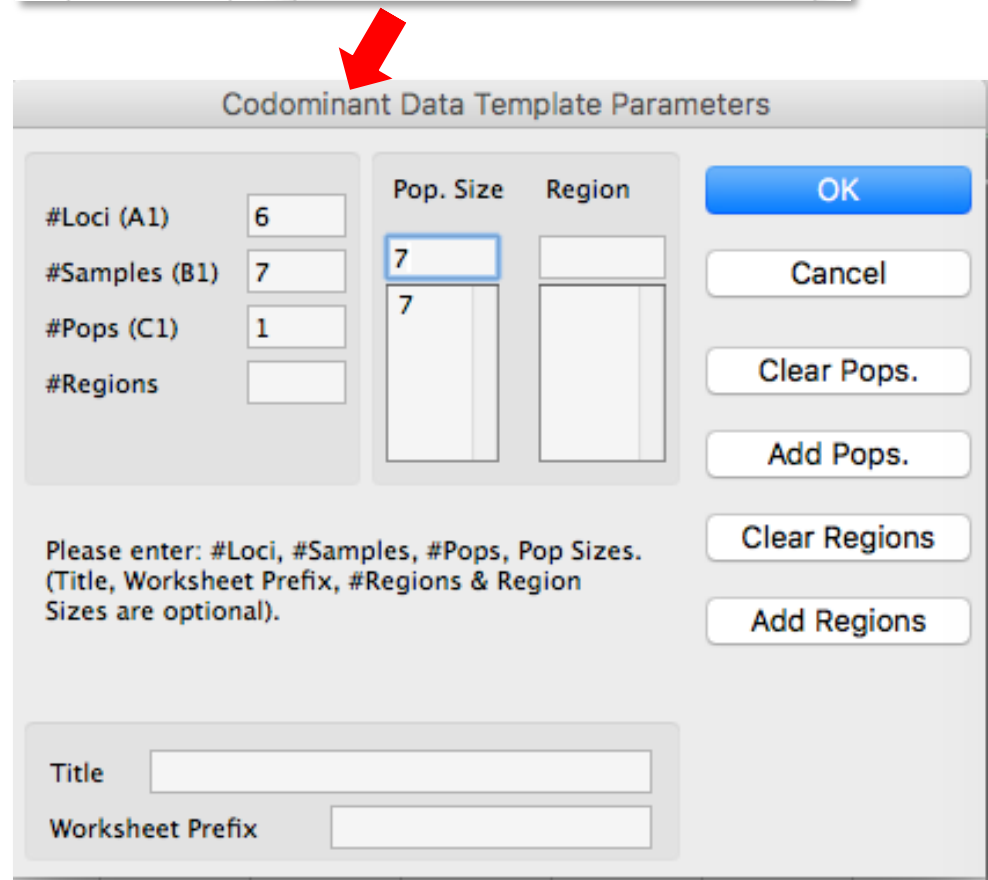
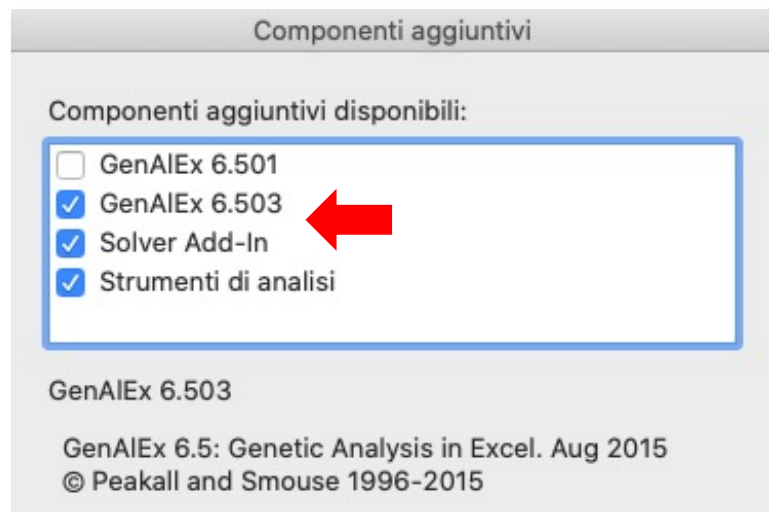
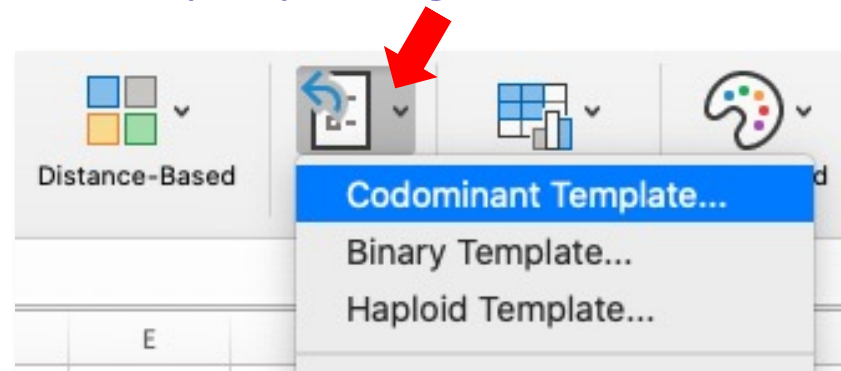
Locus UDA-025



Fourth lab activity

WE 20-12-2023

(the GenAEx software: preparing data)



Fourth lab activity

WE 20-12-2023

(some questions that could be addressed by this experiment)

Working file



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	6	7	1	7	7										
2	Codominant data template		Pop1		Region1										
3	Sample	Pop	CD195SSR		CD211SSR		Gol001		N86B11SSR2a		PacC3		UDA-025		
4	ESTRELLA	Pop1	220	220	107	107	101	101	221	221	226	226	206	206	
5	FARDAO	Pop1	220	174	112	107	103	101	221	197	226	224	209	206	
6	HARCOT	Pop1	220	180	112	107	132	101	221	197	226	216	213	206	
7	OLIMP	Pop1	218	196	112	107	103	101	197	189	218	216	209	206	
8	PRIMARIS	Pop1	190	180	112	112	132	105	197	197	218	218	213	213	
9	ROSA	Pop1	200	198	112	112	140	130	225	203	234	220	209	209	
10	UNKNOWN	Pop1	204	188	112	112	126	105	225	191	226	220	206	206	
11															
12															
13															
14															
15															

Fourth lab activity

WE 20-12-2023

(some questions that could be addressed by this experiment)

The screenshot shows a software interface with a spreadsheet and a 'Genetic Distance Options' dialog box. The spreadsheet contains data for 10 samples (ESTRELLA, FARDAO, HARCOT, OLIMP, PRIMARIS, ROSA, UNKNOWN) across 6 loci (A, B, C, D, E, F). The 'Genetic Distance Options' dialog box is open, showing settings for #Loci (6) and #Samples (7). The 'Distance Calculation' section has 'Two Cols/Locus' selected with 'Codom-Genotypic' as the method. The 'Output' section has 'To Worksheet', 'As Tri Matrix', 'Label Matrix', and 'Sample' selected. The 'Title' field is set to 'Codominant data template'.

Sample	Pop	CD195SSR	A	B	C	D	E	F
ESTRELLA	Pop1		220	220	174	112	107	
FARDAO	Pop1		220	174	112	107		
HARCOT	Pop1		220	180	112	107		
OLIMP	Pop1		218	196	112	107		
PRIMARIS	Pop1		190	180	112	112		
ROSA	Pop1		200	198	112	112		
UNKNOWN	Pop1		204	188	112	112		

Fourth lab activity (GenAEx Tutorials)

WE 20-12-2023

Welcome

Download

Tutorials

Export

Citation

Authors

GenAEx Tutorials

An Overview of Topics

We are pleased to offer a series of self-paced tutorials on population genetic analysis that employ hand calculations and exercises within GenAEx. These are drawn in part from the graduate workshops that we have offered (jointly and independently), around the world. Click the links below to download any tutorials of interest. In 2012, Tutorials 1-6 were revised to bring them up to date with the new features of GenAEx 6.5. The newer Trouble Shooting Tutorial is strongly recommended for all users. It provides helpful tips for solving some of the issues that may prevent some data sets from running.

Tutorial 1 (zip 6.9 mb)

An Introduction to Frequency-Based Population Genetic Analysis: scoring genetic markers, Allele Frequency, Heterozygosity, F-statistics, Nei Genetic Distance, Shannon Diversity Indices and Chi-square tests for Hardy-Weinberg Equilibrium



Tutorial 2 (zip 1.3 mb)

Genetic Distance and AMOVA: Haploid, Codominant and Binary Genetic Distance, AMOVA and F-statistics



Tutorial 3 (zip 4.5 mb)

Spatial Genetic Analysis: Principal Coordinate Analysis (PCoA), Mantel Tests for Matrix Correspondence and Spatial Autocorrelation Analysis

Fourth lab activity (formatting the total distance matrix)

WE 20-12-2023

	A	B	C	D	E	F	G	H	I	J
1	1	7	1	7	1	7				
2	Codominant	D	Codominant	Pop1						
3	ESTRELLA	FARDAO	HARCOT	OLIMP	PRIMARIS	ROSA	UNKNOWN			
4	0							ESTRELLA		
5	6	0						FARDAO		
6	6	4	0					HARCOT		
7	12	5	6	0				OLIMP		
8	22	12	8	10	0			PRIMARIS		
9	20	10	12	10	14	0		ROSA		
10	14	9	9	10	13	10	0	UNKNOWN		
11										
12										
13										
14										
15										

GD D +



	A	B	C	D	E	F	G	H	I	J
1	ESTRELLA	0								
2	FARDAO	6	0							
3	HARCOT	6	4	0						
4	OLIMP	12	5	6	0					
5	PRIMARIS	22	12	8	10	0				
6	ROSA	20	10	12	10	14	0			
7	UNKNOWN	14	9	9	10	13	10	0		
8										
9										
10										
11										
12										

GD D +

Fourth lab activity (formatting the total distance matrix)

WE 20-12-2023

	A	B	C	D	E	F	G	H	I	J
1	ESTRELLA	0								
2	FARDAO	6	0							
3	HARCOT	6	4	0						
4	OLIMP	12	5	6	0					
5	PRIMARIS	22	12	8	10	0				
6	ROSA	20	10	12	10	14	0			
7	UNKNOWN	14	9	9	10	13	10	0		
8										
9										
10										
11										
12										



Add "7"

	A	B	C	D	E	F	G	H	I	J
1	7-									
2	ESTRELLA	0								
3	FARDAO	6	0							
4	HARCOT	6	4	0						
5	OLIMP	12	5	6	0					
6	PRIMARIS	22	12	8	10	0				
7	ROSA	20	10	12	10	14	0			
8	UNKNOWN	14	9	9	10	13	10	0		
9										

Ten characters

Fourth lab activity (the GenAEx software)

WE 20-12-2023

Principal Component Analysis

The screenshot shows the GenAEx software interface with the 'GenAEx' tab selected. The 'Distance' menu is open, and the 'PCoA' option is highlighted. A sub-menu is also open, showing 'Analysis...' as the selected option. The spreadsheet below contains data for a Principal Component Analysis, with columns labeled A through M and rows numbered 1 through 16. The data includes population names and numerical values.

	A	B	C	D										
1	1	7	1	7										
2	Codominant	D	Codominant	Pop1										
3	ESTRELLA	FARDAO	HARCOT	OLIMP	PRIMARIS	ROSA	U							
4	0													
5	6	0												
6	6	4	0											
7	12	5	6	0										
8	22	12	8	10	0									
9	20	10	12	10	14	0								
10	14	9	9	10	13	10	0	UNKNOWN						
11														
12														
13														
14														
15														
16														



Fourth lab activity (the GenAEx software)

WE 20-12-2023

Principal Component Analysis

The screenshot shows the Microsoft Excel interface with the GenAEx software overlaid. The GenAEx 'PCoA Parameters' dialog box is open, showing the following settings:

- Input Data Type: Tri Distance Matrix, Distance Matrix as Column
- #Samples: 7
- PCoA Method: Covariance-Standardized, Covariance-Not Standardized, Distance-Standardized, Distance-Not Standardized
- Graph Options: Data Labels, Color Code Pops
- Title: Codominant data template
- Worksheet Prefix: (empty)

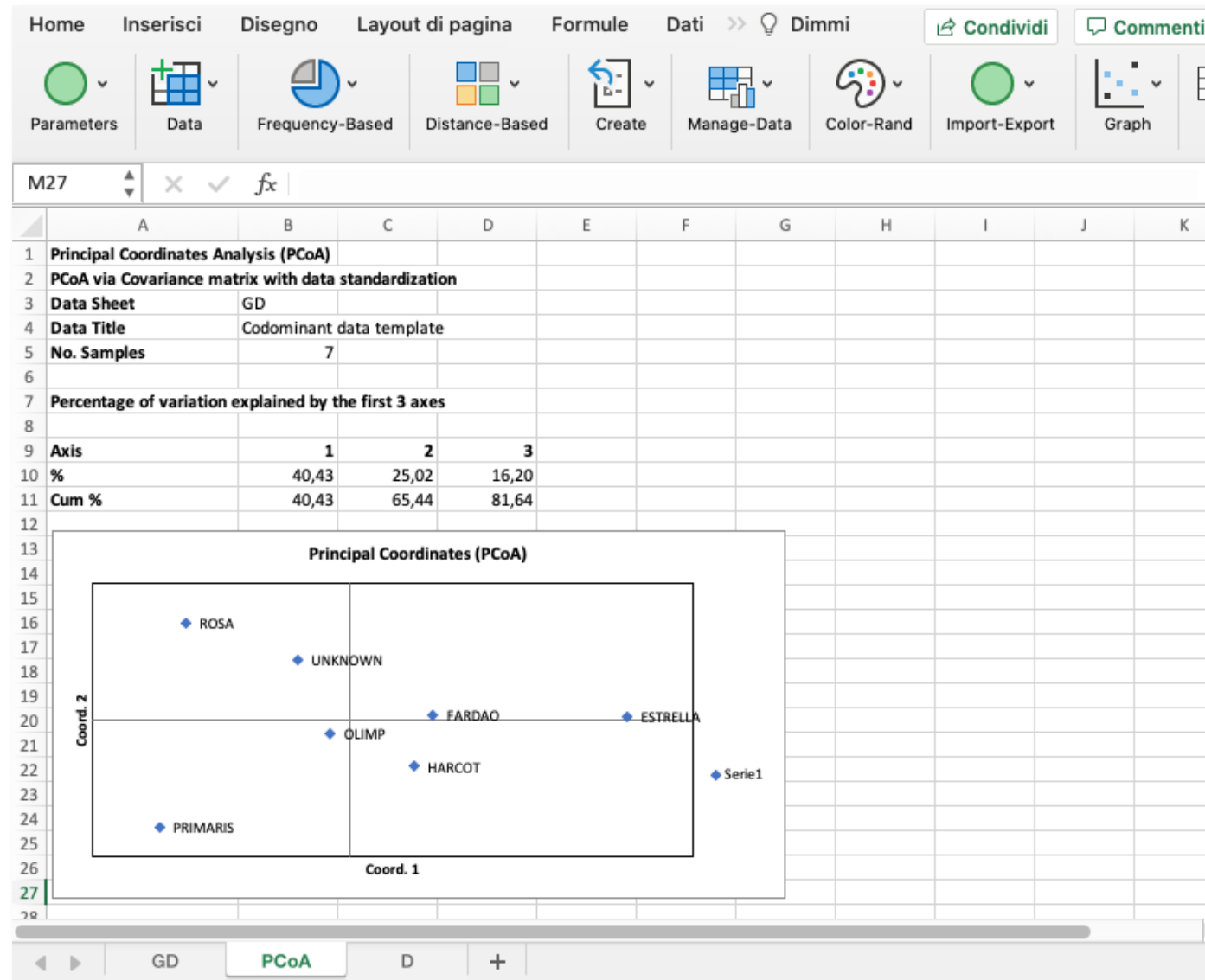
The Excel spreadsheet contains the following data:

	A	B	C	D	E	F	G	H
1	1	7	1	7	1	7		
2	Codominant	D	Codominant	Pop1				
3	ESTRELLA	FARDAO	HARCOT	OLIMP	PRIMARIS	ROSA	UNKNOWN	
4	0							ESTRELLA
5	6	0						FARDAO
6	6	4	0					HARCOT
7	12	5	6	0				OLIMP
8	22	12	8	10	0			PRIMARIS
9	20	10	12	10	14	0		ROSA
10	14	9	9	10	13	10	0	UNKNOWN

A red arrow points to the 'GD' tab at the bottom of the Excel window.

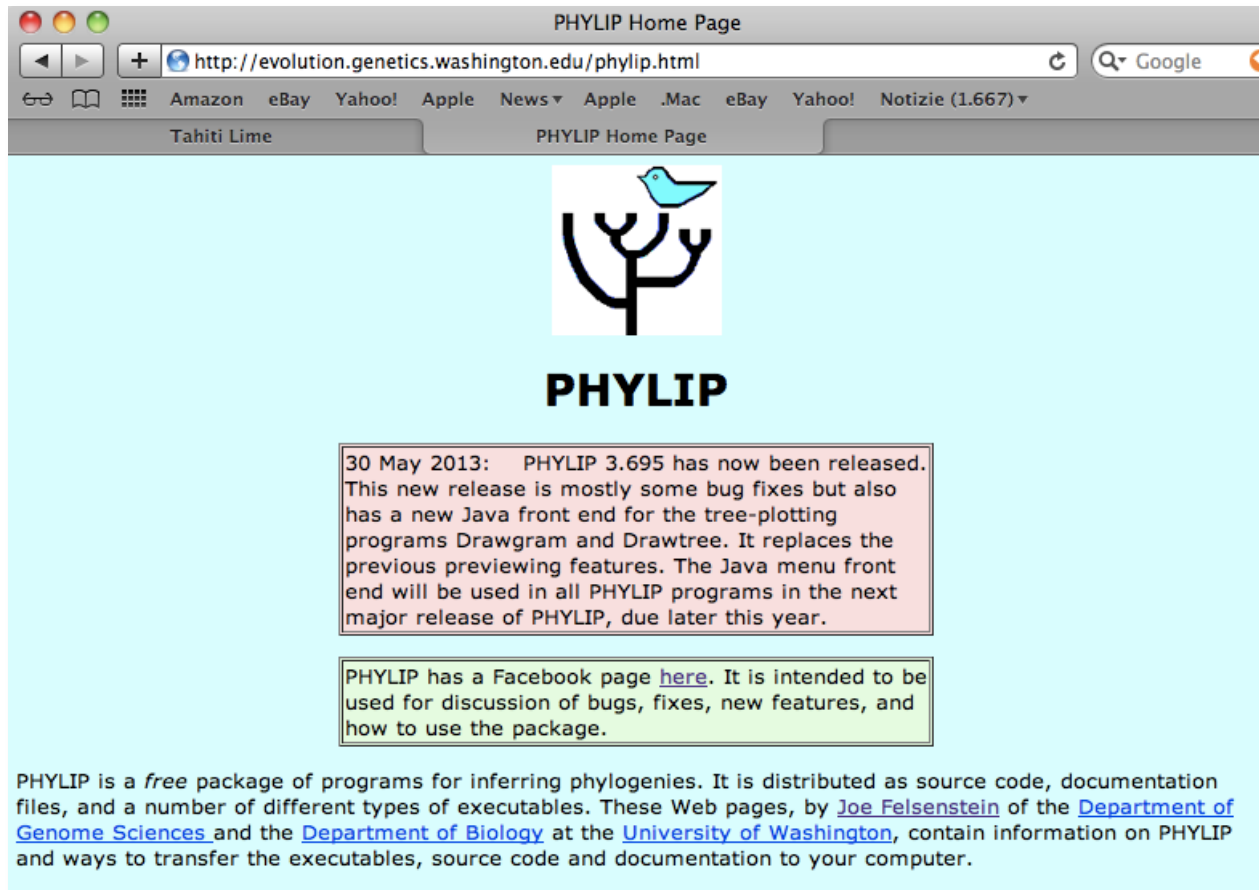
Fourth lab activity (the GenALEx software)

Principal Component Analysis



Fourth lab activity (the PHYLIP package and the relative documentation)

WE 20-12-2023




PHYLIP Home Page

http://evolution.genetics.washington.edu/phylip.html

Amazon eBay Yahoo! Apple News Apple .Mac eBay Yahoo! Notizie (1.667)

Tahiti Lime PHYLIP Home Page



PHYLIP

30 May 2013: PHYLIP 3.695 has now been released. This new release is mostly some bug fixes but also has a new Java front end for the tree-plotting programs Drawgram and Drawtree. It replaces the previous previewing features. The Java menu front end will be used in all PHYLIP programs in the next major release of PHYLIP, due later this year.

PHYLIP has a Facebook page [here](#). It is intended to be used for discussion of bugs, fixes, new features, and how to use the package.

PHYLIP is a *free* package of programs for inferring phylogenies. It is distributed as source code, documentation files, and a number of different types of executables. These Web pages, by [Joe Felsenstein](#) of the [Department of Genome Sciences](#) and the [Department of Biology](#) at the [University of Washington](#), contain information on PHYLIP and ways to transfer the executables, source code and documentation to your computer.

<http://evolution.genetics.washington.edu/phylip.html>

**A primer to phylogenetic analysis using
the PHYLIP package**

Jarno Tuimala - Fifth Edition

<http://koti.mbnet.fi/tuimala/oppaat/phylip2.pdf>

Fourth lab activity

WE 20-12-2023

(the PHYLIP package and the relative documentation)

```
fg — neighbor — neighbor — 80x24
~ — neighbor — neighbor

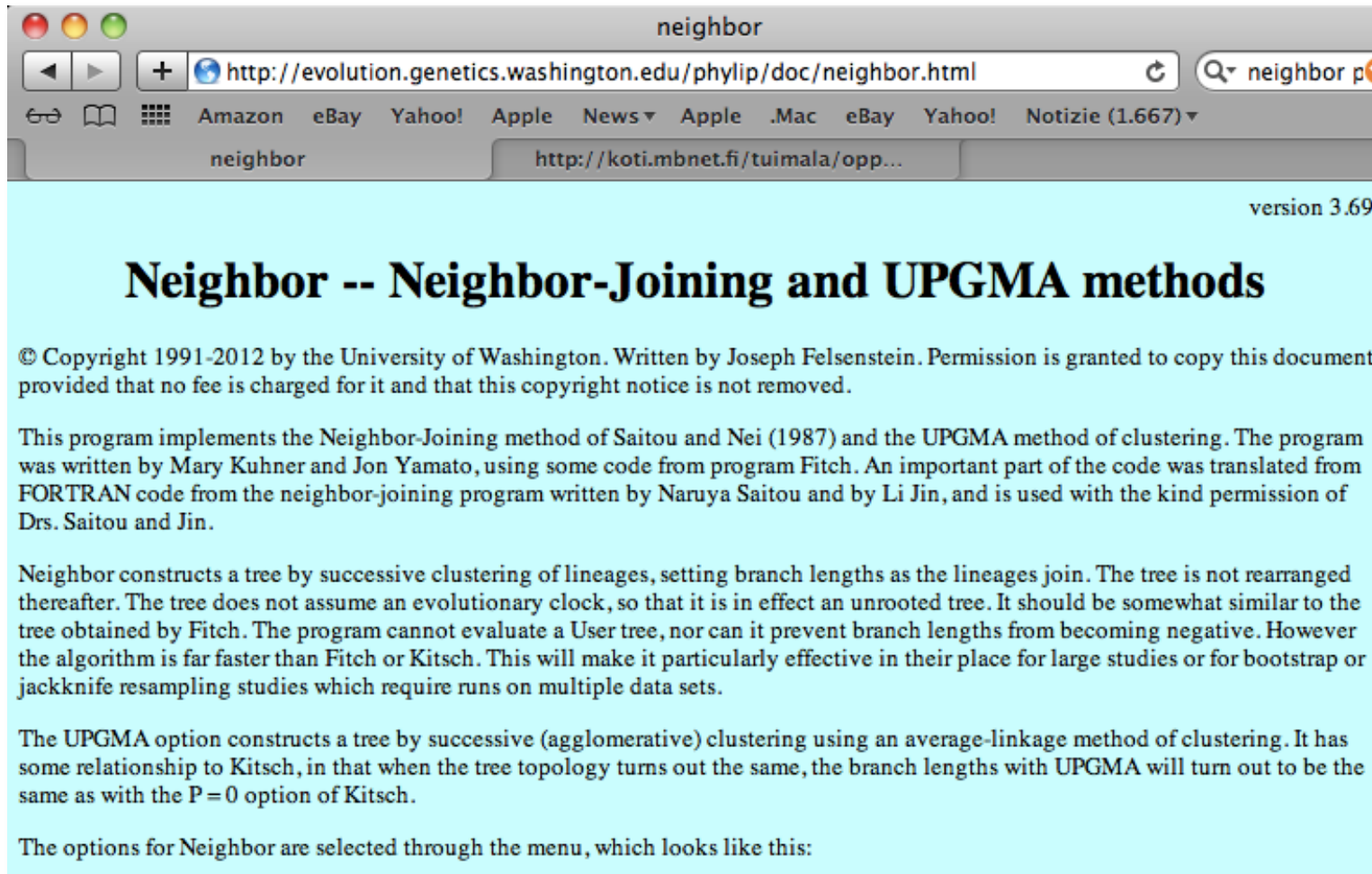
Neighbor-Joining/UPGMA method version 3.69

Settings for this run:
N      Neighbor-joining or UPGMA tree?  UPGMA
L      Lower-triangular data matrix?    Yes
R      Upper-triangular data matrix?    No
S      Subreplicates?                   No
J      Randomize input order of species? No. Use input order
M      Analyze multiple data sets?      No
0      Terminal type (IBM PC, ANSI, none)? ANSI
1      Print out the data at start of run No
2      Print indications of progress of run Yes
3      Print out tree                    Yes
4      Write out trees onto tree file?   Yes

Y to accept these or type the letter for one to change
Y
```

Fourth lab activity (creating the dendrogram through “Neighbor” of the PHYLIP package)

WE 20-12-2023



The screenshot shows a web browser window titled "neighbor" with the address bar containing the URL <http://evolution.genetics.washington.edu/phylip/doc/neighbor.html>. The browser's search bar contains "neighbor p". The page content is on a light blue background and includes the following text:

neighbor version 3.69

Neighbor -- Neighbor-Joining and UPGMA methods

© Copyright 1991-2012 by the University of Washington. Written by Joseph Felsenstein. Permission is granted to copy this document provided that no fee is charged for it and that this copyright notice is not removed.

This program implements the Neighbor-Joining method of Saitou and Nei (1987) and the UPGMA method of clustering. The program was written by Mary Kuhner and Jon Yamato, using some code from program Fitch. An important part of the code was translated from FORTRAN code from the neighbor-joining program written by Naruya Saitou and by Li Jin, and is used with the kind permission of Drs. Saitou and Jin.

Neighbor constructs a tree by successive clustering of lineages, setting branch lengths as the lineages join. The tree is not rearranged thereafter. The tree does not assume an evolutionary clock, so that it is in effect an unrooted tree. It should be somewhat similar to the tree obtained by Fitch. The program cannot evaluate a User tree, nor can it prevent branch lengths from becoming negative. However the algorithm is far faster than Fitch or Kitsch. This will make it particularly effective in their place for large studies or for bootstrap or jackknife resampling studies which require runs on multiple data sets.

The UPGMA option constructs a tree by successive (agglomerative) clustering using an average-linkage method of clustering. It has some relationship to Kitsch, in that when the tree topology turns out the same, the branch lengths with UPGMA will turn out to be the same as with the P = 0 option of Kitsch.

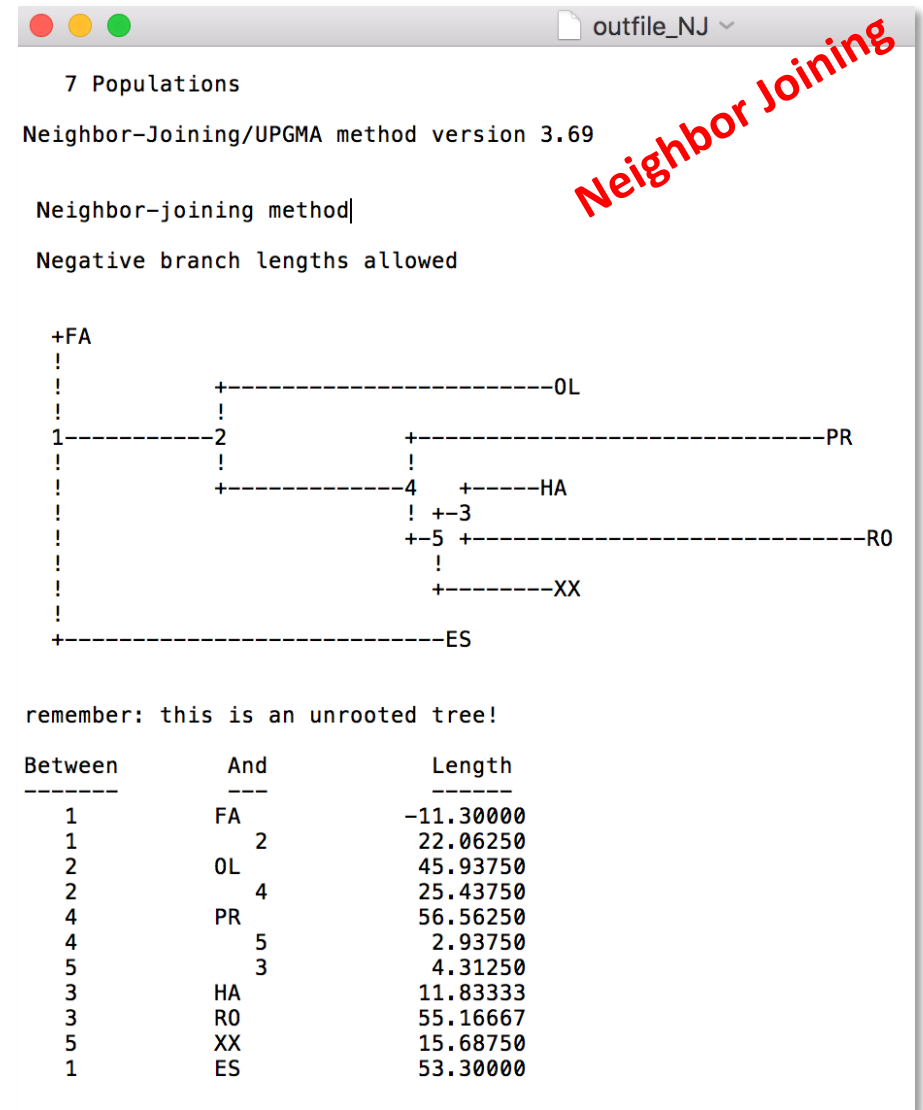
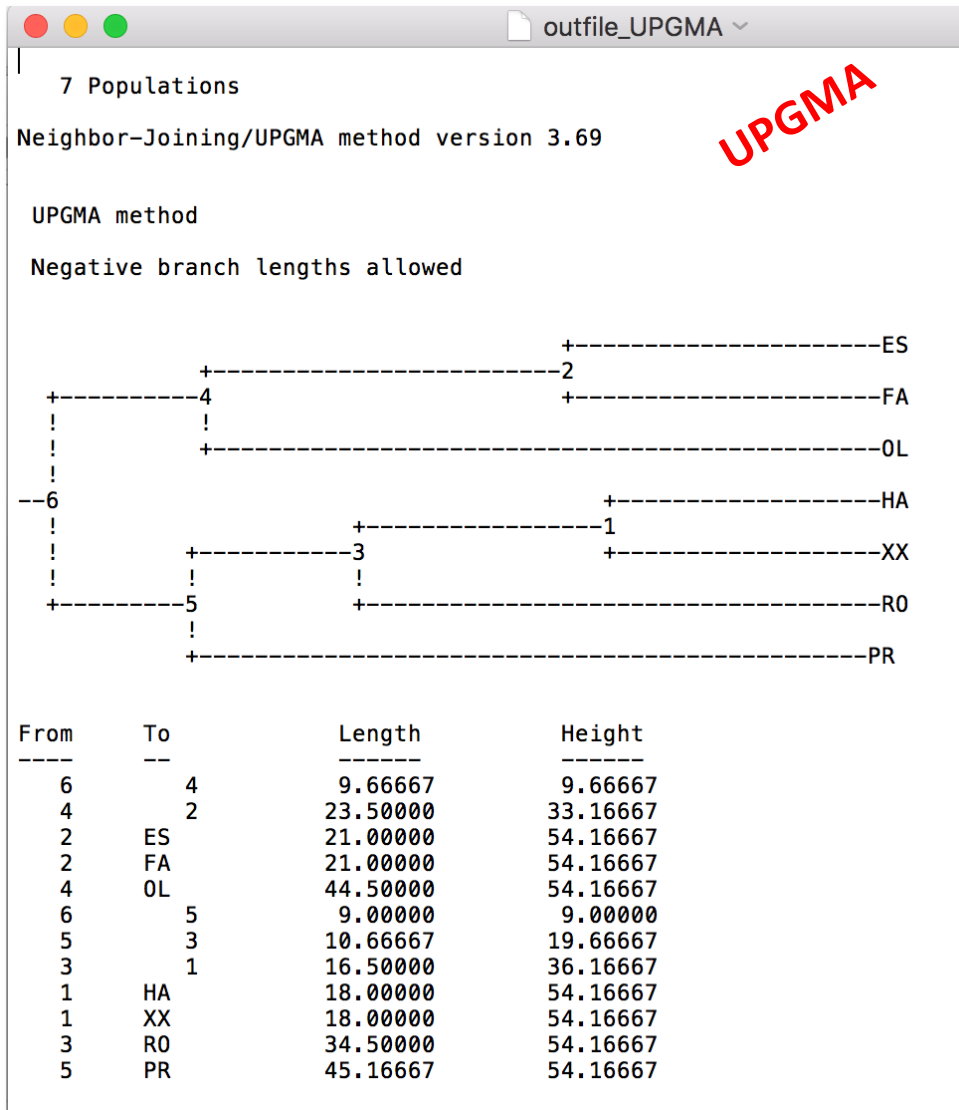
The options for Neighbor are selected through the menu, which looks like this:

<http://evolution.genetics.washington.edu/phylip/doc/neighbor.html>

Fourth lab activity

(creating the dendrogram through “Neighbor” of the PHYLIP package)

WE 20-12-2023



outtree_UPGMA

```
(( (ES:21.00000,FA:21.00000):23.50000,OL:44.50000):9.66667,
(( (HA:18.00000,XX:18.00000):16.50000,R0:34.50000):10.66667,
PR:45.16667):9.00000);
```

outtree_NJ

```
(FA:-11.30000,(OL:45.93750,(PR:56.56250,((HA:11.83333,R0:55.16667):4.31250,
XX:15.68750):2.93750):25.43750):22.06250,ES:53.30000);
```

Fourth lab activity

(identifying the genomic regions of the SSR loci analyzed)

WE 20-12-2023

Using the sequences of Forward and Reverse primers for each SSR locus it is possible to identify their location and sequence in the reference genomes of *Prunus* species (i.e. Peach, Apricot, *Prunus mume*).

A comparison can be done with the SSR allele length of the genomic sequence(s) in the reference sequence.

Hint: use the *Phytozome* database that has proven effective at locating short primer sequences.

