

EMBO Population Genomics Practical 2

Andrea Manica

qpGraph

To fit admixture graphs, we will use qpGraph, which is part of Admixtools. We will use the same dataset in Eigenstrat format as this morning. There is not R interface, so we will have to run the programs directly from the bash shell. Make sure that you connect to the server using ssh -X (so that we can use the X server to visualise pdf and edit files), and go to:

```
cd /data/andrea/practical
```

qpGraph needs two files to run, a .par file that provides information on where there data are to be found, and a .graph file that provides the structure of the graph. In the practical directory, you will find a file called qpgraph.par. You can inspect its content by typing:

```
cat qpgraph.par
```

We will start by fitting a simple graph as an example, using base.graph

You can inspect its content with:

```
cat base.graph
```

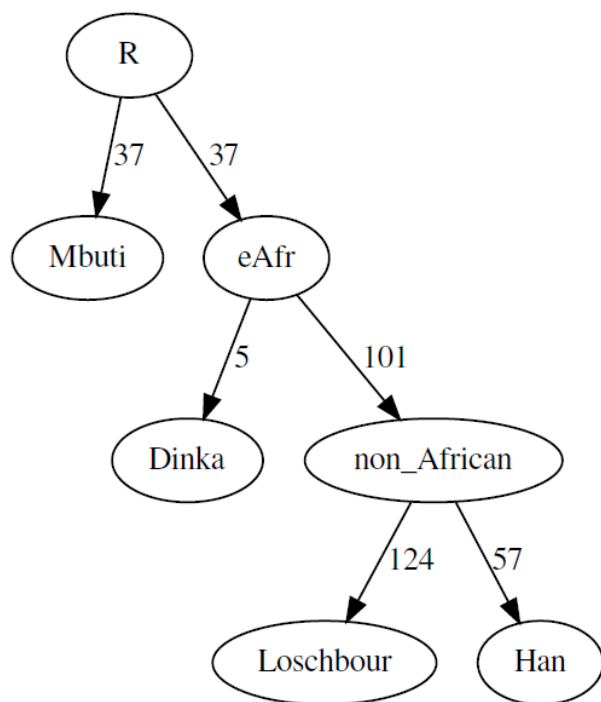
Now let's fit the graph

```
qpGraph -p qpgraph.par -g base.graph -d base.dot -o base.graphout | tee base.qpgraph.log
```

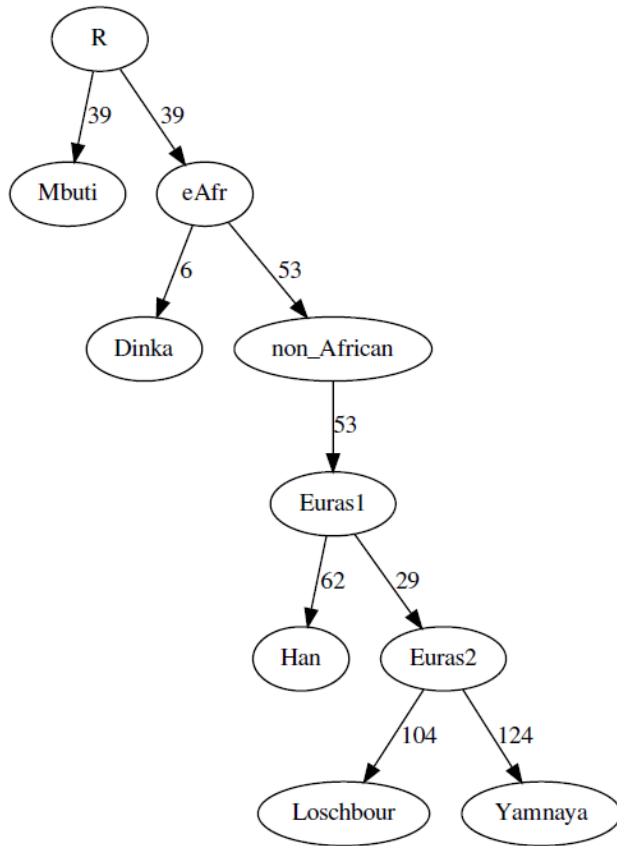
And visualise the admixture graph in a pdf

```
dot -Tpdf base.dot > base.pdf  
evince base.pdf
```

You should see something like this:



We will now expand the graph to match the one on the figure below.



We want to create a noadm.graph by modifying base.graph. First, make a copy of base.graph:

```
cp base.graph noadm.graph
```

Now, we can edit the file with a text editor. If you are connecting through ssh -X, you can simply type:

```
gedit noadm.graph
```

Otherwise, you can download the file, edit it on your machine, and reupload it via sftp or scp.

If you mispecify the graph, qpGraph does not provide you with a very informative error message. So, when you modify a graph, it is best to inspect it with qpreroot:

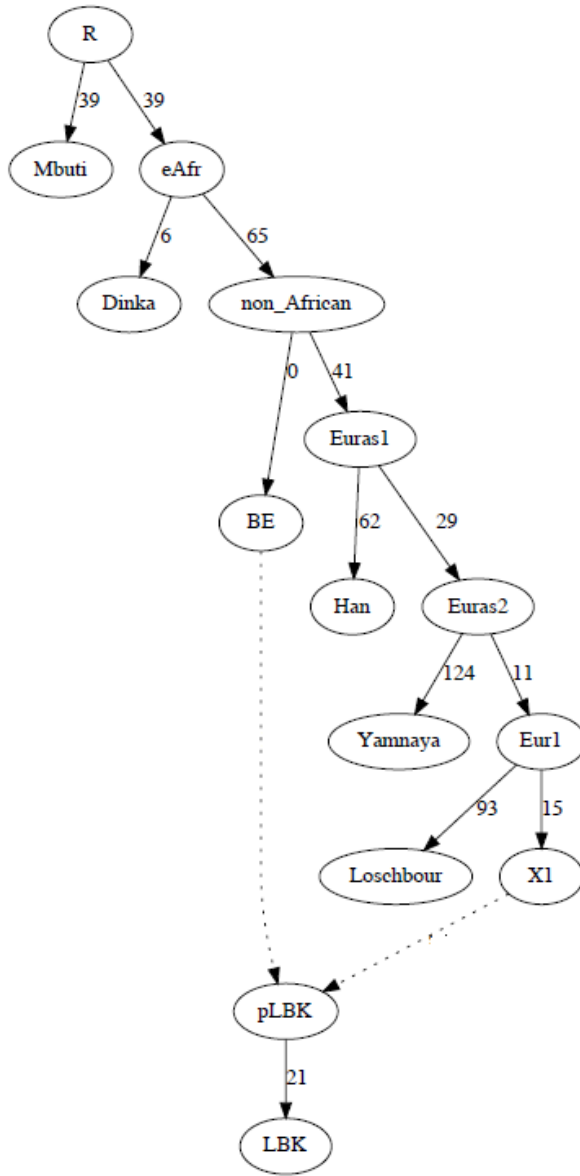
```
qpreroot -g noadm.graph -d noadm.dot
dot -Tpdf noadm.dot > noadm.qpgraph.pdf
evince noadm.qpgraph.pdf
```

If the graph looks OK, we can run qpGraph, and visualise the output:

```
qpGraph -p qpgraph.par -g noadm.graph -d noadm.dot -o noadm.graphout | tee noadm.qpgraph.log
dot -Tpdf noadm.dot > noadm.qpgraph.pdf
evince noadm.qpgraph.pdf
```

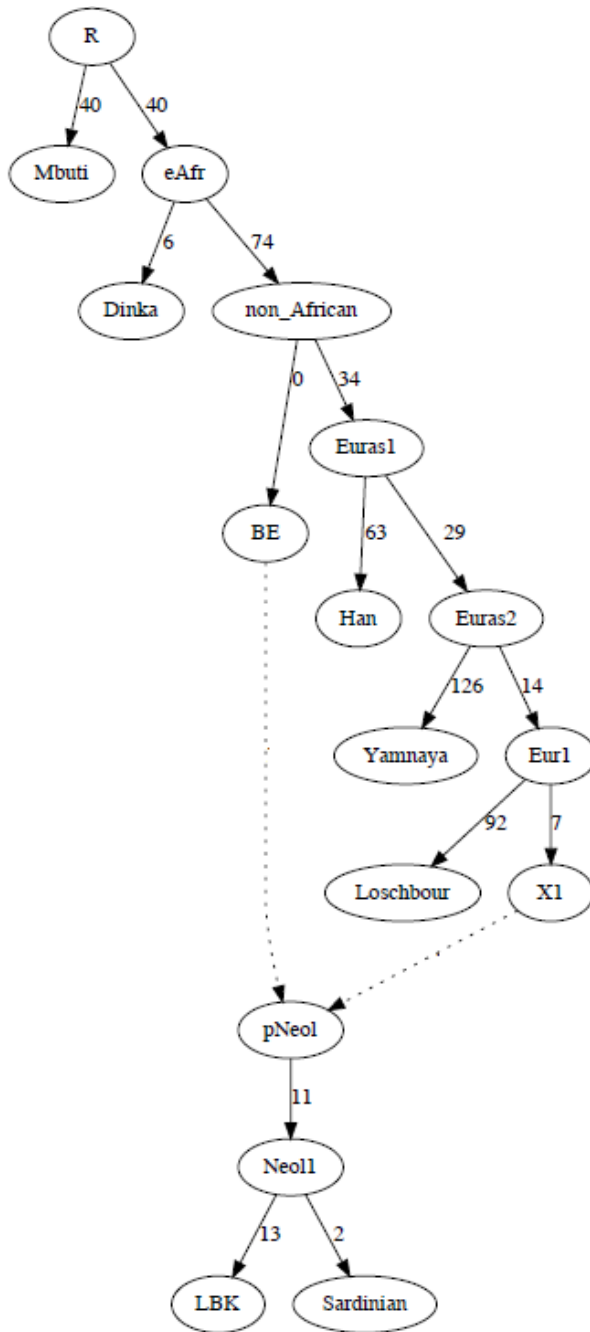
Does the graph fit the data?

Now we will add a an admixed population (LBK, which is an Neolithic farmer), giving an admixture starting levels of 80 Eurasian and 20 BE. Use the figure below as a guide on how to structure the graph.



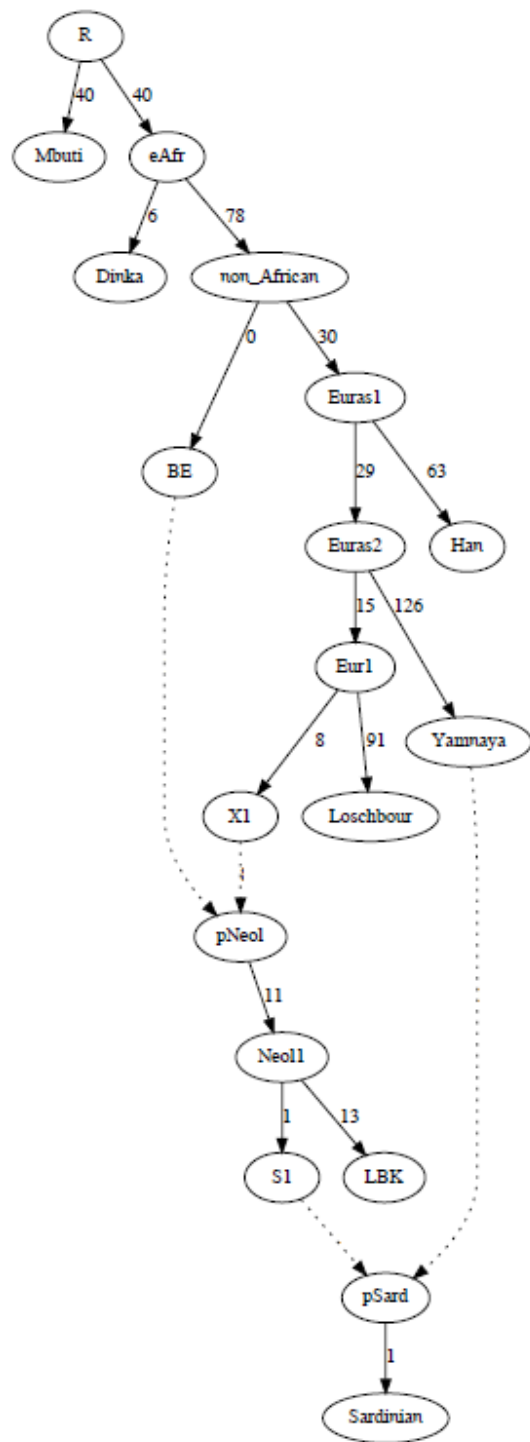
Does the graph fit the observed f-statistics?

Now add Sardinian (modern inhabitants of the Italian island) as an unadmixed population to graph, as shown below



Does the graph fit?

And finally, try to add admixture from Yamnaya into Sardinian (give it a small starting input of 5 vs 95 from the other source)



Has the fit improved? How much Yamnaya admixture does the fit show?