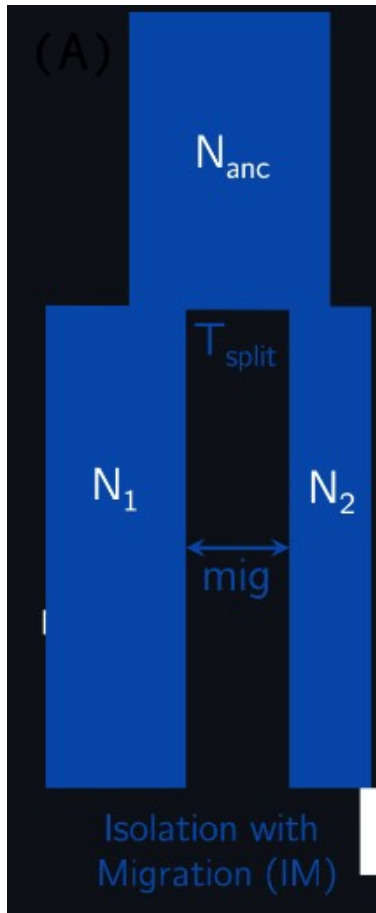


Practical – Day 3

We are analysing genomic data from two populations of *Anopheles gambiae* from two different locations. We have 50 diploid samples per location. We wish to estimate their demographic history follow this Isolation-with-Migration (IM) model.



Task 1 (morning): based on the following information, decide on suitable prior probabilities and perform random samples from them. You can use R or python but you need to check that you are covering the whole range of all prior distributions.

T_{split} : we know that these populations have been separated by at least 100 generations and a major environmental change between these two locations happened around 8,000 generations ago.

N_1/N_2 : we know from capture-recapture data that population1 is approximately 30 times larger than population 2; we also don't have a clear intuition of each magnitude but from previous findings we expect values for population 1 between 50,000 and 200,000.

mig: we expect either complete isolation after the split (rate=0) or pervasive migration (rate=0.1)

N_{anc} : we know this value with high confidence to be around 7,000,000.

Task 2 (afternoon): Assume that the observed summary stats are in `mosquito-observed.csv`

Fst,dxy,segsites1,segsites2,pi1,pi2,tajima1,tajima2

0.21341428024826015,0.09782663999999873,0.3797,0.1013,0.09135567346938553,0.03547493877551051,0.284715467675643,2.0788208139138997

Estimate the parameters of the IM model using ABC. As first trial, assume that we know all parameters but T_{split} . In particular, assume that there is no migration and $N_1=150,000$ and $N_2=5000$. Produce the posterior distribution of T_{split} .

Task 3 (afternoon): Attempt to jointly infer T_{split} , N_1 and N_2 .