PPGtk: the polyploid pop-gen toolkit

Written and maintained by Paul Blischak

E-mail: blischak.4@osu.edu

Source code: `https://github.com/pblischak/ppgtk`

Online docs: `http://pblischak.github.io/ppgtk`

Contents

PPGtk is a set of tools written in C++ for population genetic/genomic analyses with polyploids using high-throughput sequencing data. It was designed for research in non-model taxa (i.e., no reference genome), and provides a number of functions for conducting population genomics analyses in these taxa.

1 Installation

Installing *PPGtk* follows the typical workflow for Unix-based operating systems using a Makefile and the commands '`make && sudo make install`'.

1.1 Obtaining *PPGtk*

A "bleeding edge" version of the source code for *PPGtk* can be cloned from GitHub using:

```
$ git clone https://github.com/pblischak/ppgtk.git
```

from the command line. You can also download the latest stable release from GitHub by following the **releases** link in the *PPGtk* repository.

```
$ cd ppgtk-v1.0.0

[R]> dat <- read.table("tot-reads.txt")
```

1.2 Boost C++ Libraries

1.3 OpenMP

can be used for parallelization.

1.4 Compiling the executable

To enable parallelization with OpenMP, you will need to change the `OPENMP` variable in the Makefile to `yes`, like so:

```
OPENMP ?= yes
```

If you installed Boost in `/usr/local/bin`, then the paths to the header files and libraries should already be correctly specified in the Makefile. If you installed it somewhere else, then you will need to change the `BOOST_LIB` and `BOOST_INC` variables to point to where it was installed.

If you are on a Mac, type:

```
$ make darwin
```

If you are on a Linux computer, type:

```
$ make linux
```

2 Getting Started

```
$ ppgtk -n 100 -l 2000 -p 4 -t total.txt -r reference.txt -q
```

`total.txt`:

```
24 39 12 0 63 0 33 45 . . .
20 33 21 0 0 78 27 11 . . .
.
.
.
7 4 80 22 26 22 18 4 . . .
```