GeneScore
Counting Easier

Table of content

1) Intro

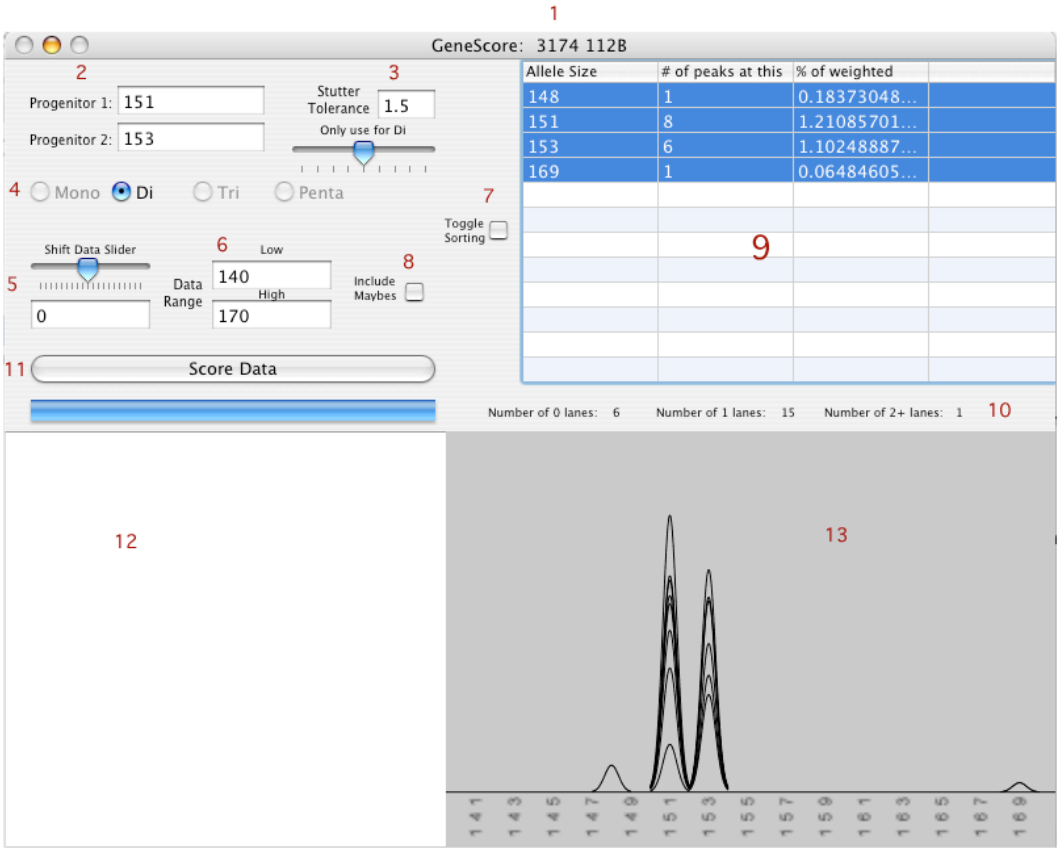
The need to automate calling of peaks arised because of how time consuming it is to print and score a large number of chromatograms. GeneScore was created to reduce the time of calling fragments and remove possible user error in calling chromatograms differently.

2) Synopsis

By importing data from chromatograms, such as size, height, and other misc. info, of alleles, GeneScore can make accurate guesses as to which Alleles are significant. This method of determining which alleles are significant is known as "Calling" or "Scoring".

3) Layout

1. File Name
2. Progenitor Insertion Box
3. Stutter Tolerance Adjuster for Dinucleotides
4. Locus Selector
5. Allele Size Manipulator
6. The Range
7. Toggle between “By Allele” and “By Lane” Sorting
8. Maybe Toggle
9. Data Table
10. Alleles Per Lane Counters
11. BIG BUTTON TO PRESS
12. Extra Data Info Box
13. Graph View



1) File Name

On the title bar, when you have a data file open, the name of the file is displayed here. Be careful when scoring your data to make sure you are scoring what you indeed want to be scoring

2) Progenitor Insertion Box

In the Progenitor 1 and 2 box you will want to insert what you believe to be the progenitors of the data you are scoring. If it is Homozygous, simply enter the same number into both boxes. If you do not know your progenitors, insert numbers close to where you think they will be. The program can use any such information to help in the scoring process. In some data samples knowing the progenitors matters, in many it does not.

3) Stutter Tolerance Adjuster for Dinucleotides

The use for this box is to adjust the tolerance of detecting alleles in stutter bands common in Dinucleotides. Long studies have shown that then accepted standard is 1.5, but for your purposes, you might wish to use a different tolerance. It is strongly not recommended you use any value under 1.0. This box has no effect on Monos, Tris, or Pentas.

4) Locus Selector

Use this to select the locus at which you are working with. Initially, only Dinucleotides are supported. Dinucleotides that were tested on, and used for creating the program are D2S123, D5S346, and D17S518.

5) Shift Data Slider

With the box or the slider, you can increase or decrease the size of the peaks(aka shift the data left and right) for an entire experiment. Sometimes an experiment can be off, and this will allow you to modify an entire experiment. It is not suggest to use this feature, and when you do, to use it very carefully. This is non-destructive way to manipulate your data, but it is doing just that, manipulating your data and it may give you results you do not want.

6) The Range

When scoring data, you are often looking at a much narrower range than the entire PCR allele spectrum. The Low is the lower bound you wish to look at and the High is the upper bound. What you plan to look at and score should fall between these two. The less sure you are of where your data is, the larger range you may wish to use. You can always refine your range after you score your data to trim out things that you are not interested in.

7) By Lane and By Allele toggle

When viewing your data in the table, it is often helpful to change how you look at your data. The default sorting/view is By Allele. Clicking this toggles changes instantly.

8) Maybe Toggle

When scoring alleles and data, there are often borderline cases. These are things we call Maybes. Clicking this toggles whether you want to include maybes as a scored call, or as a not scored called. Clicking changes representation in the Chartview(13) in real time, but not revaluate your table in real time. If you wish to see the changes to your table with or without maybes after already scoring, you will need to reclick the Score Data button(11). Examples of Maybe cases can be viewed in Appendix A

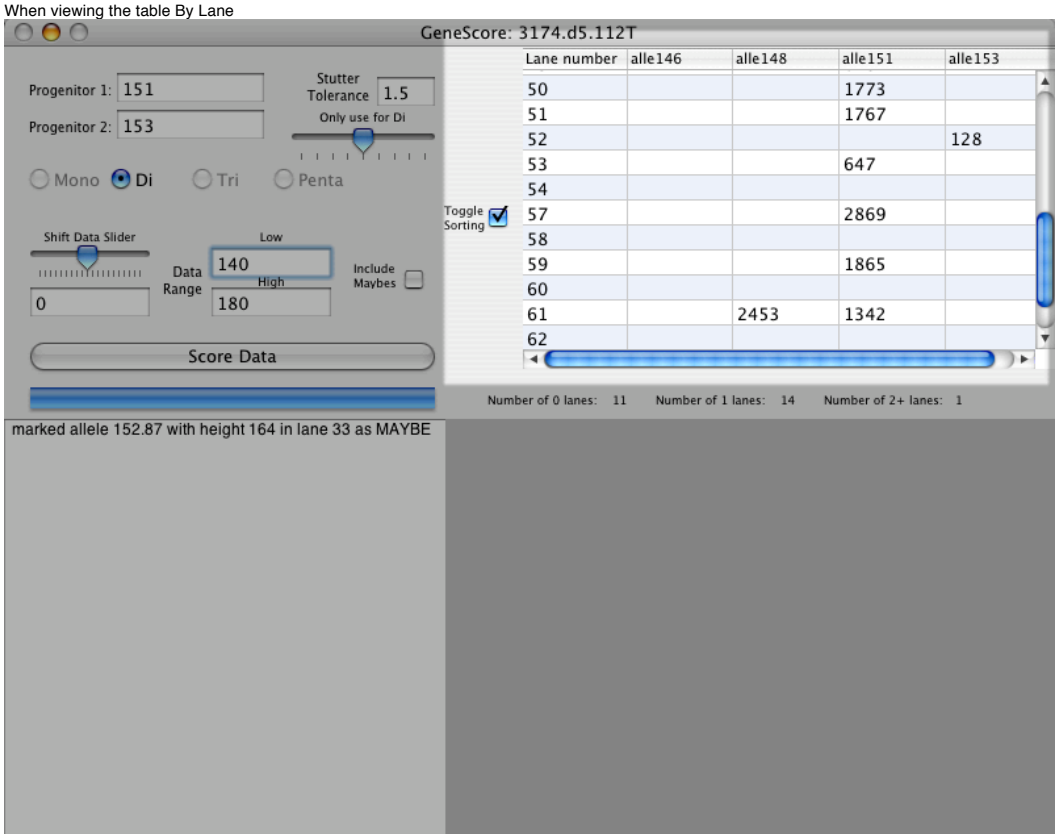
9) Data Table

This is where all the information regarding the scoring and analysis of your data goes. There are two ways to view this table: By Allele and By Lane. The default is By Allele. The table will be empty when you first run GeneScore.

In the By Allele view, there are 3 Columns: *Allele Size*, *# of peaks at this height*, and *% of weighted*.
After scoring a data, you will get a list here of data. Any alleles that have been called will show up in this tabulated form. The *Allele Size* column will have all the sizes of the alleles called. The *# of peaks at this height* column will tell you how many alleles in your entire experiment (aka your data file) were called at that size. The *% of weighted* is column is used to get an idea of what the average allele at that size stands up to everything else scored. The specific formula can be found in Appendix A. Clicking on a row will bring up a representation of all the alleles at that size selected in the Chartview(13).

In the By Lane view, the rows are each and every lane that was in the data file. The columns are those that which were rows in the By Allele view. Any alleles that were scored in a lane will show up in the appropriate column. The number is the height of the allele scored. Clicking on a lane will show you everything in that lane in the Chartview (13). Scored alleles will show up black while non-scored alleles will just be dimmed.

When viewing the table By Lane



10) Alleles per Lane Counters

These 3 numbers show you how many lanes have 0, 1, or 2+ alleles in them. This is useful for quantification of your experiment purposes.

11) BIG BUTTON TO PRESS

This is exactly what it sounds like. When you press this button, it checks all the parameters you have given it to make sure you don't violate anything. It then scores, or rescores, the data. It is necessary to repress this button if you wish to change your scoring based on maybes or without maybes. This may change in future versions, but for now you have to repress

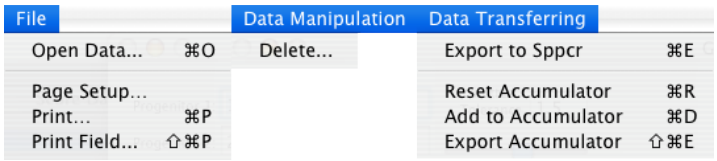
12) Extra Data Info Box

In the process of scoring data, information such as what allele where got marked as a maybe shows up here. Using this information, you can determine whether or not you wish to or not include Maybes in the scoring of your data using the Maybe toggle(8).

13) Chartview

This is where the data from your Data Table(9) is shown in a graphical representation similar to a chromatogram. This is not an exact representation, but an ideal one (if all peaks were perfectly formed). Its use is in determining relativity to other peaks, and for getting the gist of what a lane looks like. It works in both By Lane, and By Allele view.

4) Menu Options



A) File Menu

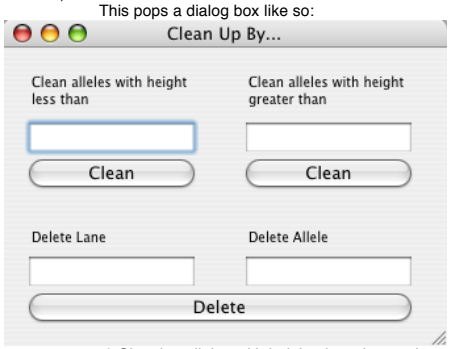
The File menu is home to where all your actions to interact with GeneScore and the Mac OS X system are.

1) Open Data...
   This allows you to select a file to be scored. The files must be in the proper format. Formatting can be found in Appendix A.
   GeneScan's export table function should properly export everything in the proper format. An example can be found in Section 5.
2) Page Setup...
   Apple's default Page Setup Dialog. Please see Apple's Documentation on how to use this.
3) Print...
   This will print the entire window except the Chartview. Support for printing Chartview is something that might be added in the future.
4) Print Field...
   This allows you to print a single aspect of the window, such as the Data Table(s).

B) Data Manipulation

The Data Manipulation menu hosts the options to manipulate your scored data. This is used to tidy up markings you feel GeneScore did not call correctly.

1) Delete...
   This pops a dialog box like so:



   a) Cleaning alleles with heights less than a size will remove all alleles with heights less than the number you put in the box.
      It does not explicitly remove them from the data set, but it does remove them from your results.
   b) Cleaning alleles greater than a size works just like less than a size, but vice versa.
   c) Delete by Lane, and Delete by Allele work in conjunction with each other. If you choose to just enter a Lane number, it will delete that entire lane.
      If you enter just an allele, it will delete all alleles at that size. If you enter both a lane and an allele, it will delete an allele that size in that lane.

C) Data Transferring

The Data Transferring menu allows for the exporting your data tables into a format easily read by SPPCR 2.0.

1) Export to Sppcr
   This prompts you for an estimated g.e., then prints the result into a separate window for you to copy and paste.
   To use this information with SPPCR 2.0, just copy everything in the window, open SPPCR 2.0, paste, then press Return.

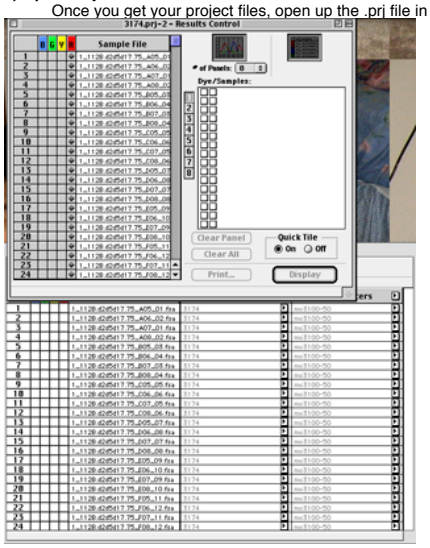
This next section has to do with the accumulation of multiple data sets. There are times when GeneScan's data is split up among 2 or more projects. Instead of trying to recreate the projects to include everything, you can use GeneScore's accumulation feature.

1) Reset Accumulator
   Before accumulating a new set, reset the accumulator so you do not have false data in there. A reset accumulator is fresh, and has nothing in it.
2) Add to Accumulator
   This will prompt you for a g.e. for the run you just scored. It will then add this to your existing, or freshly reset, accumulator.
3) Export Accumulator
   This will print to a window all the runs you have accumulated so far, as if they were a single data set, to be used with SPPCR 2.0.

5) A Sample Step-Through

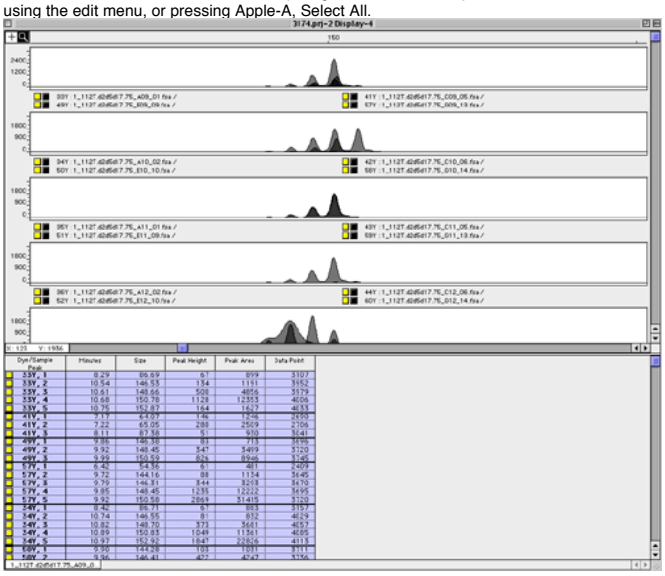
1) Open Project in GeneScan.

Once you get your project files, open up the .prj file in GeneScan. You should get 2 screens like this:

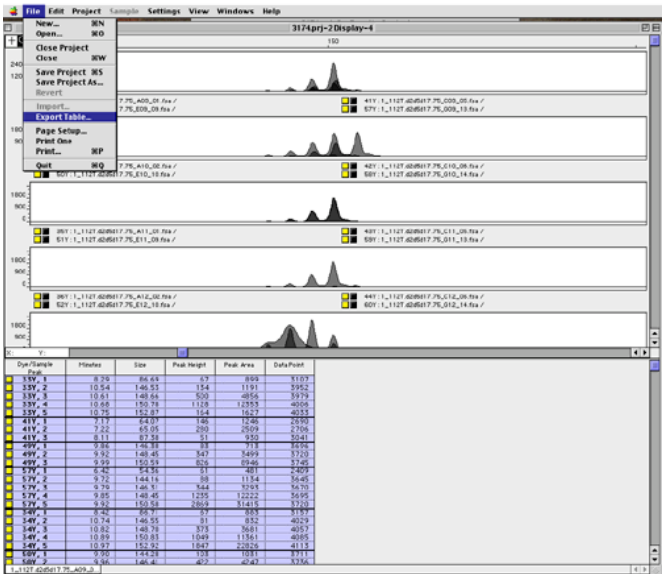


2) Export Table.

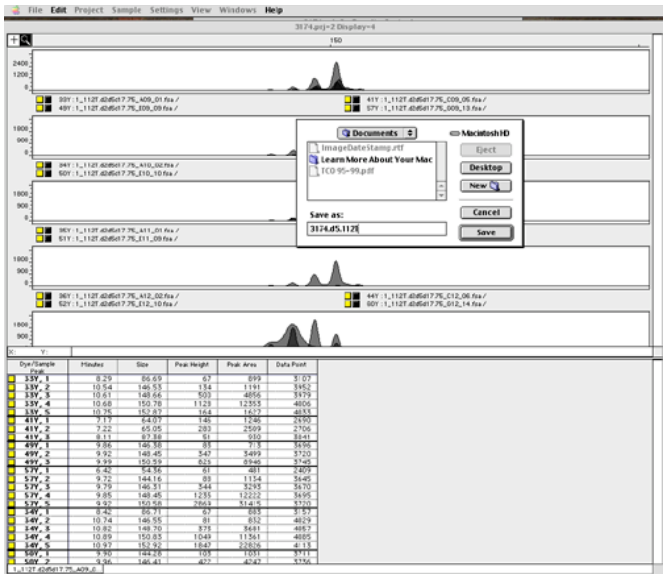
The next thing you will want to do is to select all the lanes you will want to score. After selecting those lanes, click the Display button. It does not matter the order or location of these lanes. Do not select anything besides the locus you want to score. Aka, do not include the markers. Click on the table below the chromatograms. Then either using the edit menu, or pressing Apple-A, Select All.



Once you have done this, go to the File menu and select "Export Table...".

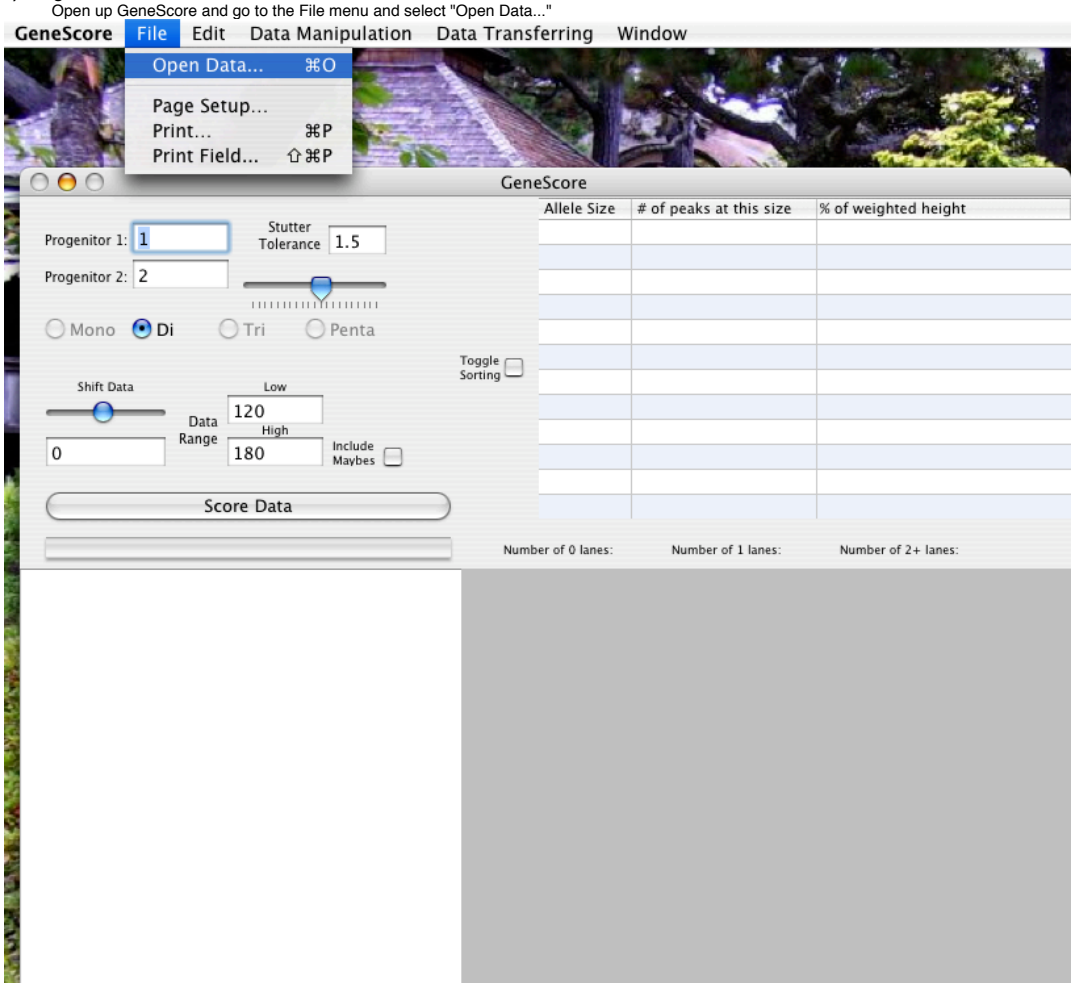


Be sure to name your file something meaningful. A name similar to your project name, such that if you need to return to GeneScan, finding the project will be painless.



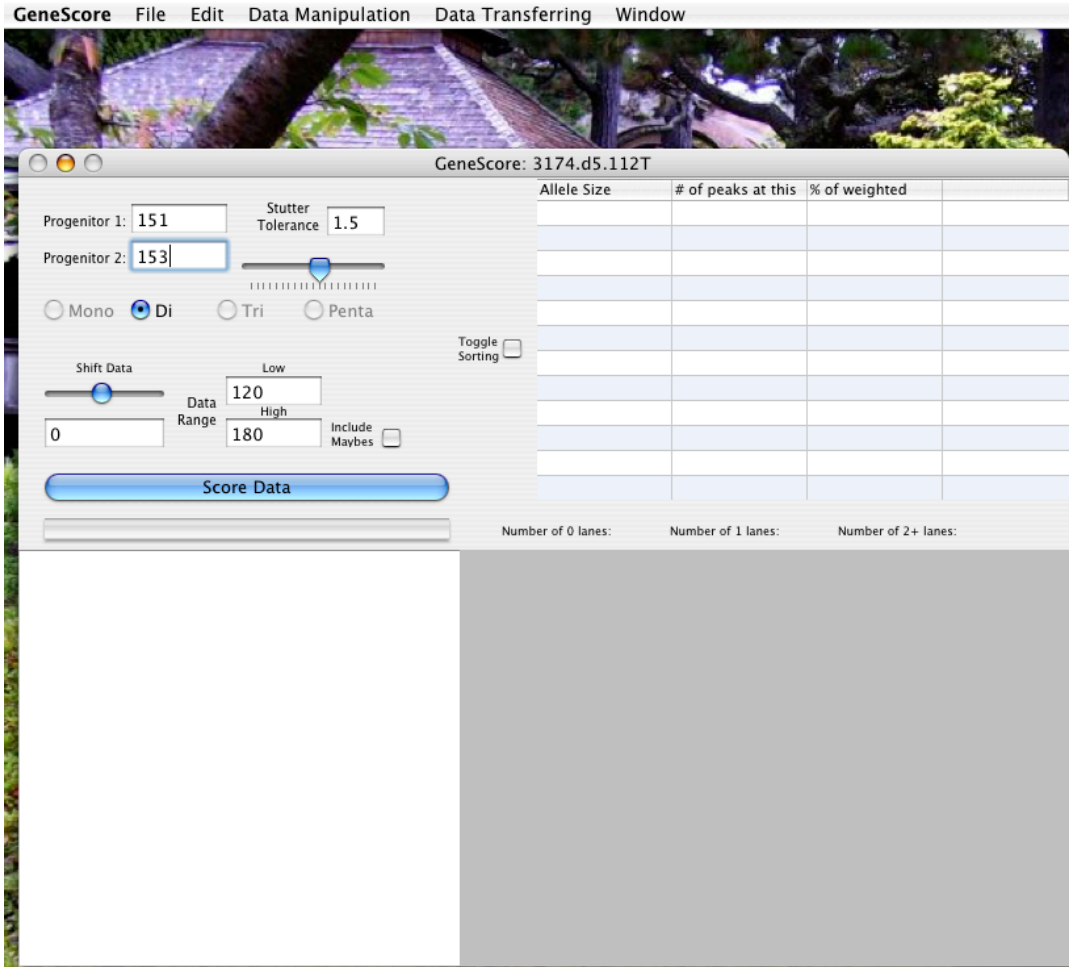
3) Bring Tables into GeneScore.

Open up GeneScore and go to the File menu and select "Open Data..."



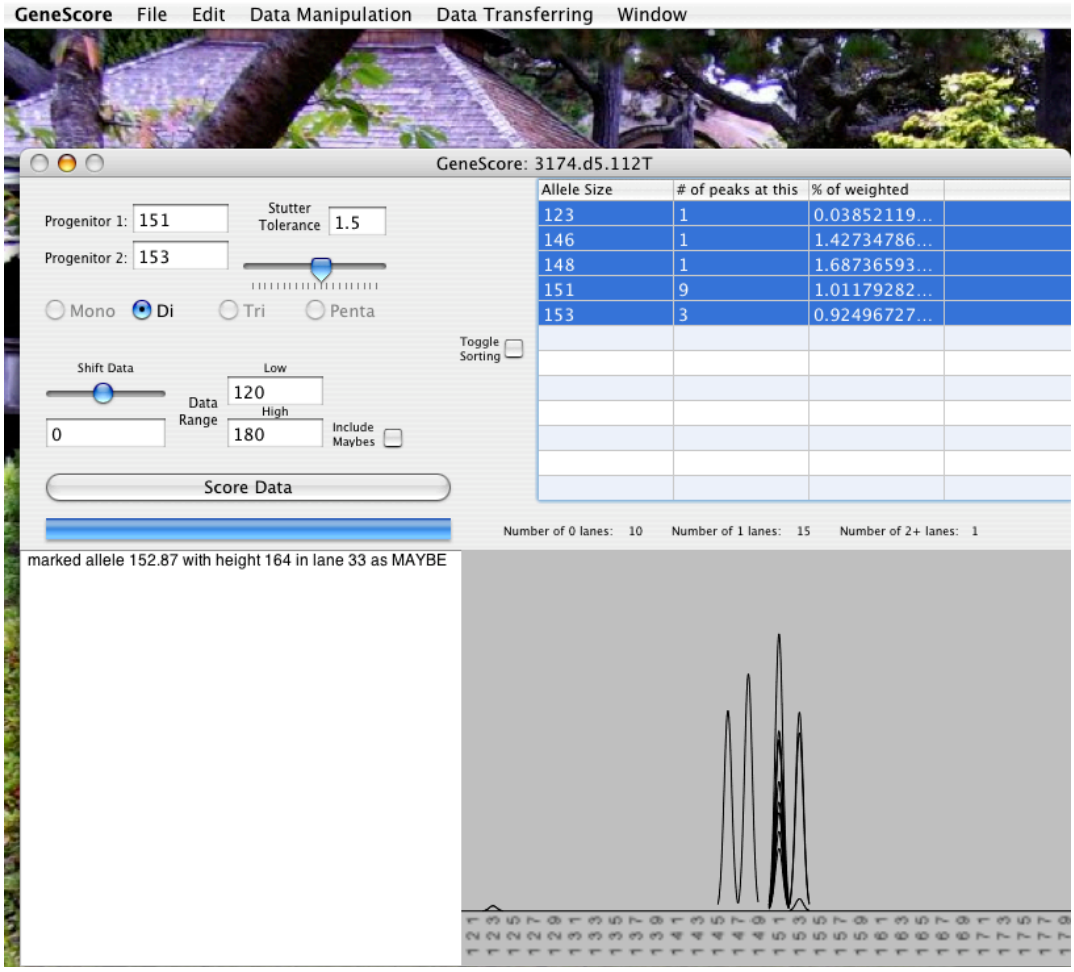
4) Parameters.

After opening your file, you will then want to modify the parameters to match your experiment. If you do not know the progenitors of your experiment, set them close to what you expect them to be. If your progenitor is homozygous, enter the same number in both boxes. Be sure to set the Data Range to the upper and lower bound you wish to look at. The less sure you are of your bounds, the larger your range should be. Rescoring your data after modifying parameters is perfectly okay.



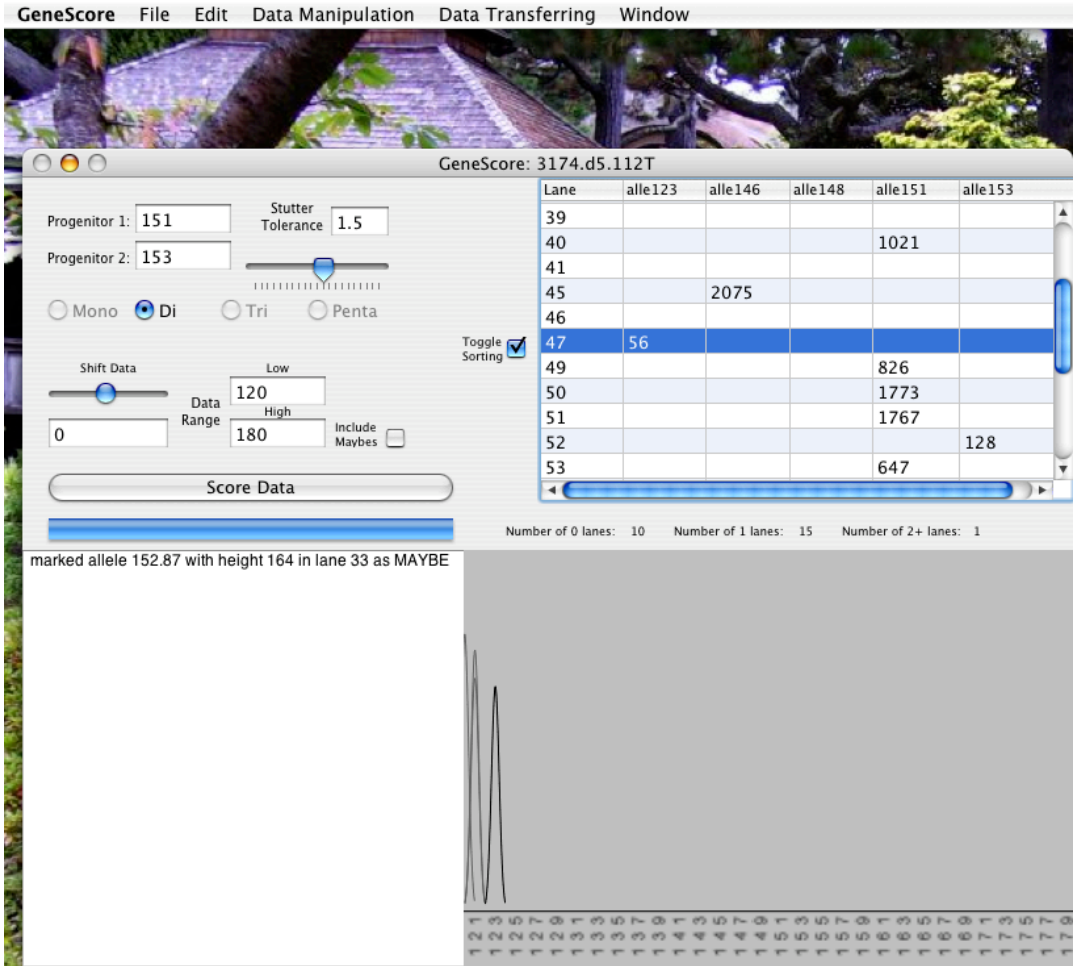
5) Scoring.

After clicking the Score Data button, you will see the progress bar move as it scores your data. Really large sets of data will take longer than smaller sets. The time it takes to score a data set increases linearly with the amount of data. The default view of the table is By Allele and shows the most useful information. Selecting items in the table will pop them up in the chart view below. This gives you a good representation of your data overall.

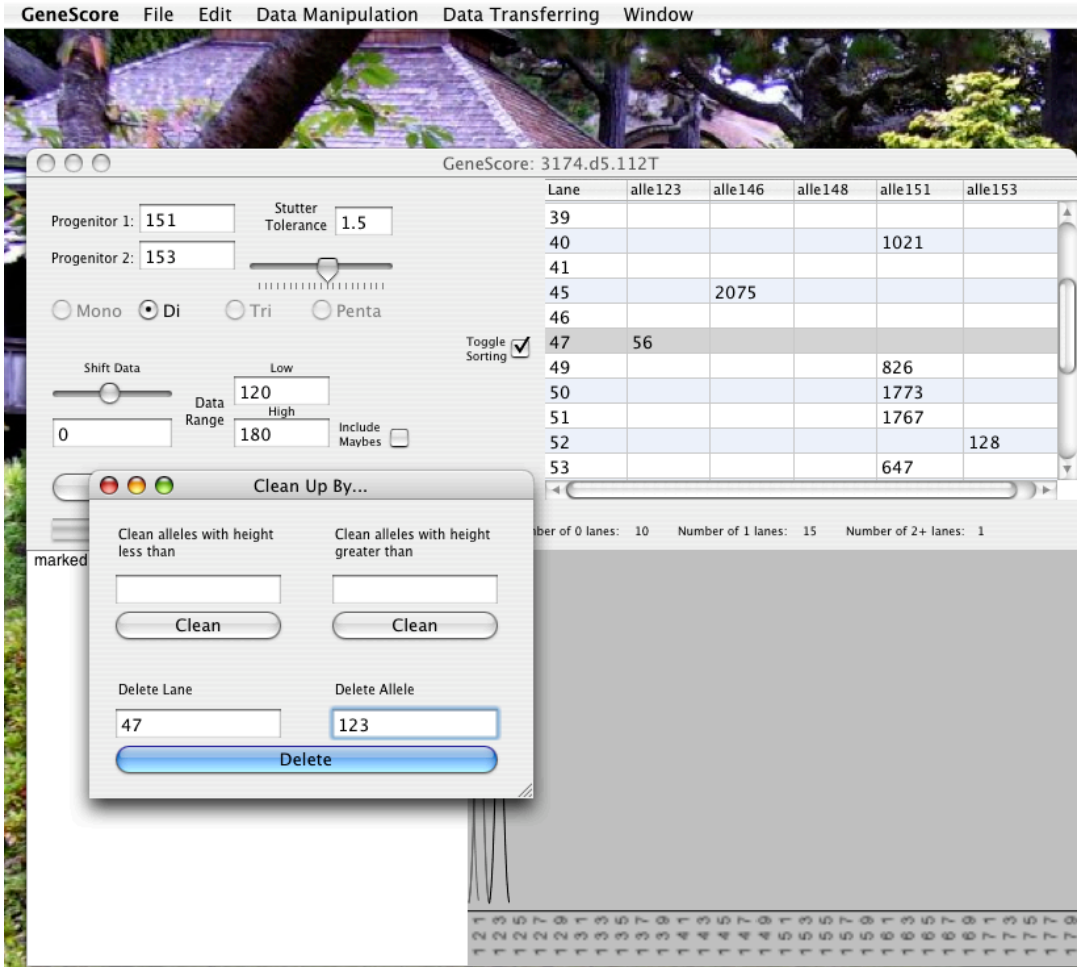


6) Fixing up.

If you see something suspicious like that allele 123, you can click on Toggle Sorting to change the table to By Lane view. Just scroll down until you find the one allele scored at 123. Clicking on the row will show that entire lane. Scored alleles show up darker, while other stuff in the lane shows up lighter. In this example we see that allele 123 in lane 47 is far too low and insignificant (as seen by its peak height of 56). We can decide to change the scoring / final results by deleting this.

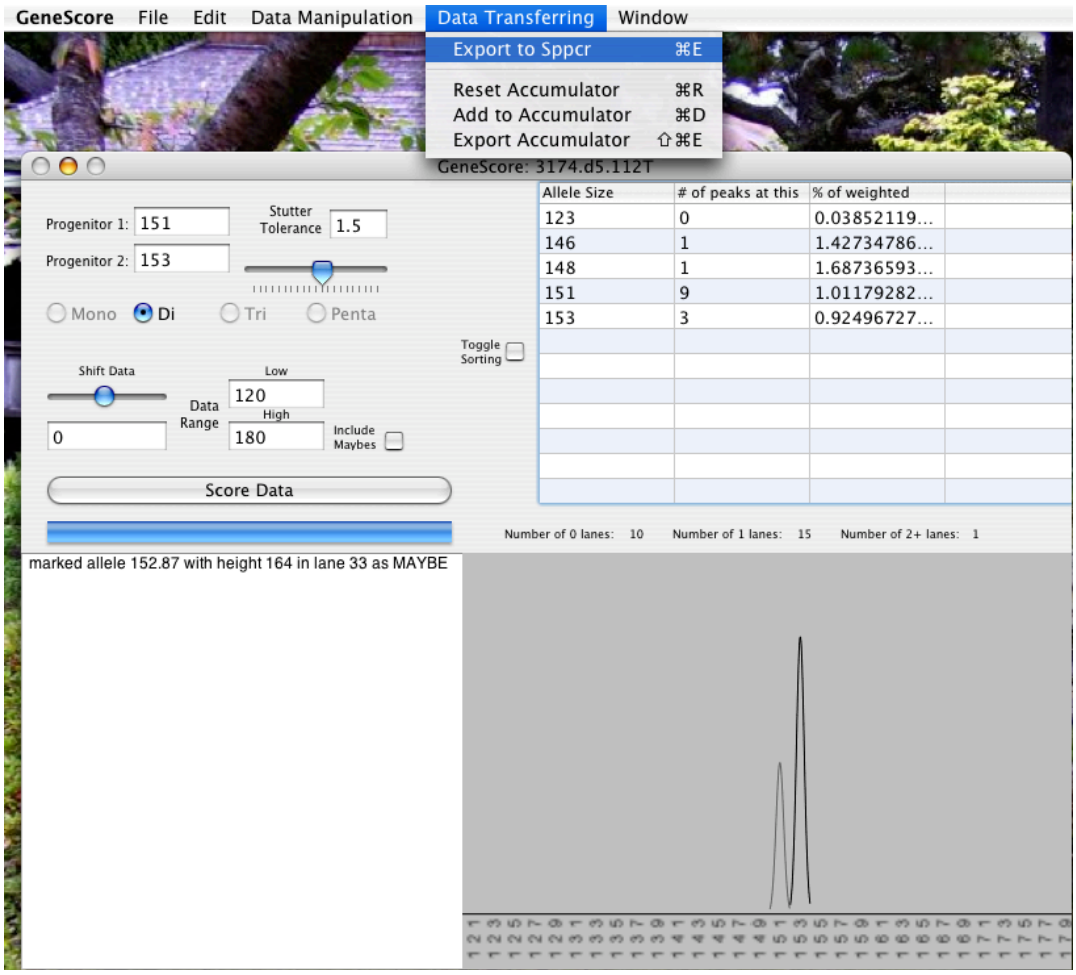


Go to the Data Manipulation menu and select Delete. A dialog box will pop up with a number of options. Since we only want to delete that one allele in that one lane, in the bottom 2 boxes, we put the lane and allele size we wish to delete, and then press the large Delete button.

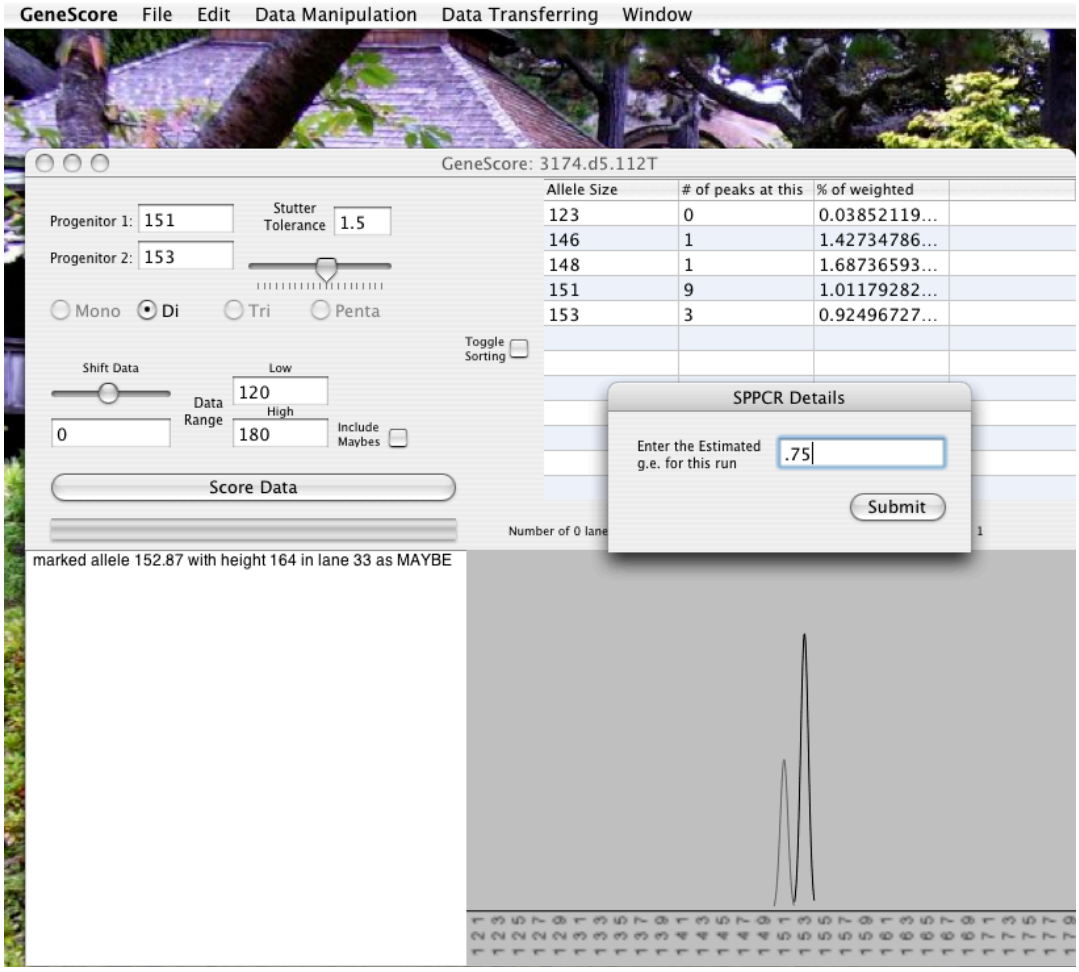


7) Exporting.

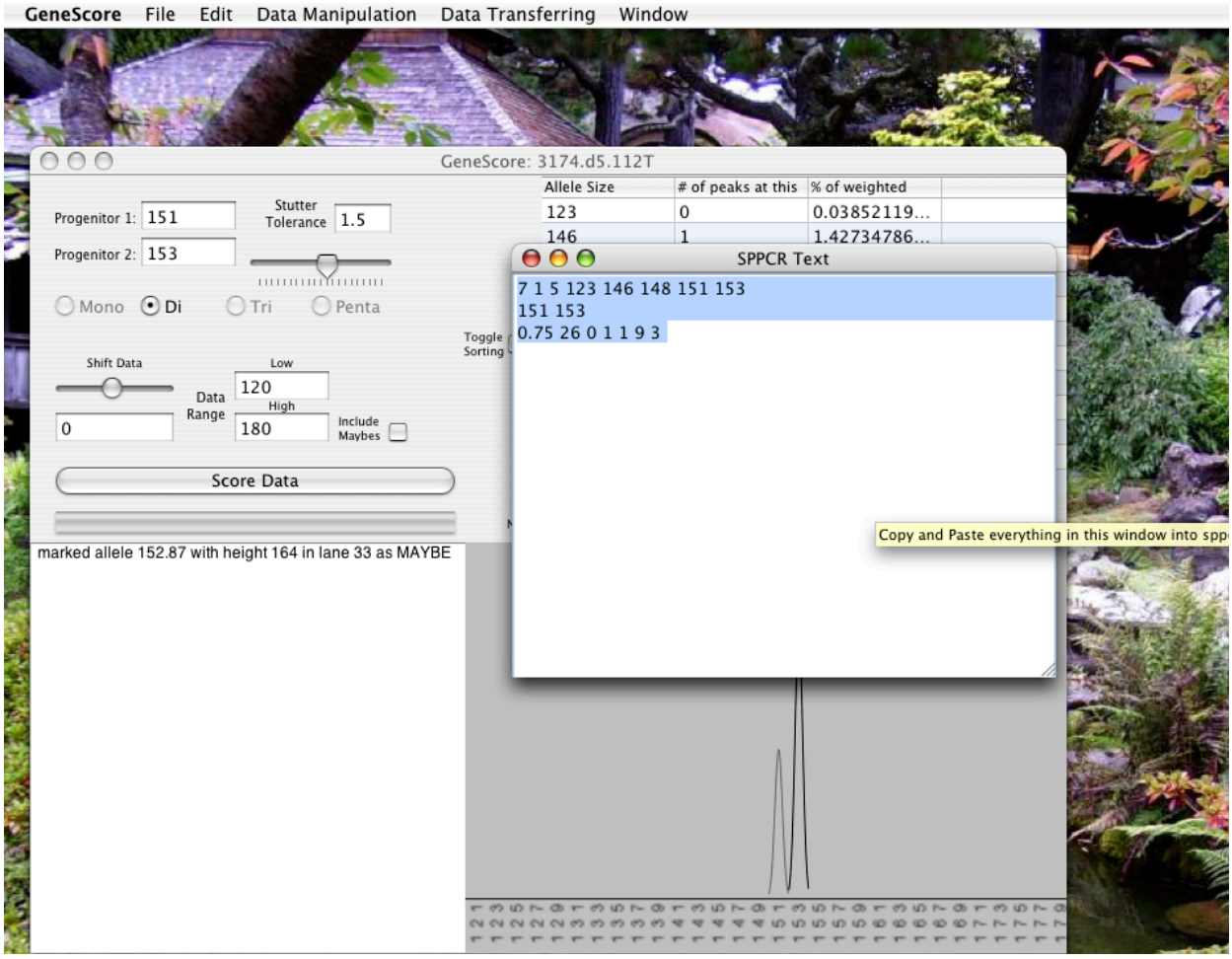
After scoring the data with Genescore, if you wish to get a statistical analysis of what it scored with Sppcr 2.0, it is necessary to export it. After scoring, open the Data Transferring menu and select Export to Sppcr.



It will then prompt you with a dialog box to enter the estimated g.e. for the run.



After clicking submit, a new window will pop up with a bunch of numbers in it. Select everything in the window and copy it. Open up Sppcr 2.0 and paste this in, and hit enter.



6) Known Problems

1) Setting the tolerance level to abnormally low levels (1 or lower) can lead to results that may not sit satisfactory.
2) Accumulator feature is assumed to be working, but testing on it has been minimal.

7) Frequently Asked Questions

How do I get rid of a scoring I believe to be false?
Go to the Data Manipulation Menu, and select "Delete...". Further instructions for the type of deletions you can make can be found in Section 4b.

Appendix A

1) Weighted Allele Equation:

$$W_{allelesize} = \frac{\sum_{n} height_{scored@size}}{\frac{n}{\sum_{all\_alleles} height^{2}}{\sum_{all\_alleles} height}}$$