Using the IMfig program

By Jody Hey

Table of Contents

Introduction

IMfig is a program for generating a figure of an isolation-with-migration model from the output file generated by running either the IMa2 or the IMa3 program.   IMfig generates an eps (encapsulated post script) file by default and will also generate files in other formats (jpeg, png, pdf).   Examples of the figures generated by IMfig are given in (Hey 2010; Hey, et al. 2017) and later in this document.

Note that some options have changed from earlier releases of this program, and some new options have been added with the release of IMa3.

This kind of figure represents an estimate of history as a fat phylogenetic tree made up of boxes (for sampled and ancestral populations), horizontal lines (for splitting times) and curved arrows (for migration). Time is represented as depth on the vertical axis, with the sampled populations/species at the top of each figure at the most recent time. Population size is represented as width along the horizontal axis. Each population is represented by a box, the height of which refers to how long it has lasted and the width of which refers to its effective size. The confidence interval for a populations size is given both by a double headed arrow extending from the right margin of that population's box and by faint boxes representing the lower and higher 95% Highest Posterior Density (HPD) intervals. Similarly 95% HPD intervals for splitting times are given by dashed lines and doubled-headed arrows. Migration arrows are printed depending on the users wishes. There are two options for only printing arrows for

migration rates that are statistically significant by the test of Nielsen and Wakeley (2001) with asterisks: "*" for p<0.05; "**" for p < 0.01, and " ***" for p < 0.001.

The source code for IMfig is a Python script, and execution requires that Python be installed (http://www.python.org/). The latest version has been tested with Python 3.6 but will probably run ok with any version later than Python 2.7.   Users are also recommended to use an eps file viewer. Getting the settings right for a particular figure can take some playing around and may ultimately require several runs of   IMfig. The program runs quickly, but the process of getting the best settings may require multiple runs while trying out various options.

Generating an IMa2/IMa3 output file suitable for IMfig

The file that is read by IMfig should be an output file generated using a run of IMa2 or IMa3 that generates estimates of population sizes and migration rates.   These runs will use a fixed topology that is included in the data file, and they will not use hyperpriors (an option in IMa3). IMfig extracts information from the phylogenetic tree in the original data file and from the following tables in the IMa2 or IMa3 output file:

- `Marginal Peak Locations and Probabilities`
- `MARGINAL DISTRIBUTION VALUES AND HISTOGRAMS OF POPULATION SIZE AND MIGRATION PARAMETERS`
- `MARGINAL DISTRIBUTION VALUES AND HISTOGRAMS OF PARAMETERS IN MCMC`
- `MARGINAL DISTRIBUTION VALUES IN DEMOGRAPHIC UNITS`

The last table (for demographic units) will be skipped by IMfig if the −d option is used on the command line. Doing an IMa2 or IMa3 analysis in demographic units requires that the program be run using demographic units, which requires in turn some information on mutation rates and generation times (see IMa2 or IMa3 documentation).

Command line options:

Running the program without any flags should cause the following text to appear on the screen:

```
IMfig program. Copyright 2009-2017  Jody Hey
IMfig command line terms  (-i is required):
help or h causes this screen to be written
-a : include ancestral population #'s in plot
-b : adjust width spacing of population boxes, values > 0, default = 1
-c : convert eps file
     -c j : make a jpeg file
     -c p : make a pdf file
     -c n : make a png file
-d : do not use demographic scale information even if in input file
-e : space split times evenly  (not proportional to time,  no confidence intervals shown)
-f : font.  Default=Arial. Use postscript fonts available on the computer
     e.g. Arial, Helvetica, Times-roman, Courier
-g : global plot scale sets the size of the plot, max = 1, default = 1
-h : arrow width, default = 1
-i : input file name
-k : print population names on an angle
-m : options for printing of arrows and 2Nm values for migration :
     -m x :  do not print migration arrows
     -m a : 2Nm migration arrows for all 2NM > 0
     -m s : 2Nm migration arrows only if m is statistically significant p <= 0.05 (default)
     -m S : 2Nm migration arrows only if m is statistically significant p <= 0.01
     -m # : '#' is a number, migration arrows appear when 2NM >= # (e.g. -m0.1)
-n : file with alternative species names
-o : output file name (with 'eps' extension) e.g. -o myoutputfile.eps, default=im_eps_file.eps
-p : fontsize (default is 14 for full scale, default follows global scale)
-q : no confidence interval boxes for population boxes printed
-r : no confidence interval arrows for population boxes printed
-s : print square, rather than landscape
-t : relative height of oldest time point, values between 0 and 1
     default value = 1/(# sampled populations+1)
-u : simple colors, blue for population boxes, red arrows for migration (default grayscale)
-v : multiple colors for population boxes, red arrows for migration (default grayscale)
-x : adjust width of plot,  >1 means wider, <1 means narrower
-y : adjust height of splittimes, relative to bottom of figure, max = 1.
```

Only the $-i$ (input file) flat is required in order to generate a figure. All other flags are optional. Several flags affect the size/shape of the plot. If you need to generate multiple figures in which the height and width of the figures are on the same scale (e.g. so that ½ inch of height corresponds to a certain length of time) then you can use the $-t$ and/or the $-y$ options. Similarly to have a unit distance on the horizontal axis mean the same thing (e.g. ½ inch corresponds to a certain effective population size) for different figures, you can play around with the $-x$ option. If you use $-g$ to make the plot bigger or smaller you may also need to adjust the font size ($-f$) and the arrow head size ($-h$).

Running IMfig

IMfig.py was tested on python version 3.6 and seems to run ok under 2.7 as well, but has not been tested thoroughly in this way.   Running IMfig.py in a shell window (e.g. cmd, terminal, bash shell etc) should be fairly straightforward. For example if your IMa2 output file is called `my_results.out` and

you want the eps file to be named `my_results.eps` then enter the following at the command prompt:

```
python IMfig.py -imy_results.out -omy_results.eps
```

(Note that this assumes the operating system knows where python is (i.e. it is in the system path), and that IMfig.py and my_results.out are in the current directory.)

Full functionality of IMfig.py requires that the python installation include the Colormath module (https://pypi.python.org/pypi/colormath/) for better color picking when using $-v$ (though it does not make a huge difference)m and the Pillow module (http://pillow.readthedocs.io/en/3.4.x/index.html) for using $-c$ to write to other file formats.   The pillow module requires in turn that Ghostscript be installed (https://www.ghostscript.com/).

IMfig.py will run without these but the corresponding command line options will not work.

Output Files

IMfig generates a vector-based encapsulated postscript (eps) file. There are many programs for viewing and editing eps files such as Inkscape and Illustrator. If the $-c$ option is used and the user's python installation includes the pillow module and ghostscript has been installed, then in addition to the eps file IMfig will generate an image in one of several other formats, including jpeg, png and pdf. These files are generated in fairly high resolution so as to be suitable for publication.

References:

Hey J. 2010. The Divergence of Chimpanzee Species and Subspecies as Revealed in Multipopulation Isolation-with-Migration Analyses. Mol Biol Evol 27:921-933.

Hey J, Chung Y, Sethuraman A, Wang Y. 2017. Phylogeny estimation by integration over isolation with migration models. Manuscript.

Examples

On the next several pages are shown some results using an IMa2 output file from an analysis of a 3 population model. On each page is given a command line, a brief description and the resulting figure.
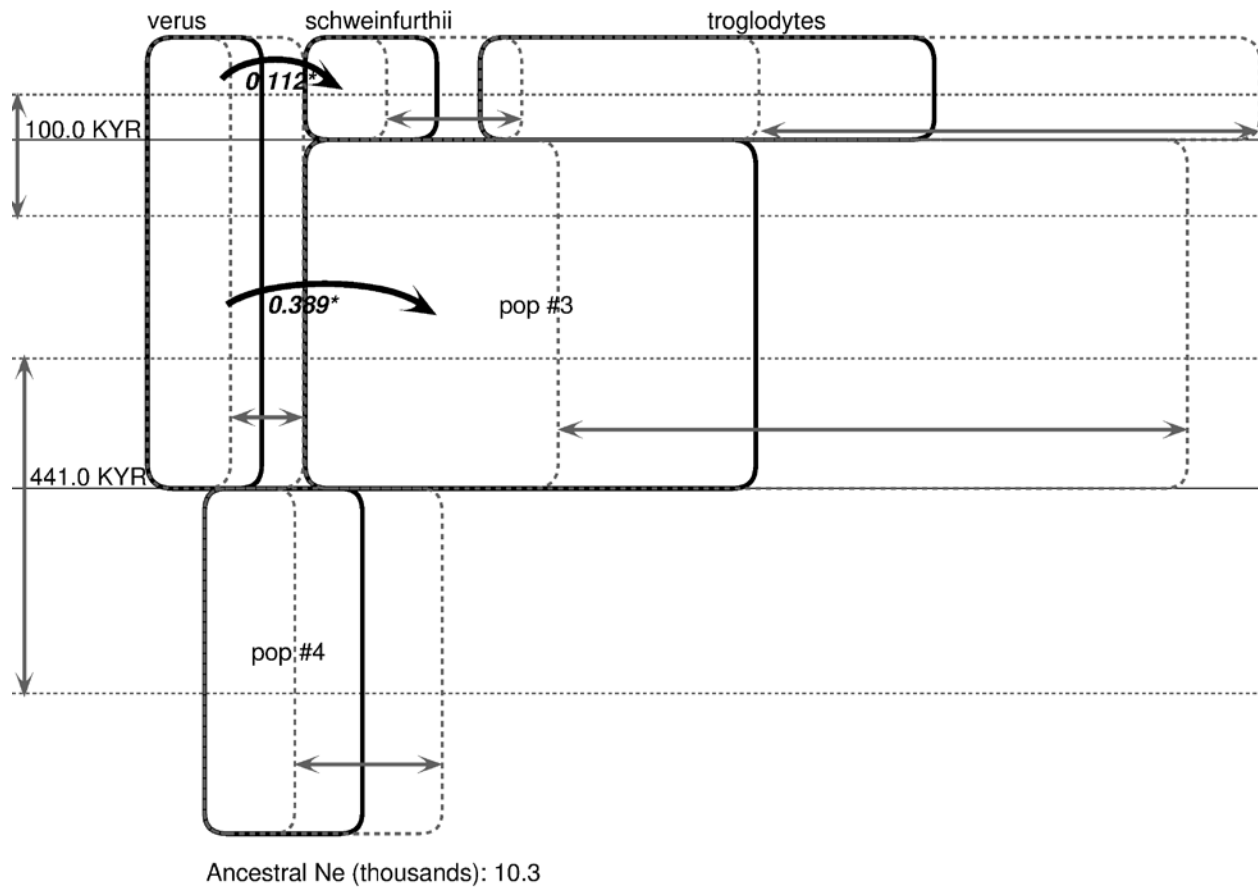
Default settings

```
python IMfig.py -iIMa2_3pop_example.out
```
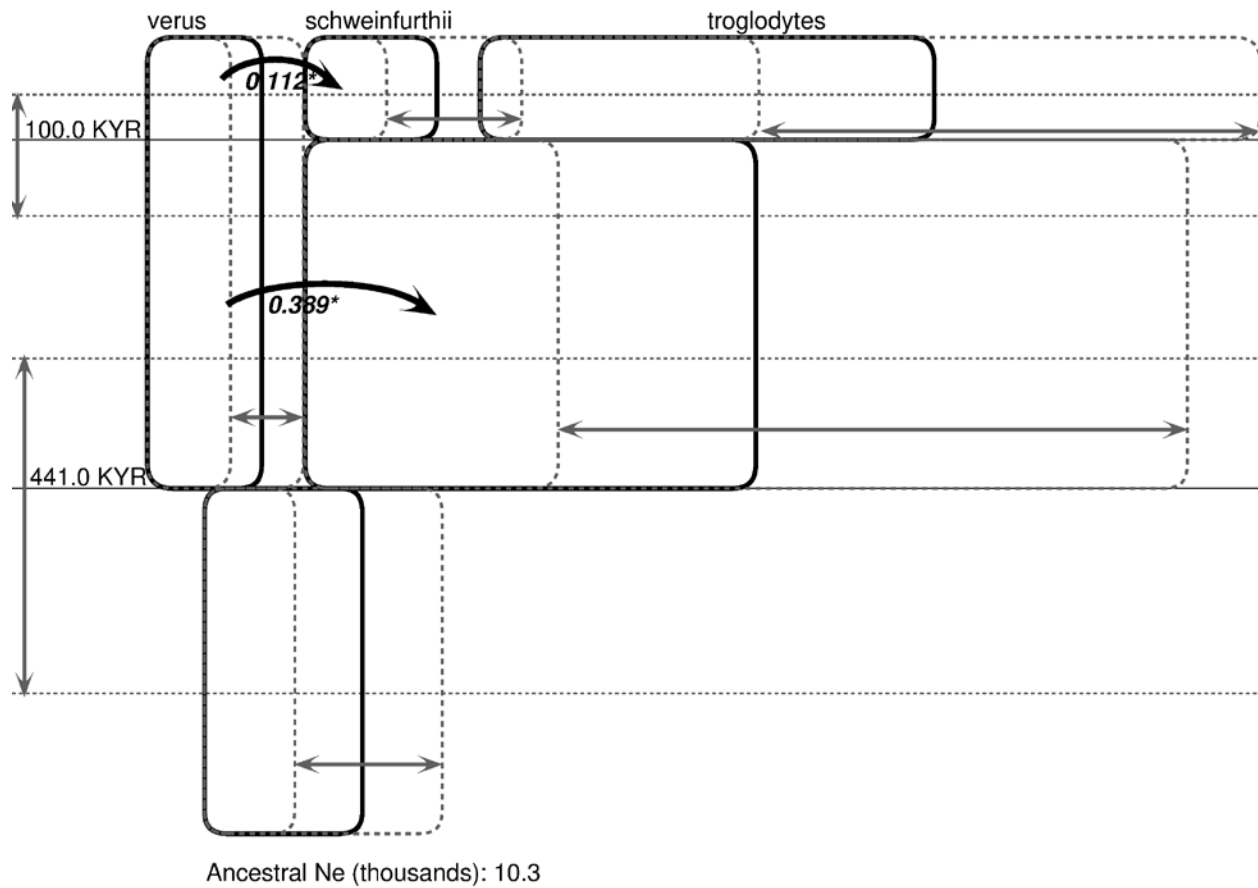
Example demonstrating the use of the –a option to have the population label number for ancestral

populations included in the figure
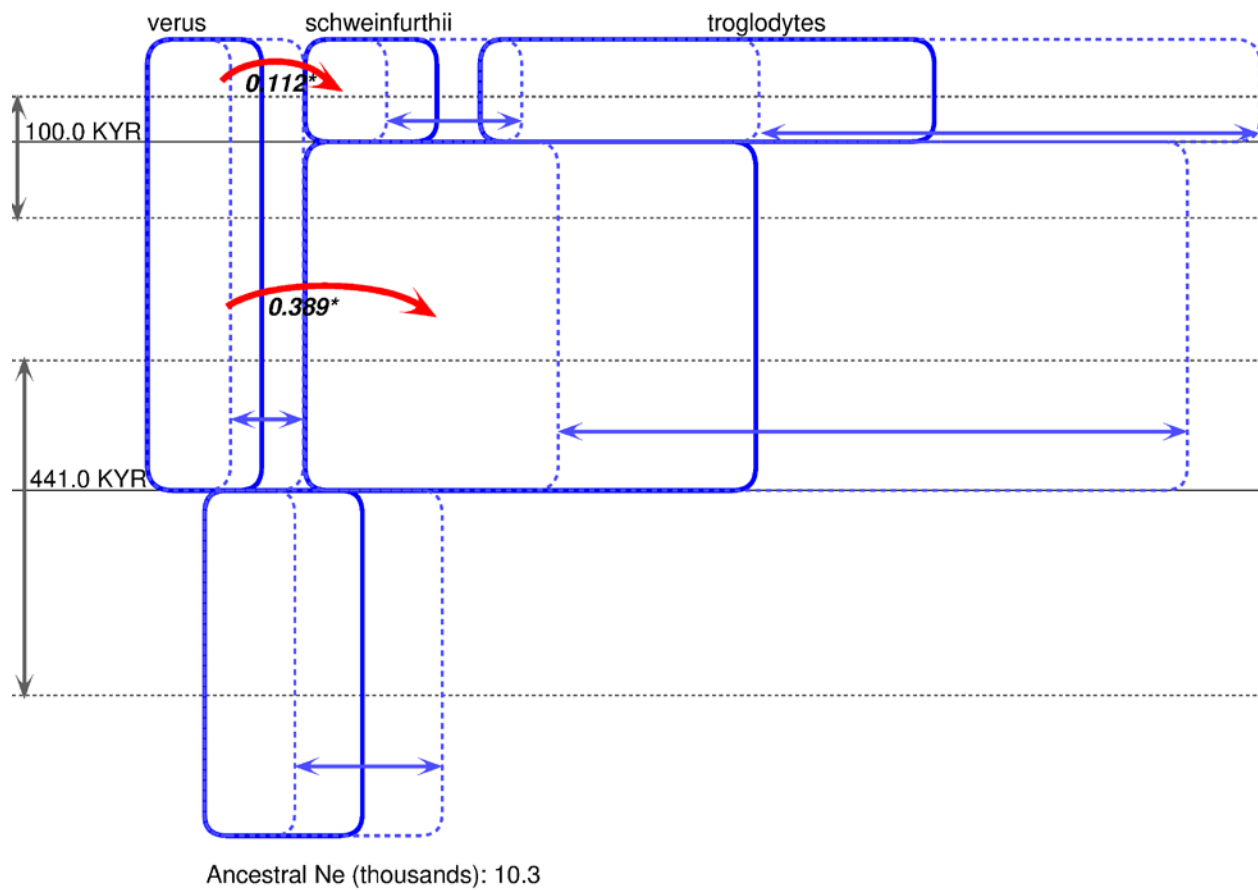
```
python IMfig.py -iIMa2_3pop_example.out -a
```

Example demonstrating the use of the –b option to alter the width of population boxes to avoid overlapping confidence intervals

```
python IMfig.py -iIMa2_3pop_example.out -b3.0
```
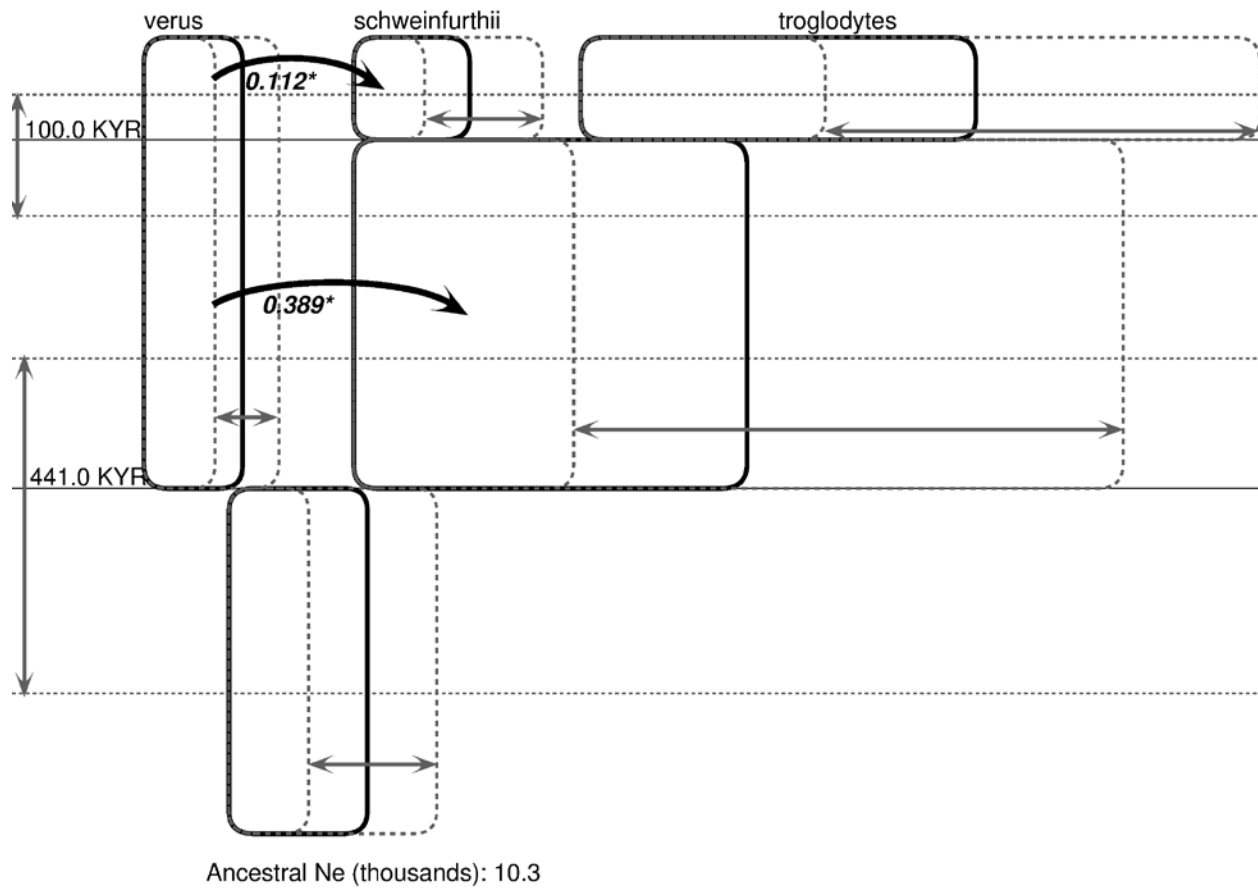
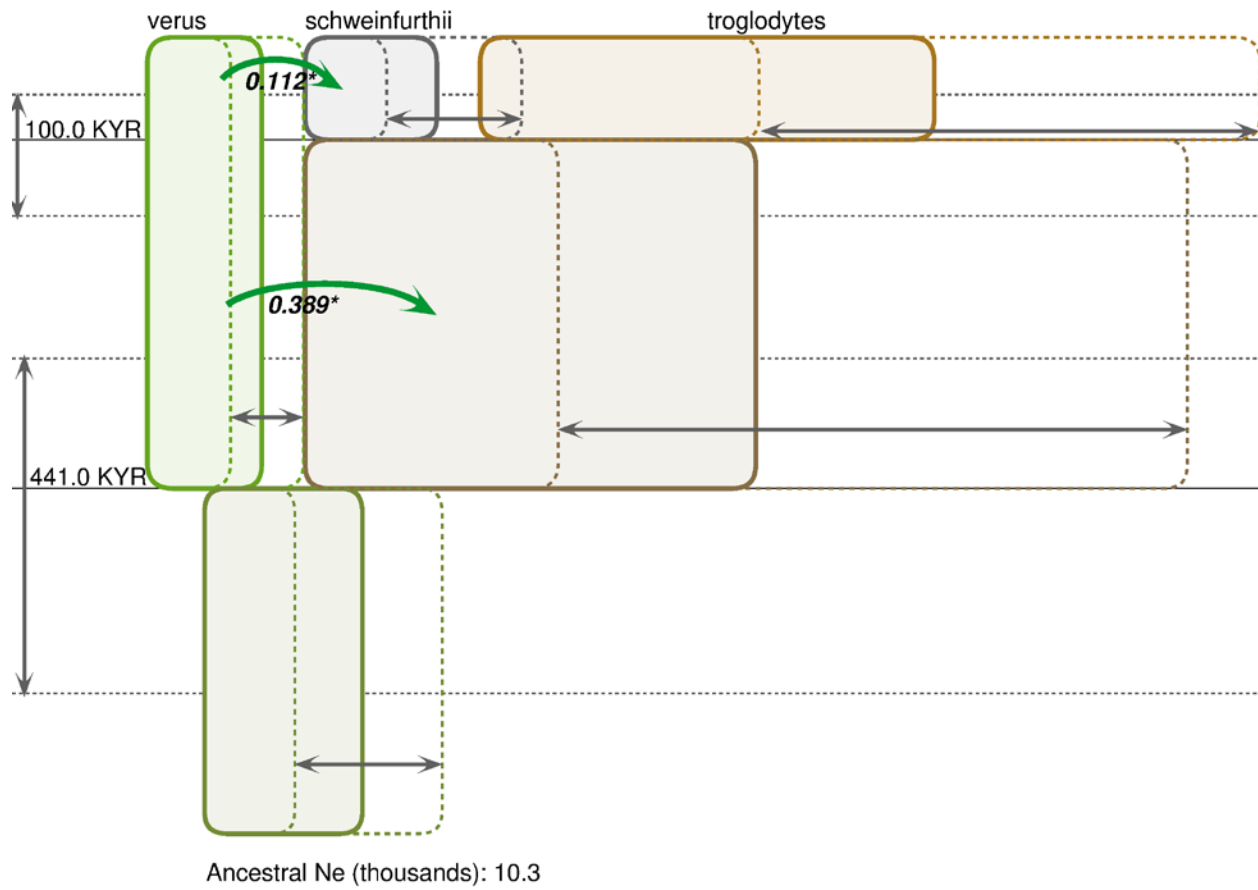Example demonstrating the use of the –c option to print in color

```
python IMfig.py -iIMa2_3pop_example.out -u
```


verus   schweinfurthii   troglodytes

0.112*

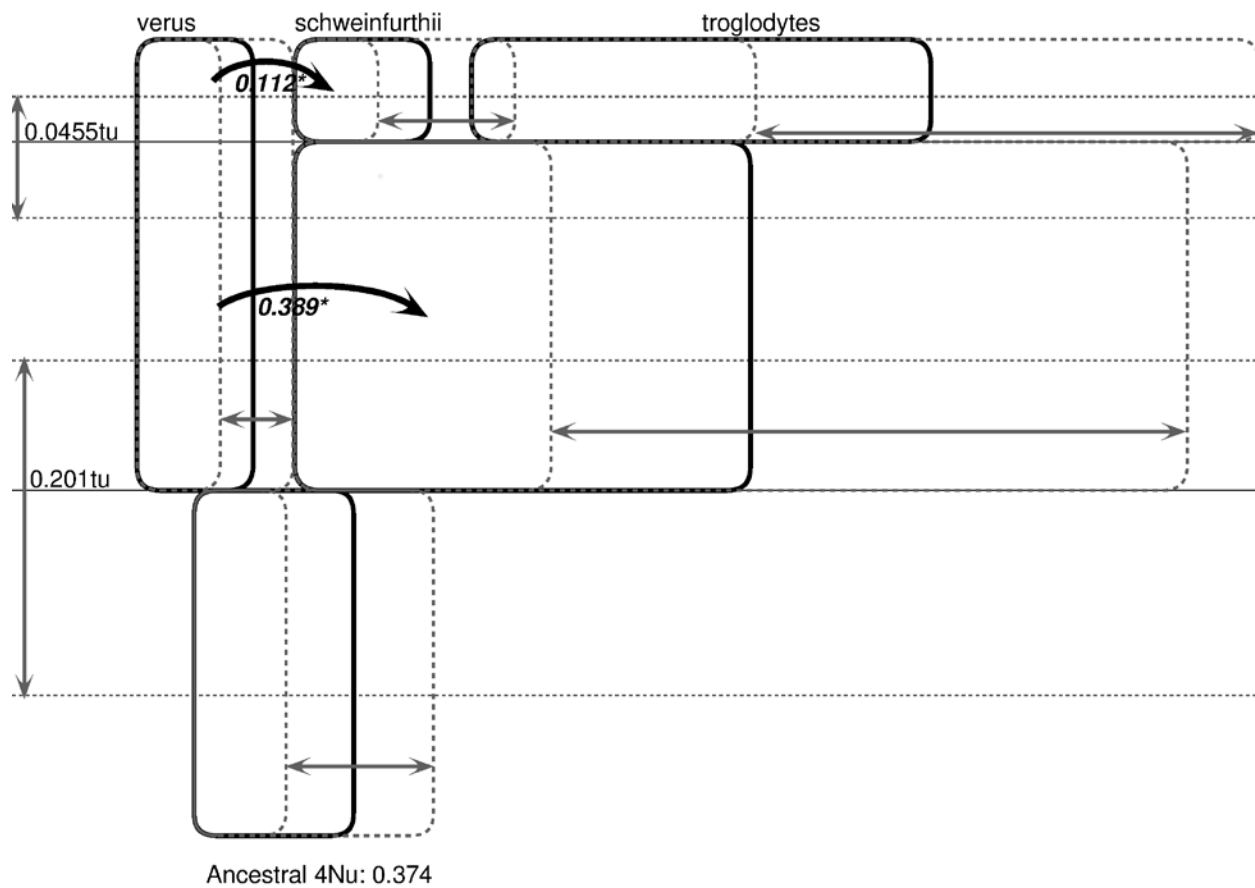0.389*

100.0 KYR

441.0 KYR

Ancestral Ne (thousands): 10.3

Example demonstrating the use of the –v option to print in color
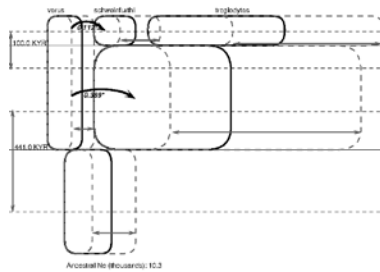
```
python IMfig.py -iIMa2_3pop_example.out -v
```

Example demonstrating the use of the –d option to use mutation rate scales rather than demographic scales

```
python IMfig.py -iIMa2_3pop_example.out –d
```


verus   schweinfurthii   troglodytes
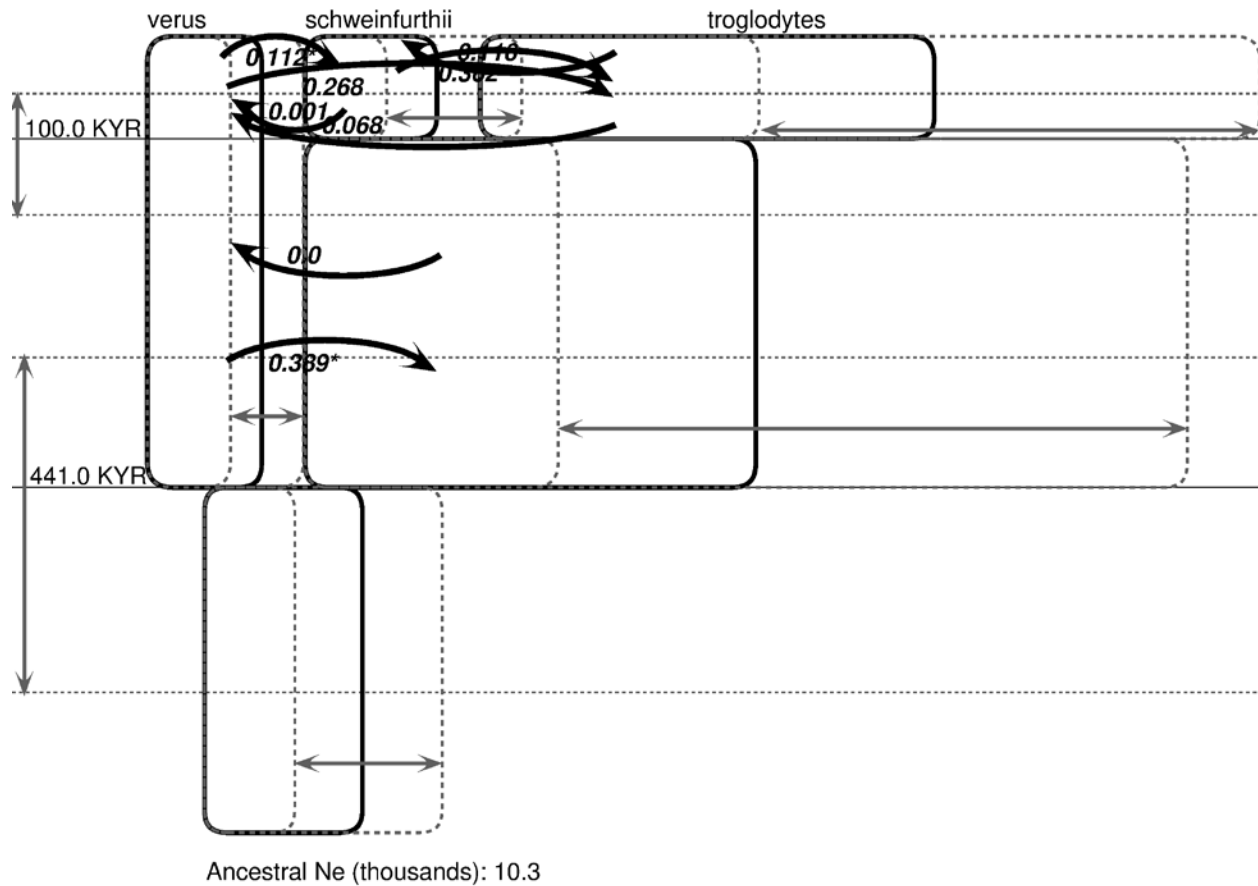
0.0455tu

0.112*

0.389*

0.201tu

Ancestral 4Nu: 0.374

Example demonstrating the use of the –g option to rescale the entire plot, in this case a much smaller figure

```
python IMfig.py -iIMa2_3pop_example.out -g0.3
```

Example demonstrating the use of the –ma option to print all arrows for all migration rates where 2NM>0

```
python IMfig.py -iIMa2_3pop_example.out –m a
```

Example demonstrating the use of the -k –n and –v options to print a figure with alternative population names, on an angle, and with color . For an input file with m populations, the file with alternative names should have m lines with line j having the alternative name for population j in the input file.

```
python IMfig.py –iIMa2_3pop_example.out –k –n popnames.txt –v
```