Du for detecting recent positive selection

How to use:

1) Currently, the *Du* program supports the four popularly used alignment format (ARLEQUIN, MEGA, NEXUS and PHYLIP). Please see the examples for the data format. If there is an outgroup sequence in the alignment, the name of outgroup sequence has to be named as "outgroup". One outgroup is supported. The *Du* program also works well for the cases when there is no any information of outgroup.

2) You need to install Java. Please visit www.java.com to install a Java Runtime Environment.

3) Use a text editor to edit properties.txt. The file tells the *Du* program where your data and the pre-calculated table are stored. Please note that properties.txt has to stay together with *Du*.

directory=D\:\\ Du\\myData
or
directory= //home//lihaipeng//data

4) Pre-calculate a table, given the number of sampled chromosomes (*n*).

initialTable *n*
or
initialTable *dataFileName*

You may need to type "./initialTable" if you are using the linux platform. Moreover, if the number of sampled chromosomes is extremely large, you may want to accelerate the calculation by using a computer cluster. In this case, the jobs could be distributed over computing nodes.

initialTable *n x*

The minimum value of x is 1, and the maximum value of x can be easily calculated by $FLOOR(0.1 \times (n-1)/2)$. And we have $x = n - \Gamma$, where Γ was defined in the reference listed below. Briefly speaking, x and Γ represent the sizes of basal branches. If the number of sampled chromosomes is larger than 10,000, please contact us for an updated version.

5) To calculate Du, the software will build a UPGMA tree based on the nucleotide differences between sequences. Then gamma and Du will be calculated. If you don't want to use the default UPGMA tree, you can use your own tree. Please be aware that the tree has to be consistency with the data, or you will receive an error message.

duTest *dataFileName*
or

duTest *dataFileName treeFile*