# G-matrix Simulator Home Version Instructions

By Adam G. Jones

Texas A&M University

## Introduction

The ***G-matrix Home Version*** program is an individual-based Monte Carlo simulation of a sexually reproducing, diploid population. The simulation deals with the case in which two quantitative traits are determined by a suite of additive, pleiotropic loci. The traits are evolving in response to an individual selection surface with a single peak. If the peak remains stationary, the population will stay near the peak and will be affected mainly by stabilizing selection, genetic drift and mutation. If the peak moves, then the population will also be subject to directional selection. This program (with some slight modifications along the way) was used to generate the data for the following papers:

Jones, A. G., S. J. Arnold, and R. Bürger. 2003. Stability of the **G**-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* 57:1747-1760.

Jones, A. G., S. J. Arnold, and R. Bürger. 2004. Evolution and stability of the **G**-matrix on a landscape with a moving optimum. *Evolution* 58:1639-1654.

Jones, A. G., R. Bürger, S. J. Arnold, P. A. Hohenlohe, and J. C. Uyeda. 2012. The effects of stochastic and episodic movement of the optimum on the evolution of the **G**-matrix and the response of the mean to selection. *Journal of Evolutionary Biology* 25:2210-2231.

The details of the model are explained at great length in those papers. This document is simply a guide to running the simulation program and interpreting the results.

## Running the Simulation Program

The program should run on any machine running any Windows operating system from Windows XP onward. It will not run on any other operating system.

***Step 1:*** Unzip the folder containing the program.

***Step 2:*** Double click on the ***G-matrix_Home_2016*** icon.

***Step 3:*** Set the parameters. To access the parameter input window, click on the button with the yellow key icon (or click the "Simulation Parameters" menu item under "Run"). A new window should pop up. The current parameter values will be displayed, and you will have the ability to change most of the parameter

values. The parameters and their significance are listed in Appendix 1, below. When you are done changing parameters, press the "Okay" button.

***Step 4:*** Decide whether or not you would like the program to save the output of each run in a separate file as the program runs. If so, examine the toolbar and find the checkbox labeled "Save Individual Runs". Check the box, and a save dialog will pop up. Choose a file location and a filename. The filename should be a basic, descriptive filename, and each file generated by the program will be appended with a different suffix indicating the run number. If you ran a simulation with 5 runs, for instance, and chose the filename "gmatrix", the program would generate files named "gmatrixRun1.csv", "gmatrixRun2.csv", and so forth. It also saves a file that ends with "Summary.csv", which contains means for each run, a mean across all runs, and the standard deviation in means across runs. Note that these files are in comma-delimited format, so they can be imported into Microsoft Excel (just by double clicking on them) and into R. To import each file into R as a data frame, use the command **read.csv("filename.csv")**.

After your simulation runs complete, you may want to perform a new set of runs with a different set of filenames. To do so, **uncheck the box and check it again to select a new filename**.

***Step 5:*** Run the simulation. Press the small arrow button to run the simulation. Alternatively, select "Run Simulation" from the "Run" menu.

***Step 5:*** Watch the graphical output. There's no way to save the graphical output, so if you want to save a copy, just hit your computer's "Print Screen" button and paste it into a drawing program.

***Step 6:*** Look at the text output. The 2016 version of the program produces fewer columns of data than the industrial strength but more confusing G-matrix program, which is also available online. The definitions of the variables are listed in Appendix 2. Note that when the program automatically saves runs in ".csv" format, the "~" symbols are replaced with "d" and all parentheses are removed to maximize compatibility with downstream programs.

## Appendix 1: Parameters

**Number of Loci:** (represented by *n* in the papers) The number of pleiotropic, additive loci determining the two traits (**must not exceed 100**).

**Number of Offspring per Adult:** (*B*) The number of offspring produced per adult. The mating system is monogamous, so this number is also half the number of offspring produced per female. The maximum total number of offspring allowed by the simulation is 20,000, so **the product of *B* and *N* (see next entry) must be less than 20,000**.

**No. Females in Population:** (*N*/2) The number of females in the population at carrying capacity. Because the simulation holds the sex ratio constant, this number is also the number of males in the population at carrying capacity. Thus, twice this number is the total adult population size (*N*).

**Mutation Rate (per 1,000,000):** ($\mu$) The number of mutations per 1,000,000 meioses. A value of 200 gives a mutation rate of 0.0002. The value entered for this parameter **must be an integer**, so the smallest permissible mutation rate is $10^{-6}$.

**Number of Traits:** This model uses two traits, so **this parameter must not be changed**.

**No. Experimental Generations:** The number of generations during which you want the program to tally and report data about the population.

**Mutation Variance (in hundredths):** ($\alpha_1^2, \alpha_2^2$) Sets the mutational variances for the two traits. Higher mutational variances will result in mutations with larger absolute effects. A value of 5 is a mutational variance of 0.05. **Only use whole numbers in this box.**

**Strength of selection (w2):** ($\omega_{11}$, $\omega_{22}$) Sets the width of the selection surface for each trait. Larger values represent weaker selection. Typical values should be between about 5 (for very strong stabilizing selection) and 100 (for weak stabilizing selection).

**No. Initial Gens (Stationary Opt):** The number of initial generations of the simulation. During these generations, the bivariate optimum does not move and the population accumulates variation due to new mutations. The goal is to reach a mutation-stabilizing selection-drift equilibrium. A few hundred generations is probably enough unless you choose to check the "Begin with no genetic var." box. Then you will probably want 5000-10000 initial generations to generate some genetic variation.

**No. Complete Replications:** The number of complete simulations to perform under the current parameter combinations. The results will be concatenated in a single text window, and at the end of the replicates, means and standard deviations across replicates will be calculated. These means, typically across 20 replicates, are the values reported in most of the tables appearing in the papers mentioned above.

**Mut. Corr. Traits 1, 2 (hundredths):** ($r_\mu$) The mutational correlation. This value must be an integer (e.g., 50 produces a mutational correlation of 0.50), and too large an absolute value will cause population extinction. The value can be positive or negative, so values between 95 and -95 are reasonable. Values as extreme as 99 and -99 are technically permissible, but may cause problems, especially in small populations.

**Sel. Corr. Traits 1, 2 (hundredths):** ($r_g$) The strength of correlational selection. Determines the extent to which interactions between the two traits are important in terms of fitness. Like the mutational correlation, the values of the selectional correlation can be between -99 and 99 (corresponding to -0.99 and 0.99).

**Str. sel'n after peak shift (w2):** ($\omega_{11}$, $\omega_{22}$) The strength of selection will shift to this value as soon as directional selection begins in the case of an optimum that moves more than once. In the case of an optimum that moves only once, the strength of selection will switch to this new value at the experimental generation indicated for the peak shift. Thus, this parameter should be used to impose a single, major change on the individual selection

surface at a specified generation. For directional selection, it's usually better to set the parameter to the same value as the "Strength of selection" above.

**Sel. Corr. after sel'n (hundredths):** ($r_g$) Like the previous entry, this parameter allows the user to impose a single shift in the strength of correlational selection. The value entered here must be an integer from -99 to 99.

**Trait optima shift (real number):** ($\Delta\theta_1$, $\Delta\theta_2$) The distance the optimum for each trait moves during each peak shift. The environmental standard deviation is 1, and in our models the phenotypic standard deviation would typically have a value between about 1 and 2. Most populations will tolerate only small peak movements averaging something on the order of 0.01 per generation.

**Calc. G every ___ generation(s):** The interval between calculations of the variables of interest. **This should be left at 1.** Otherwise, the program may exhibit strange behavior.

**Move peak at generation:** Sets the generation at which the peak shifts in the case of a peak that moves once. Sets the interval between peak shifts in the case of a peak that moves repeatedly. Thus, this parameter and the "Trait optima shift" parameters fully describe the movement of the bivariate optimum.

**Calc. equil. values over ___ generation(s):** The program calculates the means of a small subset of variables during the stabilizing selection generations. This parameter determines the number of generations over which these values are averaged. **It's probably best to leave this parameter at its default value.**

**Draw G every ___ generations:** How often you would like the **G**-matrix to be drawn on the graphical output window. A value of zero here will result in a new drawing of the G-matrix every generation.

**Drawing scale constant:** If the drawn **G**-matrix is too small or too large, change the value of this parameter, which must be a positive integer.

**Size of image (X, Y dimensions):** Use these parameters to change the dimensions of the canvas upon which the **G**-matrix is drawn.

**Location of starting optimum (x, y):** This parameter determines where on the drawing canvas the starting phenotypic mean of the population will appear.

**No. Initial Gens (Moving Optimum):** The number of generations of directional selection preceding the experimental generations during which values are tallied. These generations only apply to the case in which the peak moves repeatedly and they allow the population to equilibrate to the moving optimum if desired. No data will be reported during these generations.

**Move Peak Once, Move Peak Repeatedly:** Select the button that gives the desired behavior.

**Begin with no genetic var.:** Use this check box to decide whether the population starts with some randomly generated variation or not. All simulations in the papers listed above started with no genetic variance, so this box should be checked to replicate those results. However, if this box is checked, each simulation should start with 5000 to 10000 initial generations to allow mutations to introduce some genetic variation in the population. If this box is unchecked, a few hundred initial generations should be enough.

## Appendix 2: Output

***Note that in the .csv files, the parentheses and hyphens will be omitted and the "~" symbol will be replaced with a "d".**

**Gen:** The current generation of the simulation.

**Opt(1), Opt(2):** The position of the optima for traits 1 and 2.

**z-bar(1), z-bar(2):** The phenotypic means for traits 1 and 2.

**P(11), P(22):** The phenotypic variances for traits one and two.

**P(12):** The phenotypic covariance.

**G(11), G(22):** The additive genetic variances for traits 1 and 2. These are $G_{11}$ and $G_{22}$, the diagonal elements of the **G**-matrix.

**G(12):** The additive genetic covariance, or $G_{12}$, the off-diagonal element of the **G**-matrix.

**r(g):** The additive genetic correlation between traits 1 and 2.

**lambda(1), lambda(2):** The two eigenvalues of the **G**-matrix.

**G-angle:** The angle of the leading eigenvector of the **G**-matrix. This angle is reported in degrees from -90 to 90 on a plot with trait 1 on the x-axis and trait 2 on the y-axis. The x-axis has an angle of 0 degrees.

**~G(11), ~G(22), ~G(12), ~r(g):** The absolute value of the change in $G_{11}$, $G_{22}$, $G_{12}$, and $r_g$ between the current generation and the previous generation. These values provide one way to look at the stability of the **G**-matrix.

**~lambda(1), ~lambda(2):** The absolute change in the eigenvalues of the **G**-matrix between the current generation and the previous generation.

**~G-angle:** The absolute change in the angle of the leading eigenvector of the **G**-matrix between the current generation and the previous generation.

**Beta(1), Beta(2):** The values of the selection gradients on traits 1 and 2, calculated from the equation given in Lande (1980) – see equation (4) in Jones et al. (2004).

**Final rows of the output:** At the end of each simulation run, the means of most values are reported for that run. If there are many runs, then at the very end of all the runs (which will all be in the same text window, one after another), there will be a mean of means row and a standard deviation of means row. The mean of the means just averages the means across runs, and the standard deviation of the means is the standard deviation of these means across runs.