Genesis-Py user manual: Group 2

Erik Prinz Charles Lu Nandi Bau

May 2018

1 Introduction

This manual describes Genesis-Py, a program created to implement the functionality of Genesis using Python. It generates Principal Component Analysis (PCA) and admixture graphs for academia. The program operates on data output by other tools such as eigenstrat, the SNPRelate package for PCAs and Admixture, and CLUMMP.

Note that this manual is only for **Genesis-Py v0.1.0b** and its current functionalities. This version of the program runs on **Windows 10, Linux Ubuntu 16.04 LTS and MacOSX**.

Git repository

The Git repository for our project can be found at `https://github.com/charleslu0529/Genesis-Py/tree/master`

2 User manual

2.1 Installation of Genesis-Py

Genesis-Py **requires at least Python v3.6.5** installed on your computer. To check if you have the latest version of Python installed, you can open your operating system's command line and enter:

```
python –V
```

Visit `https://www.python.org/downloads/` to obtain and download the latest version of python.

Genesis-Py depends on the **pyplot** and **wxpython** libraries. Make sure you have the latest version of pip, then simply enter into your command line:

```
pip install –U wxPython
```

Note that if you are using a Linux system, you must instead enter:

```
pip install –U \
  –f https://extras.wxpython.org/wxPython4/extras/linux/gtk3/ubuntu-16.04 \
  wxPython
```

To **install pyplot**, enter the following commands into your command line:

```
python –mpip install –U pip
python –mpip install –U matplotlib
```

No special installation is required for **Genesis-Py** and the file can simply be executed from the root folder of the program. This has been tested on Windows 10 and Ubuntu 16.04 LTS operating systems.

2.2 Running Genesis-Py

For Windows 10 and Linux Ubuntu 16.04LTS users, double click on the latest version of Genesis-Py in the root folder of the program. If this does not work, open terminal in Linux or cmd in Windows, make sure you are located in the program's root directory and run the following:

```
python Genesis-Py_v0.1.0b
```

2.3 Known Bugs

- On Windows 10,

2.4 Supported input data formats

For admixture charts, Genesis-Py supports the input of admix data and .phe files. For PCA charts, the program supports .evec and .phe files.

2.5 Menu Structure

Genesis-Py includes a menu and a toolbar for functionality. Menu items include the *File* drop down window and the toolbar offers 10 graphic buttons with associated functions. The basic overview of the menu and toolbar is explained in the following sections with complementary images.

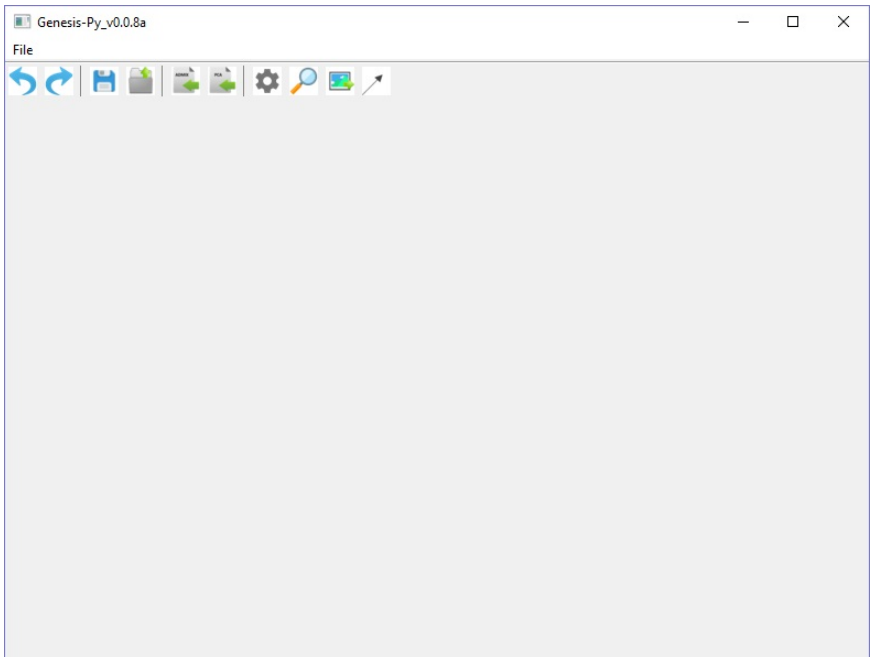

Figure 1: GUI Layout

The **toolbar** contains 10 prototype buttons, each representing a specific function you may use.

Figure 2: toolbar Layout

The functionalities linked to the buttons can be found and detailed in table 1 below. Note the implementation status is for Genesis-Py v0.1.0b.

Table 1: Table showing all toolbar buttons and functionality implementation

Button	Description	Function	Implementation Status
1	Undo	Undo last operation	Incomplete
2	Redo	Redo operation	Incomplete
3	Save	Saves data	Function Complete, not integrated
4	Load	Loads data	Function Complete, not integrated
5	Import admix	Import admixture data	Complete
6	Import PCA	Import PCA data	Complete
7	Settings	Open settings menu	Incomplete
8	Magnify	Zooms in on image section	Incomplete
9	Export image	Opens image export format dialogue	Can be done from Graph Window Only
10	Add arrow	Adds arrow with label to graph	Incomplete

The layout of the **menu** can be seen here, with the functionalities explained in table 2 below.

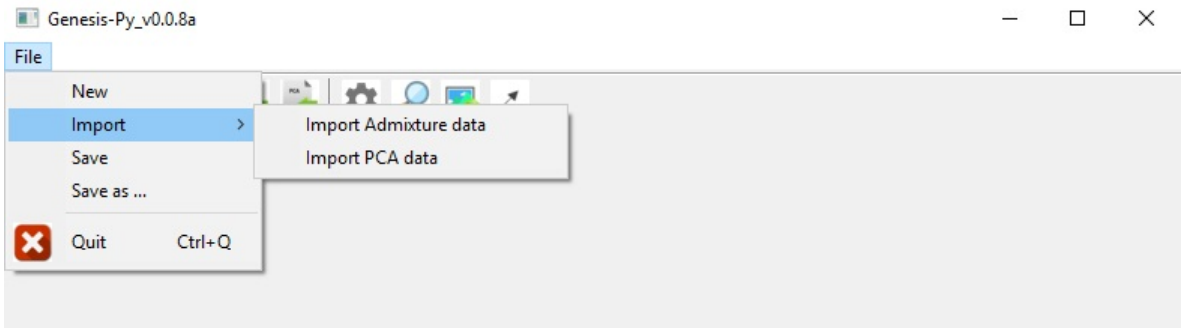

Figure 3: Menu layout

Table 2: Table showing menu item functionality and implementation status

Menu Item	Function	Implementation Status
New	Opens blank project	Incomplete
Import	Opens Import sub-menu	Complete
Import Admixture Data	Opens admixture data import dialogue	Complete
Import PCA Data	Opens PCA data import dialogue	Complete
Save	Saves current graphing data	Function Complete, not integrated
Save as ...	Saves current graphing data into new file	Function Complete, not integrated
Quit	Quit Genesis-Py	Complete

2.6 Program shortcut

Program **keyboard shortcuts** are provided and there is currently one in demonstration to allow for ease of use. **You may hold 'Ctrl & Q' to exit the application quickly**. In future, we plan to implement a saving dialogue before the exit keyboard shortcut is applied.

2.7 Admixture graphs

2.7.1 Data input format

Genesis-Py requires 2 input file types from the user to plot an admixture graph:

- An admixture data file, which contains the estimated ancestral proportions of each individual on each new line. This is produced by the *Admixture* and *CLUMMP* programs.
- A phenotype file in the *.phe* format, where each line represents an individual and the first two columns offer identity and subsequent columns are labels.

2.7.2 Inputting data

To input the files, click **Import** menu item then **Import Admixture Data**. Follow the subsequent instructions. You may also use the included toolbar button.

Note that in the current version of Genesis-Py, **.fam** files are not supported in admixture plotting.

2.8 PCA Plots

2.8.1 Input data format

For PCA plotting, Genesis-Py inputs 2 types of data files, the eigenvector, **.evec**, and phenotype, **.phe** files. Both file types are compulsory for plotting a PCA graph.

2.8.2 Inputting data

To input the files, click **Import** menu item then **Import PCA Data**. Follow the subsequent instructions. You may also use the included toolbar button.

2.8.3 Saving and loading

To save or load files, click on the save buttons present on the toolbar and in the menu. Follow appropriate instructions.

2.9 Appearance options

The colour elements of the graphs may be changed through the settings tool button.

3 Technical manual

Genesis-Py was designed using the Model-Controller-View(MCV) framework. These separate functionalities into classes under a either the model, controller or view. The model represents the data rules and logic of the system. The view represents the output and the user interface, while the controller represents the user input and commands to manipulate the model.

3.1 Model

The model classes in Genesis-Py are the graphing tools. The model classes for PCA and admixture graphing take in data controlled by the controller buttons in the user interface and operates on them to output appropriate graphs to the view. The model classes are listed below.

- PCA.pca - handles the generation of PCA graphs
- Admixutre.AdmixtureInstance - handles the generation of admixture graphs

3.2 Controller

The controller classes responsible for converting user inputs to commands for the model classes are included in the view classes due to them being represented by buttons on the graphical user interface. These are implemented and documented already in the wxPython user interface library that Genesis-Py utilises. The controllers are included in the above model classes due to them being lightweight.

3.3 View

The Genesis-Py application is the main view type class. This represents the main source of information for users, and provides the interface through which the users may call the controllers. The view classes, apart from the main program, were organised into one folder in an attempt to

- Genesis-Py - the program is the main view class
- Admix_View.import_Admix - class corresponding to admixture importing
- PCA_View.import_PCA - corresponding to PCA importing
- Settings_View - handles