

Installation

MSG currently runs on any linux platform.

Place `msg-version.tar.gz` in any directory:

```
$ tar xzf msg-version.tar.gz
$ cd msg_version/
$ make
```

Please ensure that the following dependencies are installed before running MSG (working versions are indicated in parentheses):

- Python (2.6)
- bwa (0.5.7)
- samtools (0.1.9-3)
- biopython-1.53
- Pyrex-0.9.9
- pysam-0.1.2 (apply fix*)
- R packages (HiddenMarkov 1.3-1, zoo 1.6-2, R.methodsS3 1.2.0 and R.oo 1.7.3)

*<http://code.google.com/p/pysam/issues/detail?id=22&can=1&q=dandavison0>

Setting up the MSG analysis directory

1. Create a text file named `msg.cfg`. This file will specify the location of your data files, and a few other details. You can copy and fill-in the details in the sample `msg.cfg` file provided within your `msg` directory. All data files must be located within your MSG analysis directory (links to files are acceptable).

2. Format your barcode file as below:

sp1	sp2	sp1_gen	sp2_gen	enzyme	pre-pd	linker_type	DeIQV	strict	minQV				
sim	sec	dsim-all-chromosome-r1.3.fasta	dsec-all-chromosome-r1.3.fasta	MseI	1	Dros_SR_vII	0	0	0				
GATAG	B3	Bcsec	male										
CACCG	C12	Bcsec	female										
CTTAT	F10	Bcsec	female										

3. Create a link to the `msg` software within your MSG analysis directory:

```
$ ln -s path_to_msg msg
```
4. To run MSG, simply type the following from within your MSG analysis directory:

```
$ perl msg/msgCluster.pl
```

Sample data

Please download the following data:

Data from manuscript

<http://trace.ncbi.nlm.nih.gov/Traces/home/>

- backcross data
SRA025671.2/SRR071201.4
- parental Dsim_w501
SRA025671.2/SRR074287.1
- parental Dsec_wsu1
SRA025671.2/SRR074288.1

Parental reference sequences

<ftp://ftp.flybase.net/genomes/>

- dsim-all-chromosome-r1.3.fasta.gz
- dsec-all-chromosome-r1.3.fasta.gz