The config file

The config file consists of whitespace-separated pairs of parameters and their values, given on separate lines. Lines that begin with "#" are comments and ignored by the program.

The structure of the config file should be as in the example:

```
ALPHABET  ACGT
TREE_FILE simple.tre
ROOT_SEQUENCE_FILE rootseq.txt
PRINT_LANDSCAPE_INFO true
CONSTANT_RATE true
SCALE_LANDSCAPE_CHANGE_TO_SUBSTITUTION_RATE true
NUM_LANDSCAPES 2
[END GLOBAL PARAMETERS]


[LANDSCAPE 1]
LENGTH 4
INITIAL_FITNESS file
FITNESS_FILE fitness_20aa_1_0.1.txt
NEW_FITNESS_RULE shuffle
LANDSCAPE_CHANGE_TIMING specified_branch_and_time
CHANGE_BRANCH_AND_TIME_FILE change_branch_time1.txt
[LANDSCAPE 2]
LENGTH 2
INITIAL_FITNESS file
FITNESS_FILE fitness_20aa_1_0.1.txt
NEW_FITNESS_RULE shuffle
LANDSCAPE_CHANGE_TIMING specified_branch_and_time
CHANGE_BRANCH_AND_TIME_FILE change_branch_time2.txt
```

If there is only one set of landscape rules, the lines in square brackets can be omitted.

The config file contains the following fields. Default values (if such exist) are shown in **bold** in the Example(s) column. If an option does not have default values, its values must be specified, either at all times, or depending on the values of other parameters (the latter case is mentioned in the Description column).

Parameter	Description	Example(s)	Global or landscape-specific
Basic global parameters			
ALPHABET	String that specifies the alphabet that your sequence uses (no spaces)	ARNDCEQGHILKMFPSTWYV	global
NUM_INSTANCES	The number of independent instances of landscape rules that are going to be simulated in parallel. The length of the sequence that evolves according to this landscape is given in the LENGTH parameter. Mutually exclusive with NUM_LANDSCAPES	**1** (default)	global
NUM_LANDSCAPES	The number of unrelated landscapes that are going to be simulated in parallel. The length of the sequence that evolves according to a landscape is given in its LENGTH parameter. Mutually exclusive with NUM_INSTANCES	**1** (default)	global

NUM_THREADS	The number of threads that are going to be used by the simulator. It should be **no greater than NUM_RUNS,** as there currently is no functionality for multithreading a single run (simulation of a single landscape)	**1** (default)	global
TREE_FILE	Absolute or relative path to the file containing the phylogenetic tree in Newick format	/path/to/tree/file.tre	global
ROOT_SEQUENCE_FILE	The file containing the sequence at the root. If not set, the sequence is generated randomly according to the landscape at root.	rootseq.txt	global
MUTATION_RATE_MATRIX_FILE	The file containing the |ALPHABET|x|ALPHABET| mutation rate matrix. If not set, all entries are assumed to be 1.	mutationrate.txt	global
Additional I/O options			
PRINT_LANDSCAPE_INFO	If true, then information about every landscape change time and the corresponding new fitness vectors are printed to files changetimes.merged.fasta and fitnesses.merged.fasta, respectively. WARNING: this can **greatly** increase your memory usage and execution time, so set it to true only if you are sure you need the information (or if you are using small datasets)	**false** (default)	global
Landscape-specific fields			
LENGTH	Length of a sequence that uses a particular landscape, aka the number of characters that evolve according to the same landscape history. If NUM_LANDSCAPES > 1, then, this is the length of the partition(subsequence) that is governed by the current landscape	1 (for single-site evolution simulation) 100 (for simulating a longer sequence that shares a landscape	landscape
INITIAL_FITNESS	How is the initial fitness specified? file – if the fitness is given in a file (see below for file description). If this is set, the FITNESS_FILE must be supplied lognorm – if each character's fitness is randomly generated from lognormal distribution with mu = 0 and sigma = SIGMA gamma – if each character's fitness is randomly generated from the gamma distribution with GAMMA_ALPHA and GAMMA_BETA	file lognorm gamma	landscape
FITNESS_FILE	**Required if INITIAL_FITNESS is file, ignored otherwise**. This is the name of the file containing the fitness vector as a space-separated list of numbers. The length of the list should be the same as that of the ALPHABET string	/path/to/fitness/file.txt whose contents may be: 1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1	landscape

SIGMA	The value of sigma for the lognormal distribution**. Required if INITIAL_FITNESS is LOGNORM, ignored otherwise.**	0.1	landscape
GAMMA_ALPHA	Alpha for the gamma distribution. **Required if INITIAL_FITNESS is GAMMA. Ignored otherwise.**	1.0	landscape
GAMMA_BETA	Beta for the gamma distribution**. Required if INITIAL_FITNESS is GAMMA. Ignored otherwise.**	0.5	landscape
The following fields deal with landscape changes			
Rules determining when and how the landscape changes			
NEW_FITNESS_RULE	How do we obtain the new fitness landscape? iid: sample it from the same distribution as the initial fitness (in which case INITIAL_FITNESS **must** come from a distribution, not a file) shuffle: new fitness vector is produced by randomly shuffling the previous fitness vector current_allele_dependent: the fitness of the current letter is increased every LANDSCAPE_CHANGE_INTERVAL time units by AGE_DEPENDENCE_COEFFICIENT. **This option only works with landscape change timing is at deterministic fixed intervals and sequence of LENGTH 1** user_set: use this option of the user specifies landscape change times by hand (in the CHANGE_BRANCH_AND_TIME file and specifies the new fitness vector values in the same file. **Requires** LANDSCAPE_CHANGE_TIMING to **be set to** specified_branch_and_time	iid shuffle current_allele_dependent user_set	landscape

LANDSCAPE_CHANGE_TIMING	How do we determine when to change the landscape? stochastic: landscape change is a Poisson process with parameter equal to LANDSCAPE_CHANGE_PARAMETER (time is in terms of tree length units). fixed_num_changes: the landscape is changed at fixed-length intervals whose length is such that there will be LANDSCAPE_CHANGE_PARAMETER changes along the longest path from the root to a leaf in a tree. fixed_interval_length: the landscape is changed every LANDSCAPE_CHANGE_PARAMETER time (tree length) units. specified_branch_and_time: the user manually sets the tree position of the landscape change(s) in the CHANGE_BRANCH_AND_TIME_FILE	stochastic fixed_num_changes fixed_interval_length	landscape
CHANGE_BRANCH_AND_TIME_FILE	**Required if** LANDSCAPE_CHANGE_TIMING is **set to** specified_branch_and_time. Name of the file containing tree positions of user-specified landscape changes and, optionally, the new landscape at those positions.	change_branch_time.txt The first word is a tree node, the second is the distance to that node on the branch leading up to it where the landscape is changed. The remaining entries on that line (optional) are the values of the new landscape. Example of contents of the file: C 0.5 0.1 1 0.1 0.1 D 1 0.1 0.1 1 0.1 E 1.5 0.1 0.1 0.1 1 Meaning: at 0.5 tree units before node C change the landscape to (0.1, 1, 0.1, 0.1); at 1 unit before node D change the landscape to (0.1, 0.1, 1, 0.1); at 1.5 units before node E, change the landscape to (0.1, 0.1, 0.1, 1).	landscape
Parameters detailing the rules of landscape change			
LANDSCAPE_CHANGE_PARAMETER	The single parameter governing the times of landscape change whose interpretation depends on the choice of LANDSCAPE_CHANGE_TIMING.	0.1	landscape

	If LANDSCAPE_CHANGE_TIMING = stochastic, this is the parameter of the Poisson process (mean rate, lambda) If LANDSCAPE_CHANGE_TIMING = fixed_num_changes, then this is the number of landscape changes along the longest root-to-leaf path in the tree If LANDSCAPE_CHANGE_TIMING = fixed_interval_length, it is the length of the interval between time changes (in tree time units)		
AGE_DEPENDENCE_COEFFICIENT	The number added to the current character's fitness value every time unit. **Required if and only if NEW_FITNESS_RULE is current_allele_dependent**	1	landscape
Advanced global parameters (should be specified together with the other global parameters in the beginning of the file)			
SHARED_LANDSCAPE	Do parallel tree branches share the same landscape? If true, then there is one copy of the landscape that is evolving with time and that is shared by all branches of the phylogenetic tree. If false, then parallel branches of the tree have separate independently evolving landscapes. **This option is incompatible with the current allele-dependent landscape change (see below).**	true **false** (default)	global
Q_NORMALIZATION	How is the Q matrix normalized? constant_rate: \sum_i -q_ii*p_i = 1, i.e., the mean rate of change per time unit is 1. This is the approach taken by Yang's evolver constant_for_flat: each q_ij is divided by the alphabet length – 1, so that \sum_i – q_ii * pi_i = 1 if all q_ij =1. Thus, if the landscape is flat, the mean rate of change per time unit is 1 none: no normalization	**constant_rate** (default)	global
SCALE_LANDSCAPCE_CHANGE_TO_SUBSTITUTION_RATE	Only relevant if LANDSCAPE_CHANGE_TIMING is stochastic and Q_NORMALIZATION is not constant_rate. If true, the rate of landscape change is obtained by multiplying the LANDSCAPE_CHANGE_PARAMETER by the allele substitution rate. WARNING: since this option requires recomputing the stationary distribution vector π at every landscape change, it will likely slow down execution.	**false** (default)	global
SEED	Seed for the random number generator. For debugging purposes only. Not used by default.	12345	global