

Bacmeta:
Simulation for genomic evolution
in bacterial metapopulations
Supplementary material

Aleksi Sipola, Pekka Marttinen and Jukka Corander

1 Parameter input-file: *simu.input*

Bacmeta reads parameters from a file located in the simulator folder. By default the file named *simu.input* is read, but custom file names of form *simu * .input* can be used with command line options, where * is some alphanumeric string.

Parameter settings of *simu.input* file for the example case, excluding documentation settings:

Table S1: Parameter inputs for example case

101	Output file name modifier, alphanumeric. Former files with same modifier will be overwritten!
0	Set seed or give zero to have random seed from time.
20000	Number of generations to run
0.001	Compute and save mean distance summary every 1000 generation (1) or not (0)
1000	Length of locus
300	Number of loci in bacteria
1000	Number of bacteria in population
3	Number of populations in metapopulation
1	Using migration rates from "migration.input"-file (1) or universal value given below (0)

0.05	Mean migration rate scaler
0.05	Migration probability
0.001	Mean microepidemic amount scaler
0.001	Mean microepidemic size scaler
2e-7	Mutation rate per nucleotide per population per generation
4	Recombination rate in relation to mutations
0	Recombination mean length. Give (0) if whole genes are to be recombined
18	Recombination acceptance parameter for similarity test
0	Gather recombination site metadata (1) or not (0)
0.00	Insertion rate in relation to mutations
1.7	Insertion length parameter for Zipf distribution
0.00	Deletion rate in relation to mutations
1.7	Deletion length parameter for Zipf distribution
0.02	Maximum indel length as a proportion of the gene length
0.31	Proportion of base A
0.30	Proportion of base T
0.18	Proportion of base G
0.21	Proportion of base C
0.127	$P(a \rightarrow t)$ i.e. Probability of mutation from base A to T
0.655	$P(a \rightarrow g)$
0.218	$P(a \rightarrow c)$
0.278	$P(t \rightarrow a)$
0.165	$P(t \rightarrow g)$
0.557	$P(t \rightarrow c)$
0.686	$P(g \rightarrow a)$
0.279	$P(g \rightarrow t)$
0.035	$P(g \rightarrow c)$
0.261	$P(c \rightarrow a)$
0.622	$P(c \rightarrow t)$
0.117	$P(c \rightarrow g)$
1	Randomly select bacteria for new generation (1) or keep as is (0)
1	Keep order of event types same (0) or randomize each generation (1)

2 Migration rate input-file: migration.input

If non-homogenous migration rate network is to be used, it can be given as adjacency matrix for directed graph, so that edge weights represent migration rates. In essence, the input file starts with tab-separated migration rate values forming a matrix, where value in column j of row i defines migration rate from population i to population j , as a proportion of population size. By default, Bacmeta reads file *migration.input*, but by using command line options, custom file names of the form *migration*.input* can be used too.

Migration rate matrix of *migration.input* file that was used for example cases with low λ value:

Table S2: Migration rate inputs

0	0	0.005
0.005	0	0.01
0	0.01	0