

Figure S1 Principal components analysis of simulated read count data for first three components. Colors indicating pool are the same as in Figure 2.

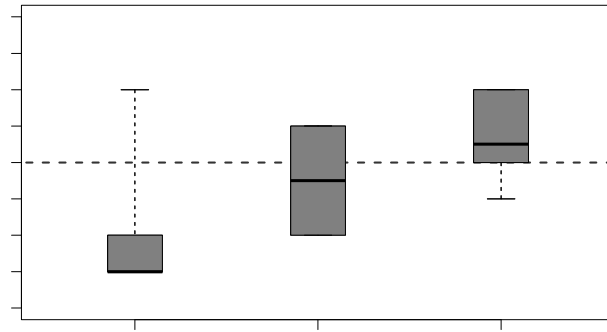


Figure S2 Number of inferred lineages made by empirical Bayes factor analysis for three different qualities of data. Ten simulations included within each group.

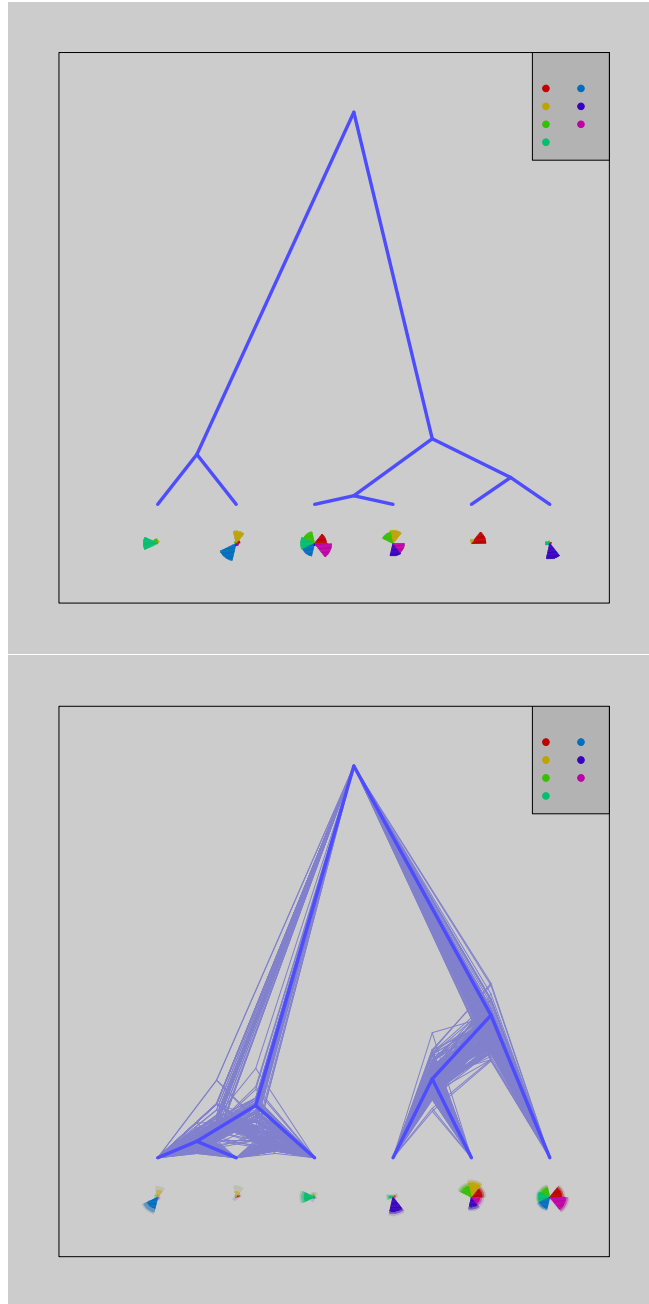


Figure S3 Example of a simulation run with 'migrating' pool proportions. Simulated model is above and inferred model below.

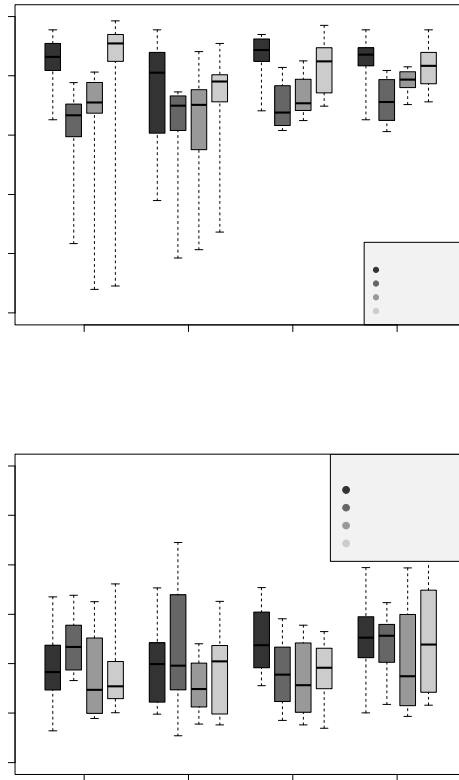


Figure S4 Performance of lineage model inference by mixture parameter by fraction of concordant SNPs (top) and mean absolute deviation in pool proportion (bottom). Mixture parameter is proportional to $\frac{1}{\psi}$.

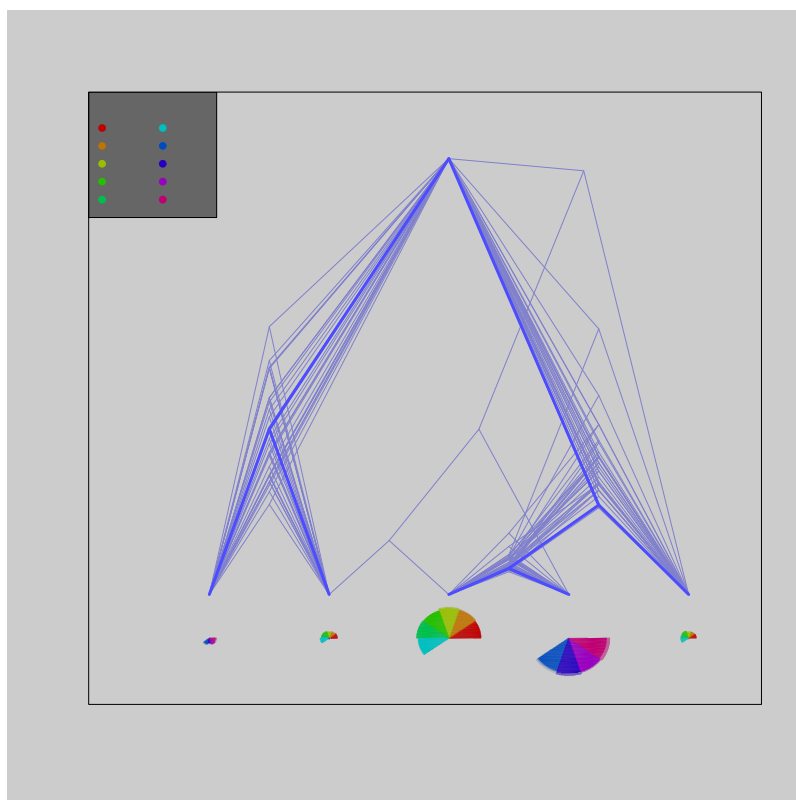


Figure S5 Inferred model for *Neisseria* data set.