

The DFE- α Wrapper offers five options:

1. Generate point estimates on 1 or 2 gene lists
2. Bootstrap *with replacement* across a single set of genes
3. Bootstrap *without replacement* across a larger set of genes N times, where N = the size of a smaller gene set
4. Shuffles the labels of two gene sets (unequal sample sizes allowed) to perform permutation tests
5. Takes one smaller N -sized list, resamples *without replacement* N subsamples from a second, larger sized list; then shuffles the labels of the two equally-sized lists to perform permutation tests

Table 1. Parameters of DFE- α WRAPPER

Parameter	Description	Special Options	Required?
-glist1	Gene List 1	set -set1 to a specific name if desired	Yes
-glist2	Gene List 2	set -set2 to a specific name if desired	No
-summary	Polymorphorama Summary Statistics File		Yes
-rep_poly	Replacement SFS from Polymorphorama		Yes
-syn_poly	Synonymous SFS from Polymorphorama		Yes
-paml_div	Divergence Counts Generated by PAML or other outside source	If excluded, will use the divergences estimated by Polymorphorama	No
-num_iter	Number of iterations to resample, <i>e.g.</i> , 1000	Set to 0 to run point estimates only	Yes, but program defaults to 0 if not selected
-set1	Optional name for Gene List 1		No
-set2	Optional name for Gene List 2		No
Resampling Schema			
-bs_w_repl	Resamples only Gene List 1 with replacement for bootstrapping		One is Required
-bs_wo_repl	Resamples only Gene List 2 withOUT replacement for bootstrapping	Requires Gene List 1 to be provided as well	

-labelshuffle	Shuffles labels of Gene List 1 and Gene List 2 for permutation test		
-bkgd	Takes an N sized set from Gene List 2, combines with Gene List 1, then shuffles labels for permutation test	Can only be used with -labelshuffle; cannot be used alone	No

Generate Point Estimates

This option will allow you to generate single point estimates of α for one or two gene sets.

One Gene Set (put code all on **one line** in your terminal):

```
perl DFE-alphaWRAPPER.pl
  -glist1 test_sets/test_set_4.txt
  -summary test_sets/summary_statistics.txt
  -rep_poly test_sets/frequencies_replacement.txt
  -syn_poly test_sets/frequencies_synonymous.txt
  -paml_div test_sets/paml_divergence.txt
  -num_iter 0
  -set1 testSet
```

Two Gene Sets (put code all on **one line** in your terminal):

```
perl DFE-alphaWRAPPER.pl
  -glist1 test_sets/test_set_4.txt
  -glist2 test_sets/test_set_35.txt
  -summary test_sets/summary_statistics.txt
  -rep_poly test_sets/frequencies_replacement.txt
  -syn_poly test_sets/frequencies_synonymous.txt
  -paml_div test_sets/paml_divergence.txt
  -num_iter 0
  -set1 testSet4
  -set2 testSet35
```

Bootstrap With Replacement Across a Single Set of Genes

This option will allow you to X number of bootstraps across a single gene set. For example, you can iteratively resample a gene set 10 times:

```
perl DFE-alphaWRAPPER.pl
  -glist1 test_sets/test_set_4.txt
  -summary test_sets/summary_statistics.txt
  -rep_poly test_sets/frequencies_replacement.txt
  -syn_poly test_sets/frequencies_synonymous.txt
  -paml_div test_sets/paml_divergence.txt
  -num_iter 10
  -set1 testSet4
  -bs_w_repl
```

Bootstrap Without Replacement Across a Single Set of Genes

This option will allow you to perform X number of subsamples without replacement across a single gene set using a smaller gene set for sample size N . For example, you can iteratively resample a larger background set of genes 10 times using a smaller -glist1 to get an $N=4$ each resample:

```
perl DFE-alphaWRAPPER.pl
  -glist1 test_sets/test_set_4.txt
  -glist2 test_sets/test_set_35.txt
  -summary test_sets/summary_statistics.txt
  -rep_poly test_sets/frequencies_replacement.txt
  -syn_poly test_sets/frequencies_synonymous.txt
  -paml_div test_sets/paml_divergence.txt
  -num_iter 10
  -set1 testSet4
  -set2 testSet35
  -bs_wo_repl
```

Permutation Test with Label Shuffling on Two Gene Sets

This option will allow you to perform X number of label shuffling events where gene lists 1 and 2 are combined, labels shuffled, and the resulting pool randomly subdivided into a N_1 - and N_2 -sized set. Sampling is done without replacement. For example, you can iteratively shuffle a gene set $N_1=4$ and set $N_2=35$ ten times to create ten new N_1 - and N_2 -sized sets:

```
perl DFE-alphaWRAPPER.pl
  -glist1 test_sets/test_set_4.txt
  -glist2 test_sets/test_set_35.txt
  -summary test_sets/summary_statistics.txt
  -rep_poly test_sets/frequencies_replacement.txt
  -syn_poly test_sets/frequencies_synonymous.txt
  -paml_div test_sets/paml_divergence.txt
  -num_iter 10
  -set1 testSet4
  -set2 testSet35
  -labelshuffle
```

Permutation Test with Label Shuffling on Two Gene Sets Where One is a Larger, Background Set

This option will allow you to perform X number of label shuffling events where first gene list 2 is randomly sampled to get an N_1 -sized subset. This subset is combined with gene list 1, labels shuffled, and the resulting pool randomly subdivided into two N_1 -sized sets. Sampling is done without replacement. For example, you can iteratively shuffle a gene set $N_1=4$ and set $N_2=39$ ten times to create ten new N_1 -sized sets:

```
perl DFE-alphaWRAPPER.pl
  -glist1 test_sets/test_set_4.txt
  -glist2 test_sets/test_set_39.txt
  -summary test_sets/summary_statistics.txt
  -rep_poly test_sets/frequencies_replacement.txt
```

```
-syn_poly test_sets/frequencies_synonymous.txt  
-paml_div test_sets/paml_divergence.txt  
-num_iter 10  
-set1 testSet4  
-set2 Background  
-labelshuffle  
-bkgd
```

Note: Both Label Shuffling options will check for overlap of lists and, if found, correct for it. The above example performs such a correction for overlap between gene list 1 (N =4) and gene list 2 (N = 39).