IOWA STATE UNIVERSITY

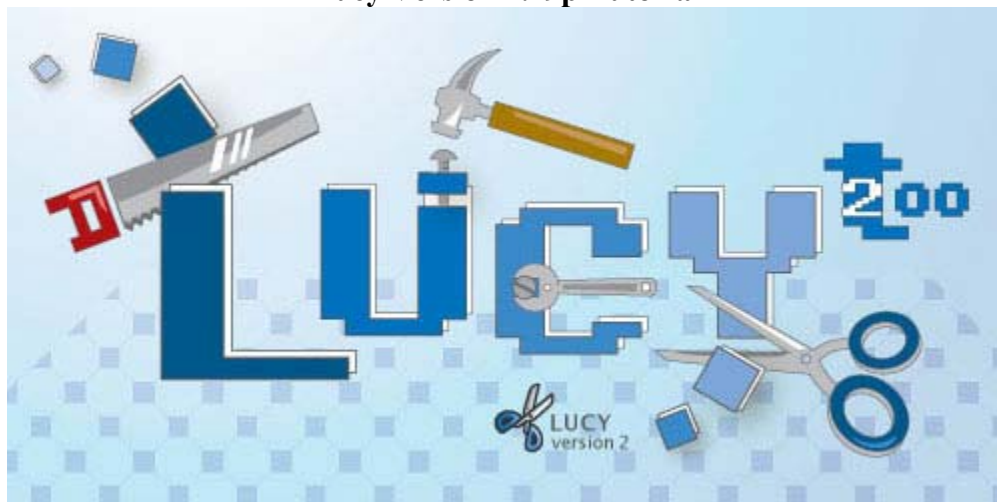**Complex Computation Laboratory**

Lucy Version 2.19p Tutorial

Hui-Hsien Chou
Denise Mooney
Song Li

Lucy Version 2.19p Tutorial

Most of the bioinformatic tools operating on DNA data assume that the sequences are trustworthy. However, raw data obtained from sequencing machines often violate this assumption. The tool Lucy1 was previously created to solve the raw data quality assurance problem.

Lucy2 is a raw DNA sequence trimming and visualization tool based on the command-line tool Lucy1. Lucy1 was originally designed for high-throughput DNA sequencing centers to easily "plug-in" to their data streams; therefore it is operated exclusively through a command-line syntax. This style of operation was not suitable to individuals who may just like to trim a few sequences since it is inconvenient for users to operate with Lucy1 command-line syntax and to memorize its parameters. With Lucy2, these users can perform Lucy1's functionality through a much friendlier user interface. In addition, Lucy2 gives the option for users to physically remove the bad region before saving the data.
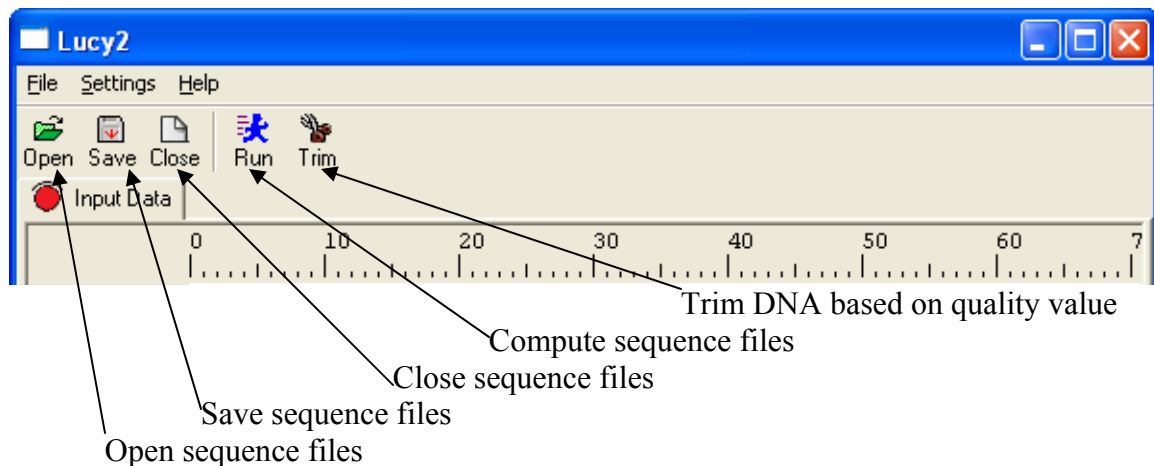
Lucy2 is an open source project. The binary executable files are currently available on Windows, Linux and MacOS.

Introduction to Tutorial

This tutorial is designed to give users a basic understanding of the functions of Lucy2. Users should be able to perform the following tasks upon completion of the tutorial:
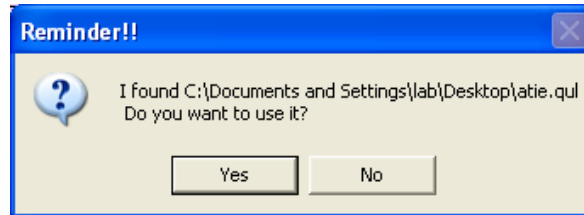
- Load a gene sequence
- Compute the sequence using parameters defined by the program
- Understand and identify meaning of the various highlighted colors, underlined sequences and text colored sequences
- Cut DNA based on Lucy quality value

Guided Tour of Lucy2

Trim DNA based on quality value

Compute sequence files

Close sequence files

Save sequence files

Open sequence files

Loading sequence files

Lucy2 accepts most genome sequences as a sequence file (.seq), however there are others that may be acceptable. Simply left click on *Open* or select Files > Open files from the pull down menu. In this tutorial, we will use the *Atie.seq* downloadable from our website. Be sure to save the file to your computer as a .seq file, not as a .txt file. Also download all files associated with it available on our website. Select the file and choose open. A pop up box will display:
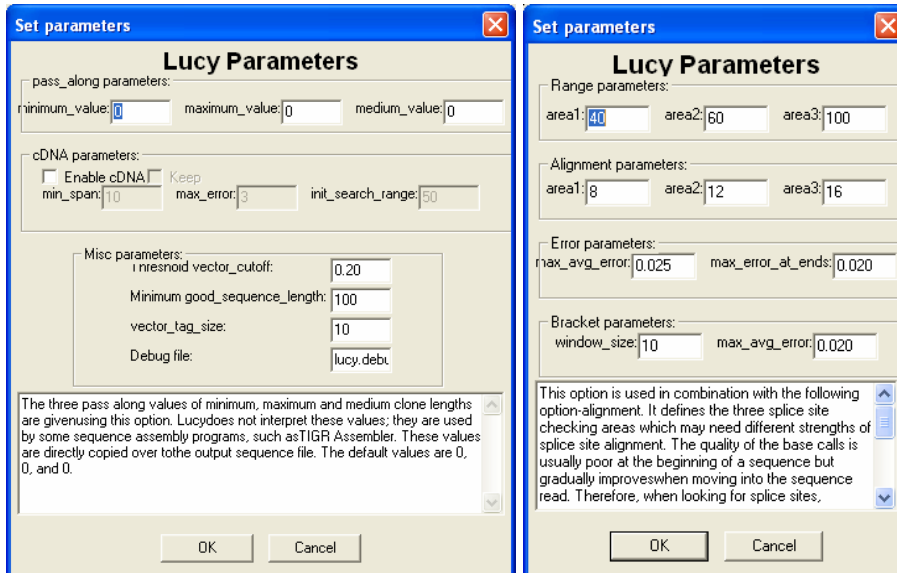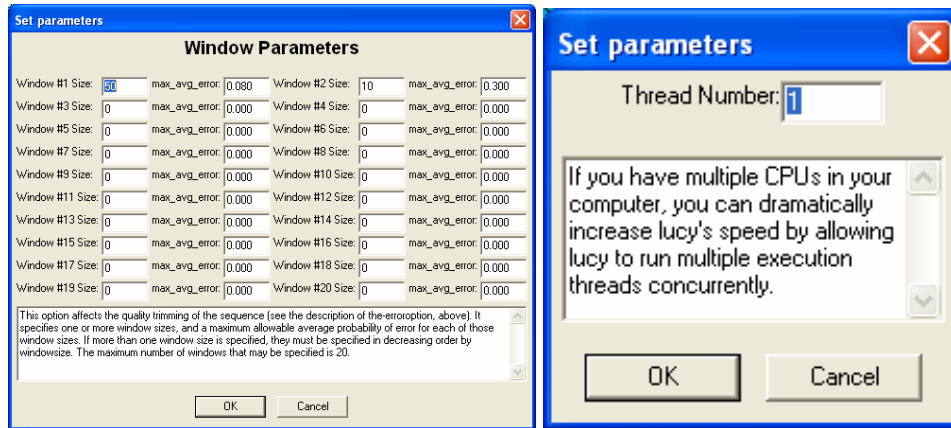

This dialogue box searches for parameters and quality values already set from the original Lucy. Quality files have the extension .qul, in addition, there are two parameter files named PUC19 and PUC19splice that will also be used during this tutorial. Choose *Yes*.

The sequence is now loaded into the program. Use the scroll down menus to view the sequences and numbering.
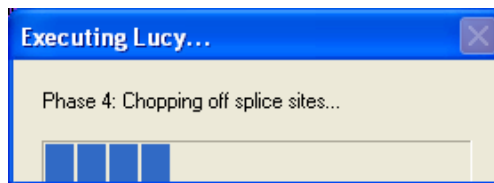
Computing Quality Value

Lucy2 allows the parameters to be specified before calculations. Please see the paper, *DNA sequence quality trimming and vector removal* by Hui-Hsien Chou and Michael H. Holmes for more information about setting parameters. The following four screens are the windows available to set parameters;
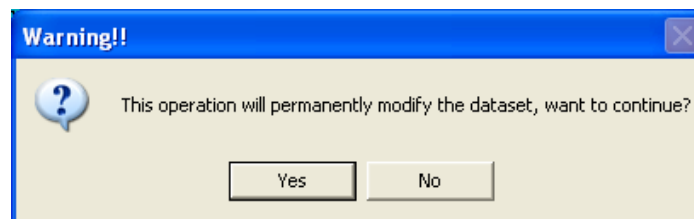
To compute quality values based on the programs pre-set parameters, left click the *Run* or choose File > Run from the pull down menu. The following screen will appear as Lucy calculates quality values:
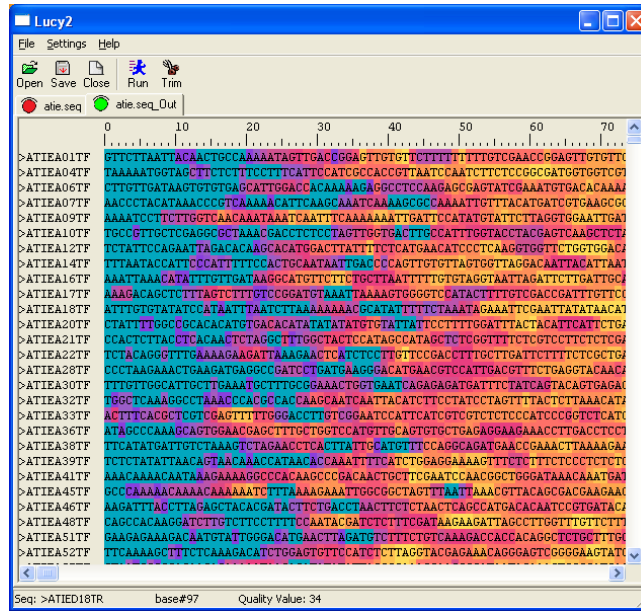

To cut out the poor quality regions of the sequence file loaded, simply left click on *Trim* or select File > Run from the pull down menu. The following screen appears;

Lucy2 will permanently delete the poor quality regions from the input file so make sure there are backup copies of the sequence file. *Select Yes*.

The sequence file now will now be free of the low quality data calculated by Lucy as shown below;

You have completed the Lucy Version 2.19p tutorial. Please stay updated for further versions of Lucy2 at http://www.complex.iastate.edu.

Complex Computation Laboratory
Iowa State University
Hui-Hsien Chou
Denise Mooney
Song Li