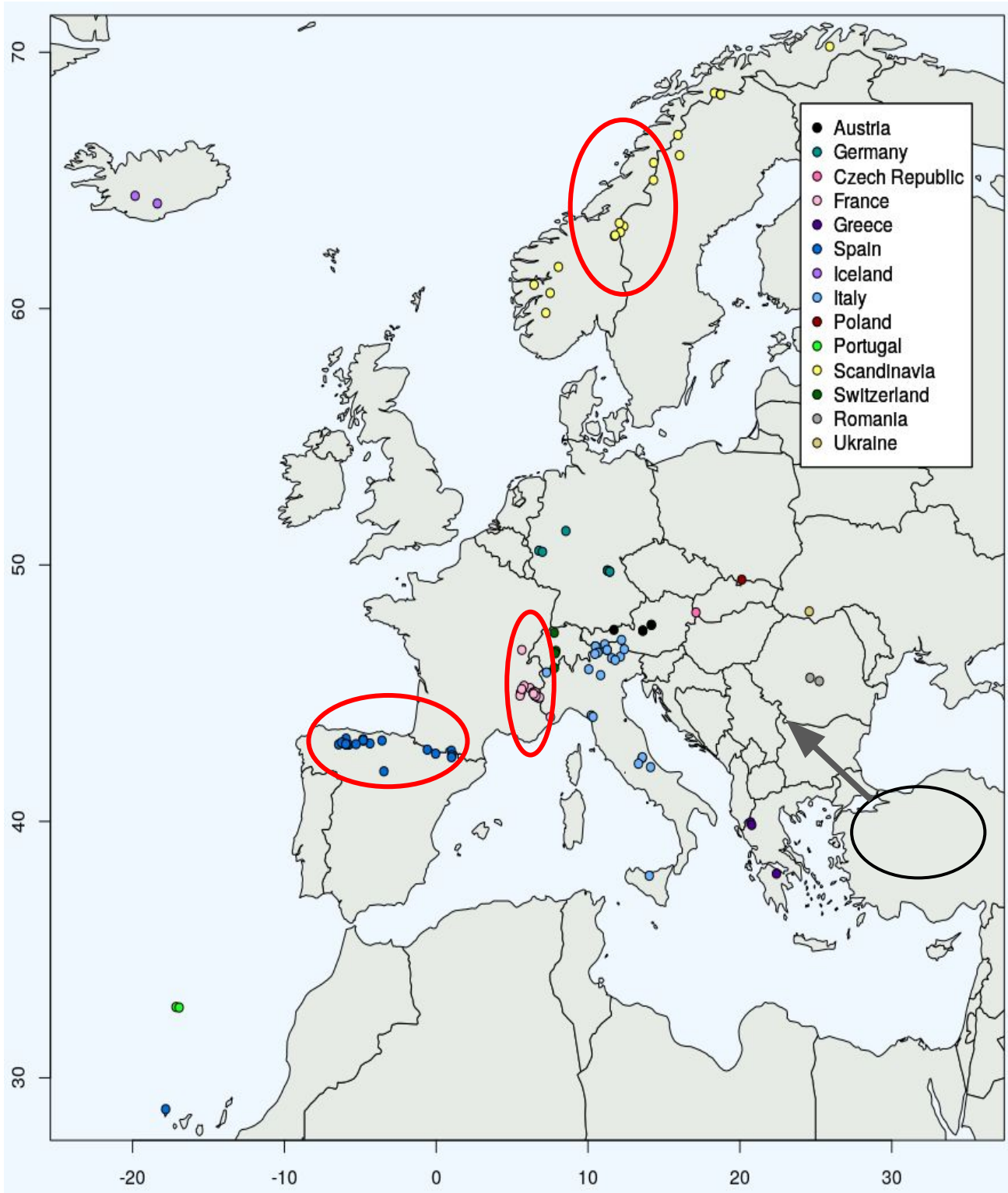# Demographic History of European *Arabis alpina*

Mehak Sharma, Dr. Andrea Fulgione

Background

Arabis alpina is:

- A herbaceous Arctic-alpine plant of Europe and Northern Africa.
- An important model organism to study the evolution of perenniality as a life history trait in extreme environments.
- *Arabis a.* emigrated from Asia minor 0.37-1.39 mya ago. [1]

Topology Test-Fastsimcoal2

Fastsimcoal2 [4] is a site frequency spectrum-based coalescent simulator that was used to infer the demographic history. It calculates the log-likelihood of the best set of parameters for a demographic model given allele count data(Site Frequency Spectrum). An initial topology test was performed by comparing the fit of the three possible unrooted trees for the populations studied(Nor6, Gal, E4). The fit was compared using the log-likelihood, AIC, and AIC weights.
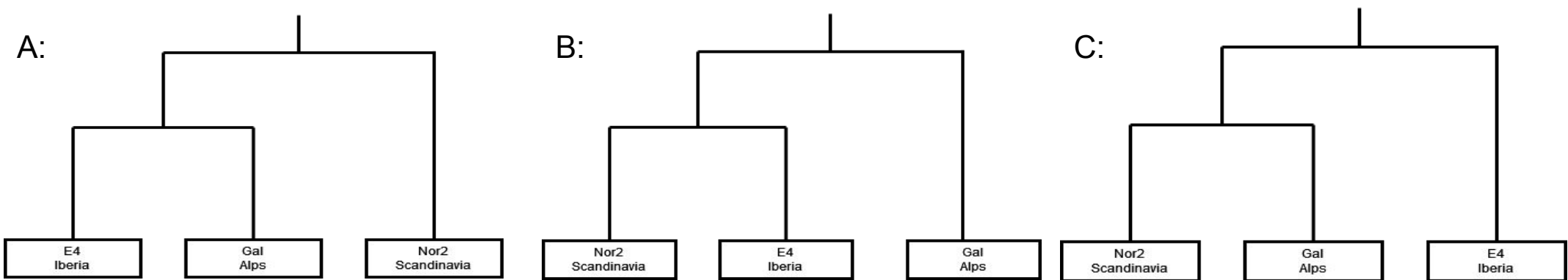

Model	Log-likelihood	AIC	Aic Weights
A:((Iberia,Alps),Scandinavia)	-26811004.301	53622018.602	~0
B:((Scandinavia,Iberia),Alps)	-26988266.28	53976542.56	~0
C:((Scandinavia,Alps),Iberia)	-26466992.831	52933995.662	1

From the values of log-likelihood, AIC, and the AIC weights, most support can be seen for model C which shows a **more recent split between Scandinavia and the Alps and an earlier split with Iberia.**

Topology Test-Twisst

Topology weighting for iterative sampling of subtrees was also used to infer the topology. The contribution of each topology in the full tree across the whole genome in segments is calculated. Since unrooted trees are compared, *Arabis montbretiana* was included in the analysis as an outgroup.
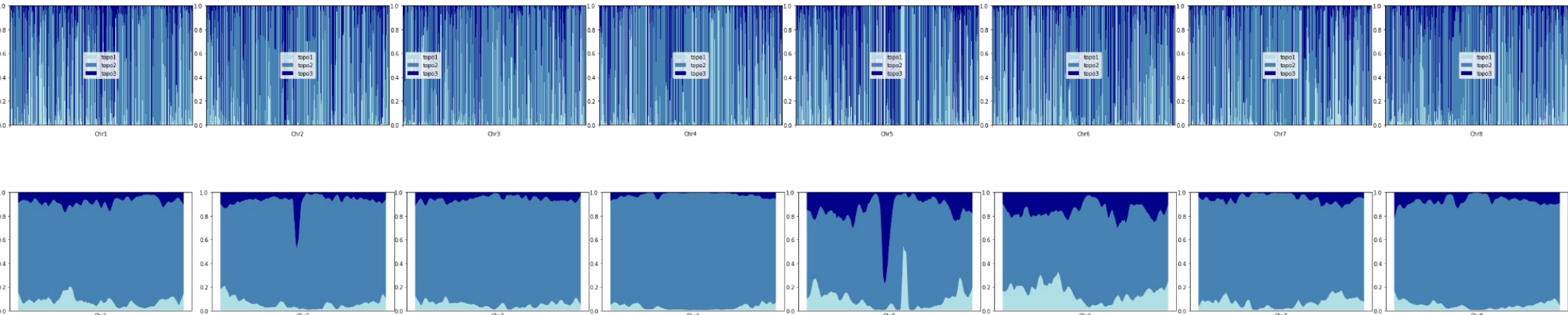
|							
T1: 0.20	0.20	0.19	0.18	0.23	0.23	0.20	0.19
T2: 0.58	0.60	0.62	0.65	0.51	0.54	0.62	0.61
T3: 0.20	0.20	0.19	0.17	0.26	0.23	0.18	0.20

The most likely scenario is topology 2 ((Scandinavia, Alps)), Iberia, outgroup) which symbolizes that the Scandinavian and Alpine populations are most closely related and likely to have a more recent split than that with the Iberian populations.

#topo1 ((Scandinavia,Iberia),Alps,mont)
#topo2 ((Scandinavia,Alps),Iberia,mont)
#topo3 ((Scandinavia,,mont),Iberia,Alps)

Demographic Detail

The demographic detail was inferred in a stepwise manner where the complexity of the model was slowly increased by iteratively adding the demographic process, which decreased the AIC the most in each step.

The starting base model was taken as a simple split model with constant population sizes. To the base model, several demographic processes were added, and the combination that decreased the AIC the most was selected to act as a base model for the next step.
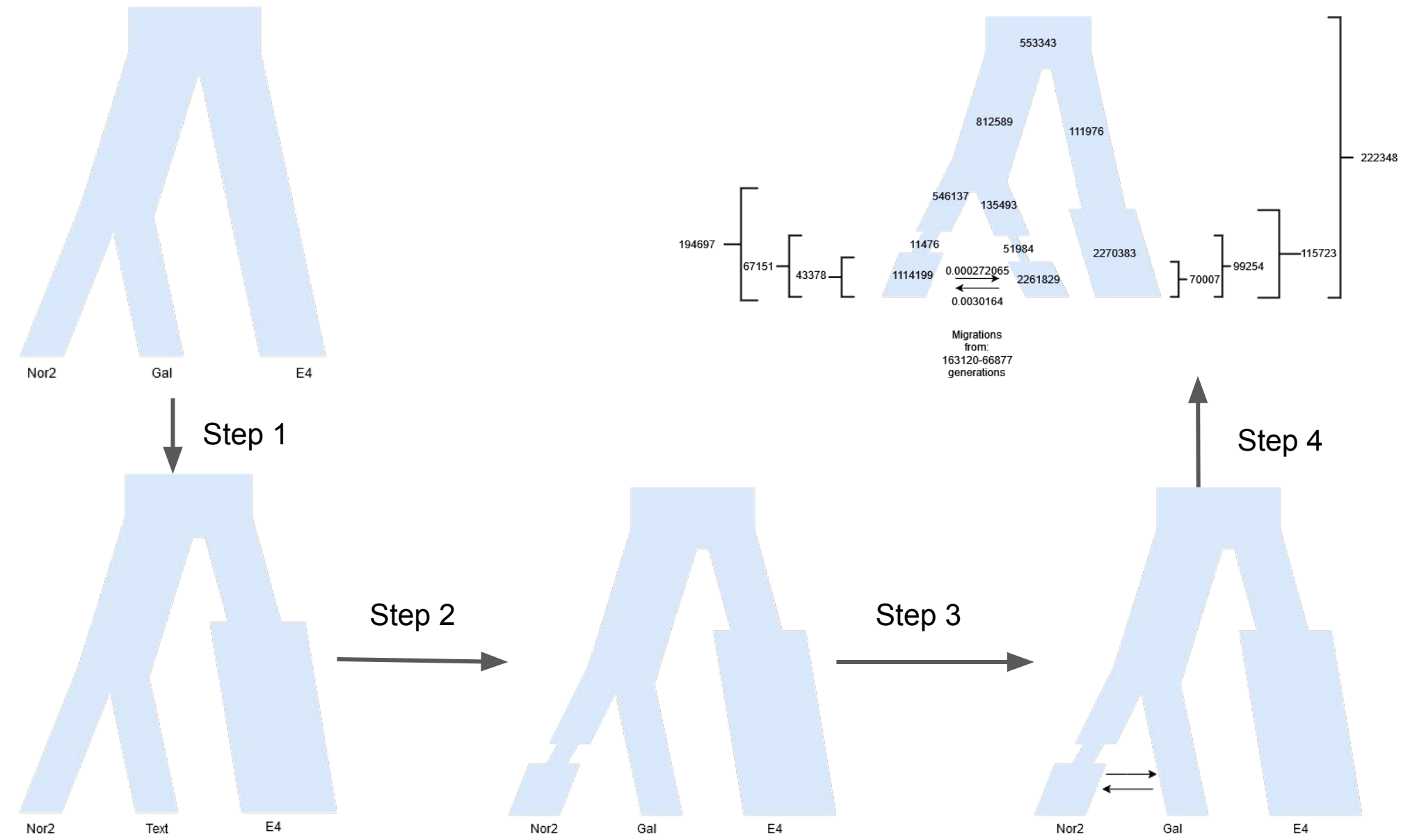
Six Population Model

Once the topology was finalized, we added one more population form each of the three major regions along with the ones already been used in the three population models to construct a six-population model.
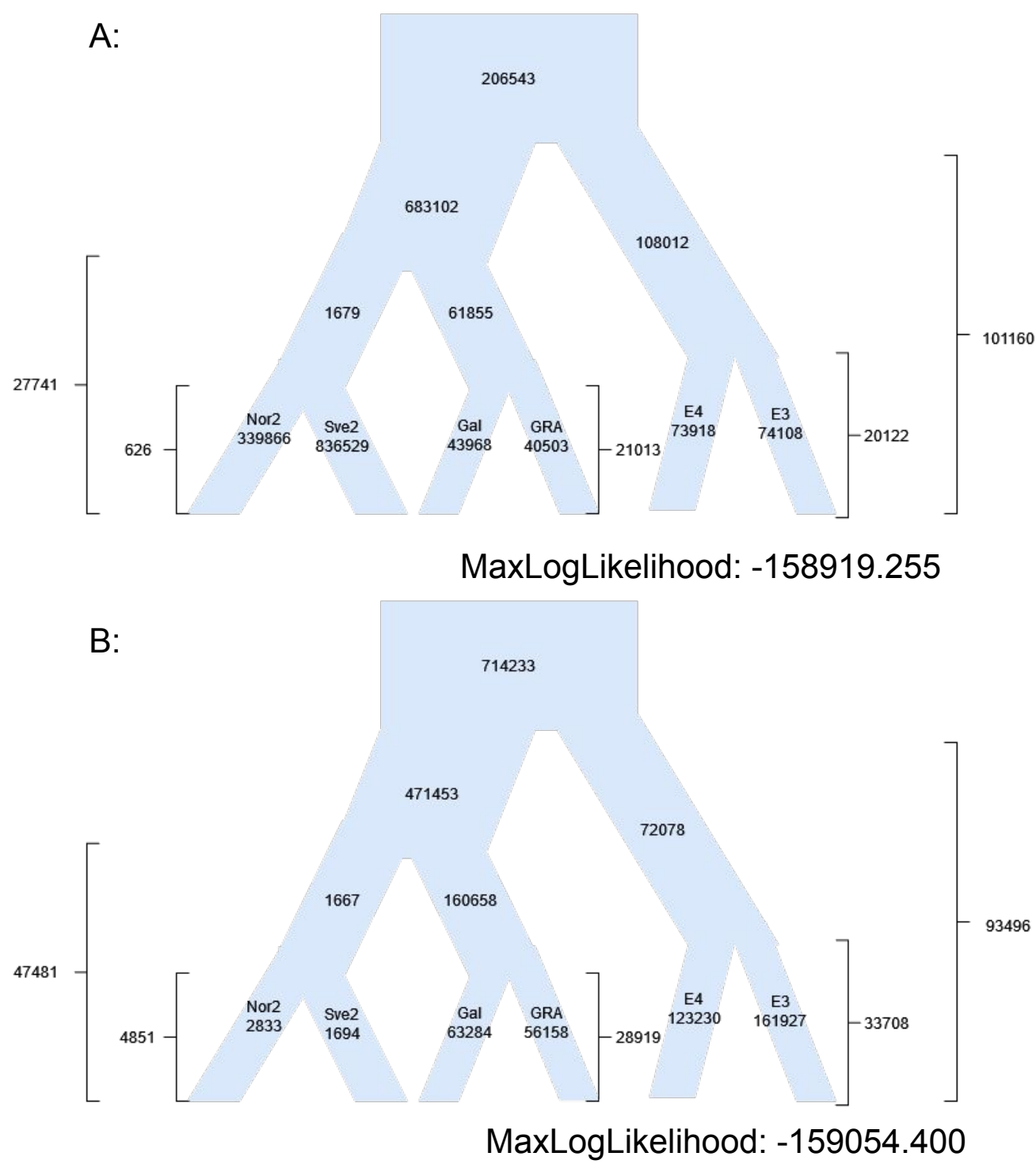
Two major groups of the optimized parameter sets were observed, the main difference being the effective population sizes of the Scandinavian populations and the split time between them.

This trend might be due to the fact that there are fewer SNPs in the Scandinavian populations, and fastsimcoal2 attributes equal weight to each SNP.

Thus even if the parameter set B fits the sfs from Scandinavian populations much better, set A might show a higher likelihood because there are more SNPs in other populations, and it fits them slightly better.

Conclusion

Within the European populations, there is a more recent split between the Scandinavian and Alpine populations and an earlier split with the Iberian populations. A model constructed by taking one population per major group might be useful to infer the possible demographic processes and add complexity in terms of the number of parameters per group, but to better resolve the parameters, genetic variation at various levels of structure within groups is required. Thus here, a six-population model with within-group resolution results in better demographic parameters. However, the complex parameter space and weighting of variants need to be understood to draw conclude realistic scenarios.

Acknowledgements

References

1) Koch et al., 2006 ,Molecular Ecology, doi: 10.1111/j.1365-294X.2005.02848.x.
2) Woetzel et al., 2021, Molecular Ecology Resources, doi: 10.1111/1755-0998.13490.
3) Martin and Van Belleghem, 2017, doi: 10.1534/genetics.116.194720
4) Excofffier et al., 2021, Bioinformatics, doi: 10.1093/bioinformatics/btab468.

MAX PLANCK INSTITUTE FOR PLANT BREEDING RESEARCH