FREGENE datasets: 'Population B'

SUMMARY OF MODELLING ASSUMPTIONS USED TO SIMULATE THE SUBDIVIDED POPULATION.

Population B mimics the evolution of 10.5K diploid individuals along 200K generations. Each sequence is 1 Mb long. The population is subdivided into three equal-sized subpopulations (each containing 7K sequences). Migration between subpopulations is considered as isotropic and the common migration rate per sequence is set such that the Fst, measuring the genetic distance between subpopulation, equals 10%. Two simulations are available: one neutral (see results in `Output_neutral`) and one with selection (see results in `Output_selected`). Scripts used to generate these simulations are also provided (`fregene_POPNB_neutral.sh` and `fregene_POPNB _sel.sh` respectively). Modelling assumptions are recorded in the input files (in `Input` directory) and summarized below. Finally some `R` scripts that generate figures describing these simulations can be found in `R_Scripts` together with the corresponding figures (in `Figures`).

General parameters	
Chromosome Length	$1\ Mb$
# Generations	$200,000$
# Sequences	$21,000$
Per-site mutation rate	2.3×10^{-8}
Recombination model	
Per site crossover rate:	1.1×10^{-8}
Per site Gene conversion rate:	4.5×10^{-9}
Proportion of recombination events occurring in hotspots	80%
Hotspot length:	$2.0kb$
Gene conversion length:	$0.5kb$
Mean distance between hotspots:	$8.5kb$
Selection parameters (if applicable)	
Prop. of sites under selection:	5×10^{-4}
Proportion of selected sites locally under selection:	0.5
Mean # generations before selected sites are switched off	$50,000$
selection coefficient:	$s \sim 0.1 \times \mathcal{N}(0.005, 0.05^2)$ $+0.9 \times \mathcal{N}(-0.01, 0.005^2)$
dominance coefficient:	$h \sim 0.8 \times \mathcal{N}(0.5, 0.2)$ $+0.3 \times \mathcal{N}(1.2, 0.2^2)$