FCGS Instructions

Abstract

This manual introduces the core functions of the FCGS tool. FCGS provides a Java implementation of the FCGS comparison algorithm, which supports the graphical user interface (GUI) mode and command line calling mode for use of this algorithm. The user can integrate the FCGS algorithm function in their project by selecting the FCGS algorithm function interface in the reference program. The main purpose of this project is to provide the user with an example of the implementation of the FCGS algorithm so that they can better understand the processes and concepts involved with this algorithm.

Project Download

Project source code address：https://github.com/plantdna/fcgs

Project Structure

The project directory contains the source code directory fcgs and the directory data for the test data files. The source code directory of the project contains five submodules (FcgsBase, FcgsCore, FcgsService, IPidsCore, and PidsCore) and a main program module fcgs. The five submodules provide some basic interfaces and implementation for the algorithm, while the main program module provides the functional implementation of performing the comparison algorithm through the graphical interface and command line modes. In the "Data" directory, we provide six files, including microsatellite, SNP KASP, and SNP-array marker types of data, each of which provides "Source" and "Target" fingerprint data in CSV file format. The user can use these files for functional testing. The basic directory structure of the fcgs project is as follows:

```
plantdna/fcgs
        ├──fcgs -- FCGS main project
        |     ├──fcgs -- FCGS main project
        |     ├──FcgsBase -- Basic code
        |     ├──FcgsCore -- Comparison core code
        |     ├──FcgsService -- Comparison service
        |     ├──IPidsCore -- Fingerprint comparison interface
        |     ├──PidsCore -- Implementation of fingerprint comparison
interface
        ├──data -- Test data file directory
        ├──README.md
```

Operating Environment

The FCGS algorithm program runs in advanced Java Development Kit (JDK) version 1.7 (Oracle Corporation, Redwood Shores, CA, USA) or above. Supported operating systemz include: Windows, Linux, and MacOS. Before executing the FCGS algorithm, the user must install the correct version of the JDK package according to their operating system type. The download address is:

https://www.oracle.com/technetwork/java/javase/overview/index.html

Program Packaging

This program uses Java language for development and the project uses Maven for developing and packaging. The code can be edited and run using IntelliJ IDEA or development tools such as Eclipse. Therefore, it is necessary to pre-configure the parameters of the Maven plug-in for the development tool so it can use Maven to automatically download the required JAR packages and automatically compile and package the source code of the fcgs project.

The fcgs project can be packaged with development tools to export two JAR package files from the main module program. These two files provide the graphical interface mode and command line mode, respectively. It is recommended that the user end the exported JAR file name with GUI or Command (CMD) to distinguish between the two functional JAR files. The jar file description is as follows.

1) The package ending with GUI is named fcgs-gui.jar. This package runs in an operating system with a graphical interface that allows the user to select data files, set parameters, and compare them graphically. When building the fcgs-gui.jar package, the user must select the main program class file as "FcgsGuiRunner.java."

2) The package ending with CMD is named fcgs-cmd.jar. This package is mainly used for the invocation of the comparison algorithm in the command line mode. This can be compared using the given parameters and can be easily integrated into different systems to provide comparison services. When building the fcgs-gui. jar package, the user must select the main program class file as "FcgsCmdRunner.java."

Description: Development tools such as IntelliJ IDEA and Eclipse generally use the latest version.
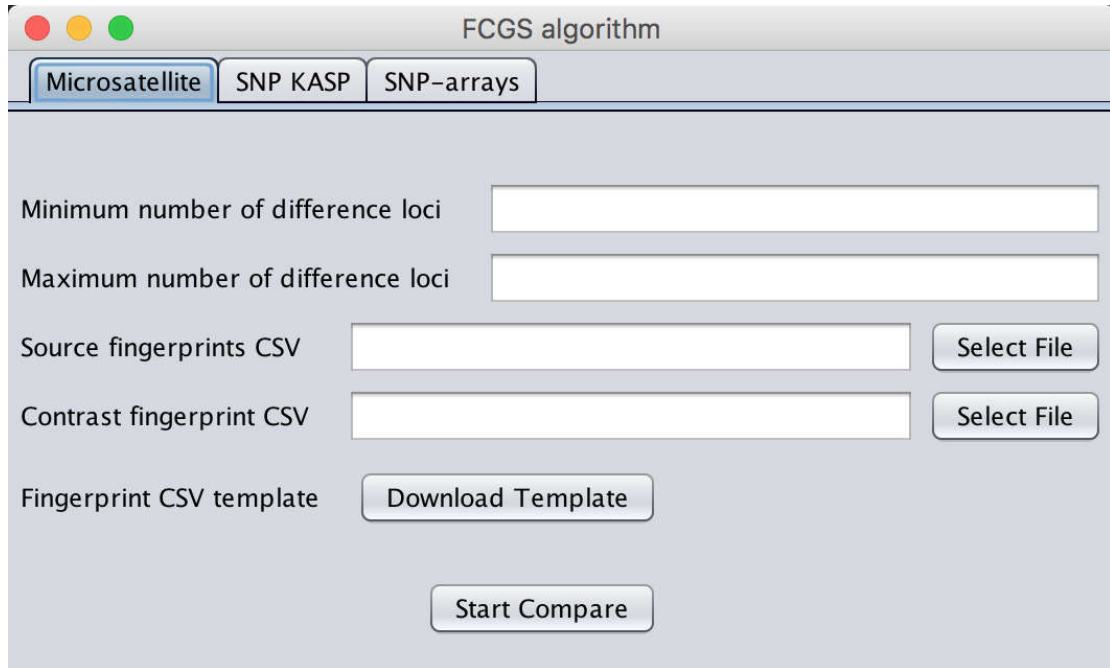
GUI Mode

Startup Command

java –jar fcgs-gui.jar

Note: If the Java environment variable is set, the user can also double-click the Fcgs-gui. jar package to execute it directly. If the file is set incorrectly, the user can use the command above to open it.

Operation Workflow

Main Interface

Three types of fingerprint comparison data are supported, as shown in the figure above.

1. Microsatellite fingerprint data comparison corresponding to the "Microsatellite" page in the figure.

2. SNP KASP fingerprint data comparison corresponding to the "SNP KASP" page in the figure.

3. SNP-array fingerprint data comparison corresponding to the "SNP-arrays" page in the figure.
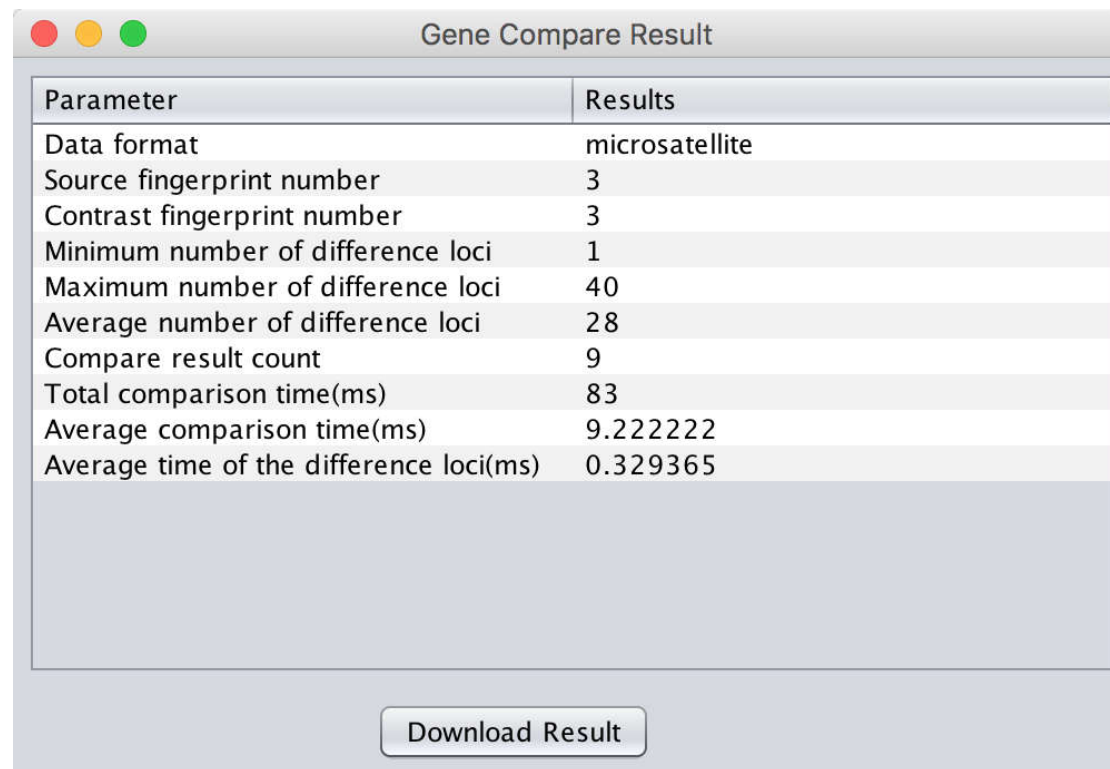
The fingerprint comparison parameters in the figure above are as follows:

Parameters	Parameter Description
Minimum number of difference loci	Used to filter difference loci in results so that only results within a given range are displayed. Parameter values may not be given.
Maximum number of difference loci	
* Source fingerprints CSV	To compare the path of the fingerprint CSV file. Input required.
* Comparison fingerprint CSV	Compare the path of the fingerprint CSV file. Input required.
Fingerprint CSV template	Download the fingerprint data template file. Different molecular marker types have different data formats.
Start Compare	Start fingerprint comparison
Parameters with * are required input	

Instructions for downloading the fingerprint template file:

Click on the "Download Template" button to download the template to the current tab page belonging to the corresponding molecular marker type format file. For example, the content of the downloaded file template in the figure above is "Microsatellite." The user must switch to the corresponding tab page to download the corresponding data template. If the data format is inconsistent with the content of the template, then the comparison operation will fail.

Comparison Results Interface

The parameters of fingerprint comparison results in the figure above are as follows.

Field	Field Description
Data format	Molecular marker data format, including: Microsatellite, SNP-KASP, SNP-arrays
Source fingerprint number	The number of fingerprints to be compared in CSV file
Comparison fingerprint number	The number of fingerprints to be compared in CSV file
Average number of difference loci	The average value obtained by dividing the sum of the difference loci in all the comparison results by the number of comparison results which can reflect the information size of the difference loci in the comparison results

Compare result count	The number of comparison results is obtained after comparing two fingerprint lists
Total comparison time(ms)	Time to call FCGS algorithm for comparison (ms)
Average comparison time(ms)	Average time spent per comparison result (ms)
Average time of the difference loci(ms)	Average time spent per difference locus (ms)

Click on the "Download Result" button to download the comparison result CSV file.

Description of the comparison result field:

$$\text{Average comparison time} = \frac{\text{Total comparison time}}{\text{Compare result count}}$$

$$\text{Average time of the difference loci} = \frac{\text{Average number of difference loci}}{\text{Total comparison time}}$$

CMD Mode

Startup Command

java -jar fcgs-cmd.jar microsatellite microsatellite-source.csv microsatellite-target.csv 0 40

The command format

java -jar fcgs-cmd.jar [Data format] [Source fingerprints CSV] [Comparison fingerprint CSV] [Minimum number of difference loci] [Maximum number of difference loci]

The parameter types of the command are described as follows:

Parameters	Parameter Description
* Data format	Molecular marker data type, optional parameter values: Microsatellite, SNP-KASP, SNP-arrays. (Required)
* Source fingerprints CSV	The path of the comparison fingerprint CSV file. (Required)
* Comparison fingerprint CSV	The path of the comparison fingerprint CSV file. (Required)
Minimum number of difference loci Maximum number of difference loci	These two parameters must be both given or not given to filter the alignment results
Parameters with * are required input parameters	

After executing the comparison command, the comparison algorithm will save the generated comparison results to the same-level directory in the path of the fcgs-cmd.jar package and store them in CSV data file format.

Help

If the user is not clear about the abovementioned documents or requires any technical support, he or she can send an email to jiangbinboy@126.com and contact our technical staff.