mpileup2pro.cpp

October 7, 2022

Author: Takahiro Maruki

C++ program for making pro files of nucleotide read counts from mpileup files.

This C++ program is for preparing pro files of nucleotide read counts from mpileup (http://www.htslib.org/doc/samtools-mpileup.html) files.

Input files. The input files consist of 1) a tab-delimited text file of the reference nucleotide, which can be made using Ext_Ref_Nuc.pl in Package-GFE (https://github.com/Takahiro-Maruki/Package-GFE); 2) an mpileup file made using Samtools (Li *et al*. 2009); and 3) a text file listing individual IDs. The order of the individuals in the list needs to correspond to that in the mpileup file.

Output file. The output file is a tab-delimited text file. Columns 1, 2, 3 denote the scaffold ID, position, and reference nucleotide, respectively. Each of the subsequent columns denotes counts of nucleotide reads A, C, G, and T separated by slashes in each individual.

Reference

If you use this program, please cite the following paper:

Maruki, T, Ozere, A, and Cristescu, M. E. Genome-wide identification of single nucleotide mutations from time-series mutation accumulation data. In prep.

Instructions

Below are specific procedures for using the program:

1. Compile the program by typing the following command:

g++ -o mpileup2pro mpileup2pro.cpp -lm

3. Run the program by typing the following command:

./mpileup2pro -ref RefNuc.txt -id IDs.txt -mp In.mpileup -out Out.pro

- The `-id', `-mp' and `-out' options specify the list of individual IDs, mpileup file, and output file name, respectively.

- A usage help message explaining these options can be shown by typing the following command:

./mpileup2pro -h

Copyright notice

This program is freely available; and can be redistributed and/or modified under the terms of the GNU General Public License as published by the Free Software Foundation; either version 3 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

For a copy of the GNU General Public License write to the Free Software Foundation, Inc., 59 Temple Place, Suite 330, Boston, MA 02111-1307 USA

Contact

If you have difficulty using this software, please send the following information to Takahiro Maruki (takahiro.maruki@mcgill.ca):

      1. Brief explanation of the problem.

      2. Command used.

      3. Part of the input file.

      4. Part of the output file.