

GeneMarker Tutorial

I suggest creating different projects for each separate locus (or combination of loci). This streamlines the editing process and allows to quickly reveal amplification patterns across a specific locus. I suggest the creation of 4 Folders for each Project. One for Raw Data, one for GeneMarker Projects, one for GeneMarker Panels (if you'll be using Panels for automated Allele scoring and binning) and one for Allele Report export files.

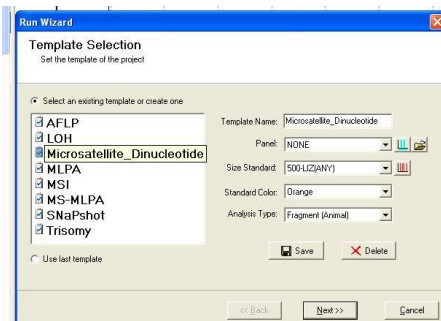
1. When GeneMarker launches click **Open Data**. (NOTE: Once a project is created and saved the **Open Project** selection is appropriate.)



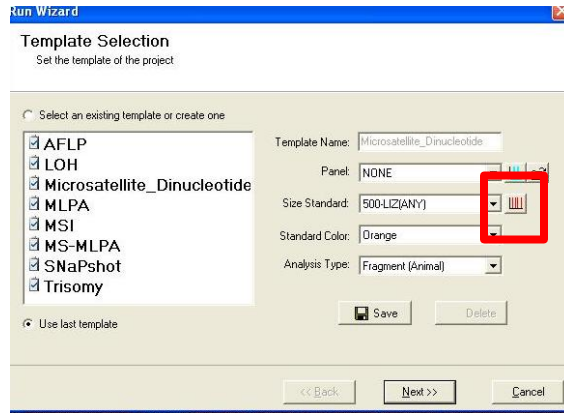
2. The Open Data Files window will appear. Here you can add one folder, multiple folders or individual files.
 - a. One folder: click **Add Folder...** navigate to folder and click **OK** in the Browse for Folder Window, Then click **OK** in the Open Data Files window.
 - b. Multiple folders: click **Add Folder...** again navigate to folder and click **OK**, repeat until all folders are displayed in the Data File List. Then click **OK** in the Open Data Files window.
 - c. Individual files: click **Add...** navigate to folder and select files (NOTE: hold control on the keyboard to select multiple files from a folder), Click **Open**. Then click **OK** in the Open Data Files window.
3. Once Files are finished loading you will be prompted to Run analysis. Click **Run**.



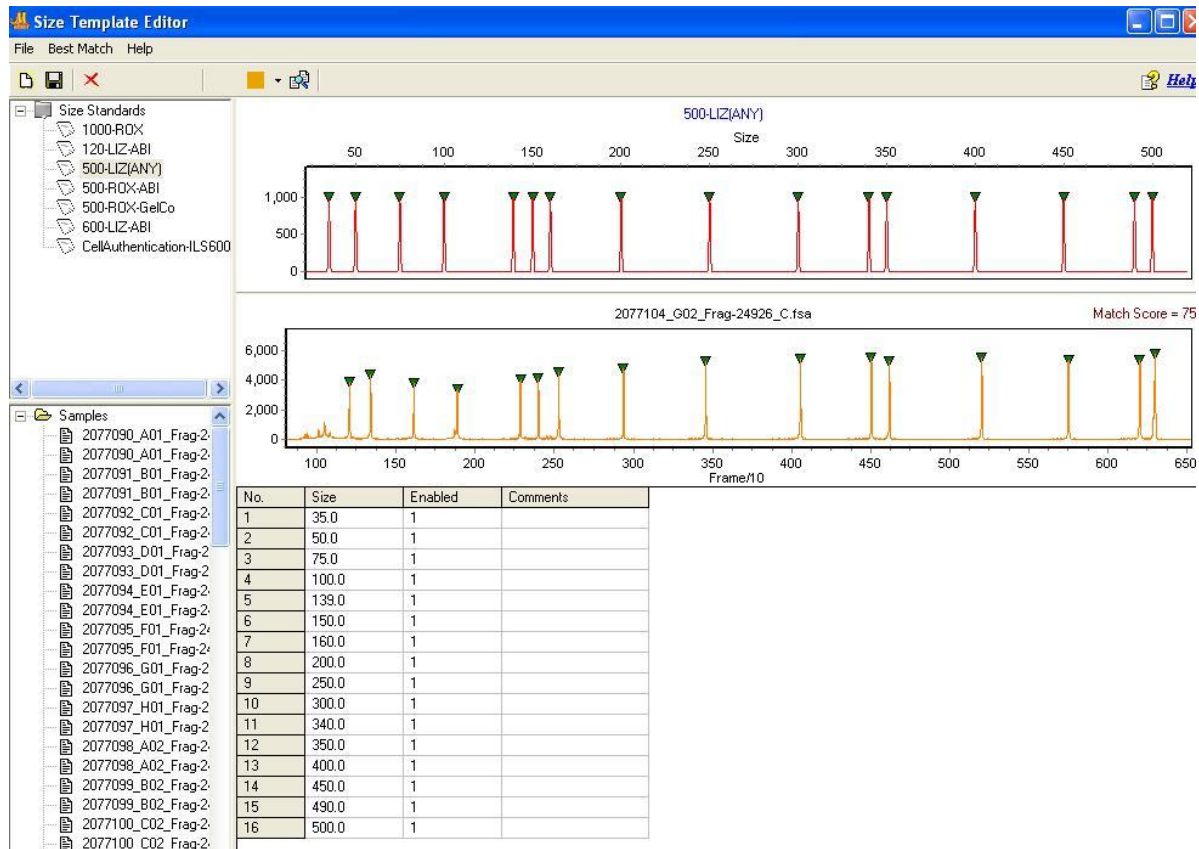
- a. Select an existing template from the list on the left and it will appear in the Template name field in the top field on the right. (NOTE: in future runs this will pre-populate automatically based on the last used template.)



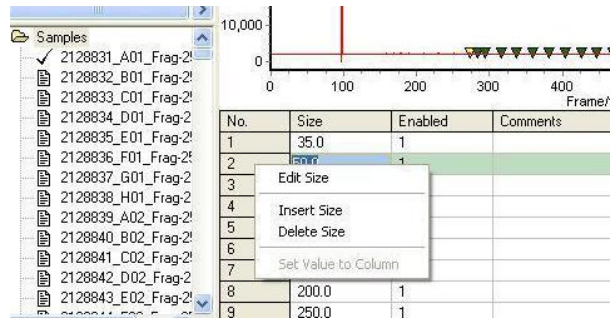
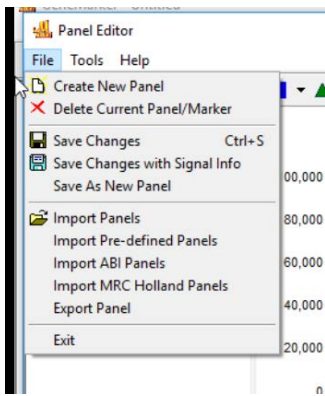
- b. Leave default for Panel: NONE. If you would like assistance with setting up panels we can schedule a half hour session to run through the exercise for some of the loci. (NOTE: you must have data files for each locus to create the panel).
- c. Size Standard. Most of our users will select GS500. (NOTE: the exception to this is for Gel Co Rox-500). In the screen shot below we have renamed the size Standard to 500-Liz(ANY). Regardless of name the most important thing is to make sure what is selected matches the correct number of fragments at the correct sizes. To do this Click the red lines icon next to the Size standard drop down.



Confirm the size standard being used for comparison is set-up properly in the software. For example Liz-500 has 16 peaks at the following sizes, once confirmed click OK to close this box.



If you need to add a size standard, for Gel Co. Rox-500 or Liz-600, click File >Create New Panel



Click on the No. column to have an edit Size Box appear to allow insertion or deletion of sizes. Once finished click OK. Gel Company-Rox-500—contains 25 single-stranded labeled fragment of: 70, 80, 90, 100, 120, 140, 160, 180, 190, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 425, 450, 475, 490, and 500.

Genescan Liz-600—contains 36 fragments that are labeled with the 5th dye LIZ®. The size standard peaks are 20, 40, 60, 80, 100, 114, 120, 140, 160, 180, 200, 214, 220, 240, 250, 260, 280, 300, 314, 320, 340, 360, 380, 400, 414, 420, 440, 460, 480, 500, 514, 520, 540, 560, 580 and 600 bps. Note: we only offer this on an as needed basis, for projects with fragments above 500 and for long term projects that require it for combining with previous data.

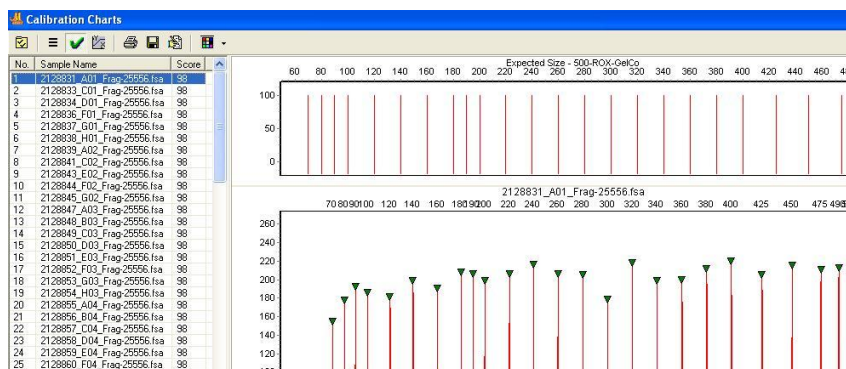
4. Click **Next**, Click **Next** (leave defaults) (NOTE: the only Data process variable I recommend changing is the Allele Call **Min Intensity Peak Detection Threshold**. Increasing it can eliminate allele calls assigned to stutter and baseline noise. If panels are created this isn't usually necessary. Read the user manual for more indepth explanations of the remaining variables). Click **OK** (leave defaults).
5. After Data processing is complete, choose save project and give your project a name. I suggest naming it by the locus/combination. Or simply click **X** to close the window if you wish to skip saving it for now.

This step is CRITICAL! Inspect the internal size standard to validate it is calibrated correctly for each ladder peak, for each and every sample. This may sound daunting, but it goes quickly and avoids scoring alleles incorrectly. (Which can lead to incorrectly adding rare alleles into a dataset.)

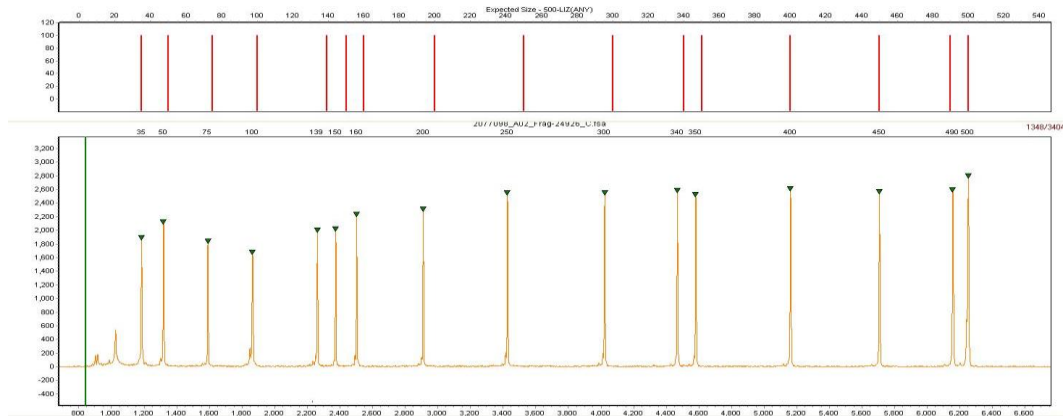
6. Click the Size calibration icon.



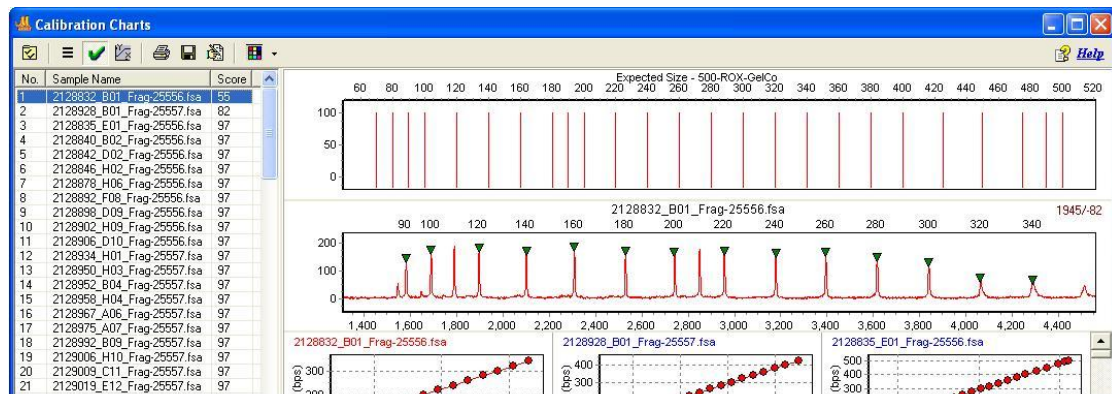
When the Calibration Charts window appears note the score column. Anything above an 85 is a good indicator that the peaks present in the sample were matched to the expected sizes of the selected size standard. Scan the list and reviewing each sample just to be sure there isn't anything wonky.



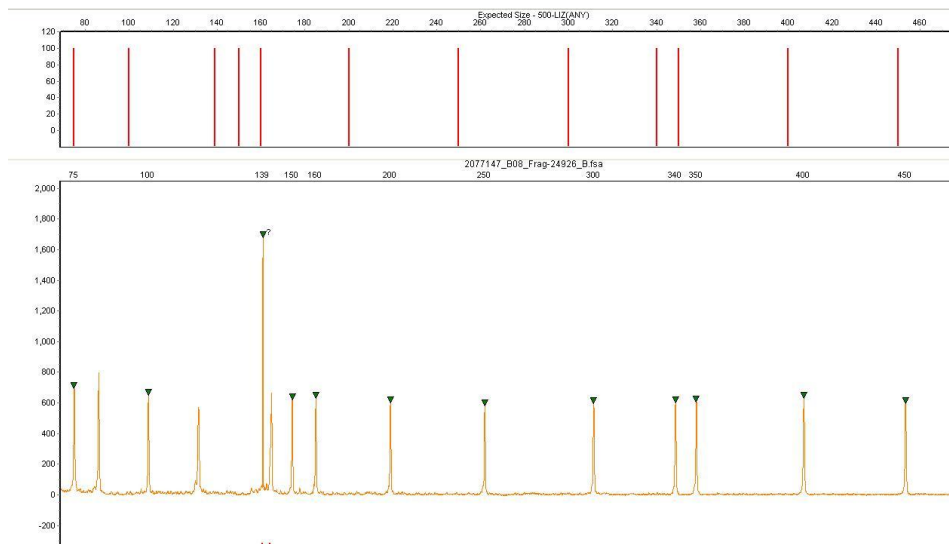
Here is an example of a perfect ladder:



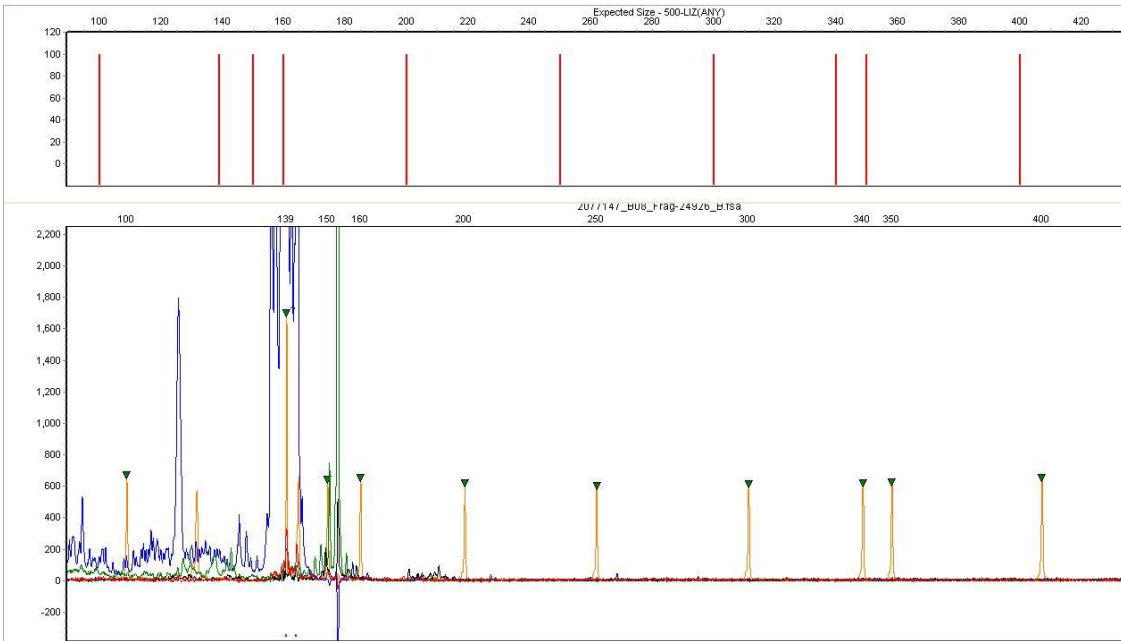
Make sure to go to the very bottom of the list, because it's sorted to show the highest values at the top. A value lower than 85 needs to be reviewed. Or click on the Score column heading and it will reverse sort lowest to highest values. In the example below, the last 8 peaks are cut off from the size standard. This is due to slow migration through the capillary. We usually re-run these samples for free if we can. However if your dye labeled peaks fall within the first correctly scored peaks it's not necessary to re-run the sample because the allele is being scored based on the closest peaks.



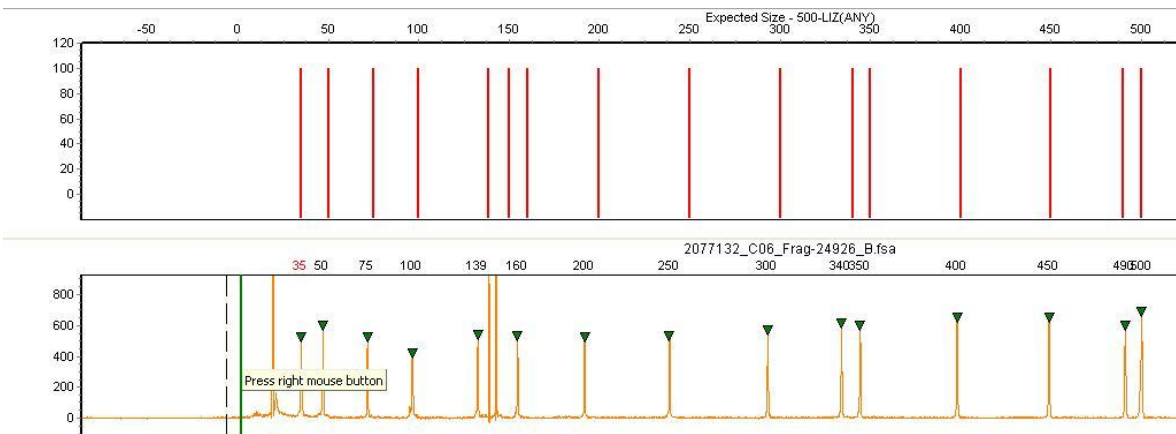
Here an example of a ladder that was scored incorrectly. Notice the 139 peak was assigned to a very tall peak? And then every peak to the left is incorrect.



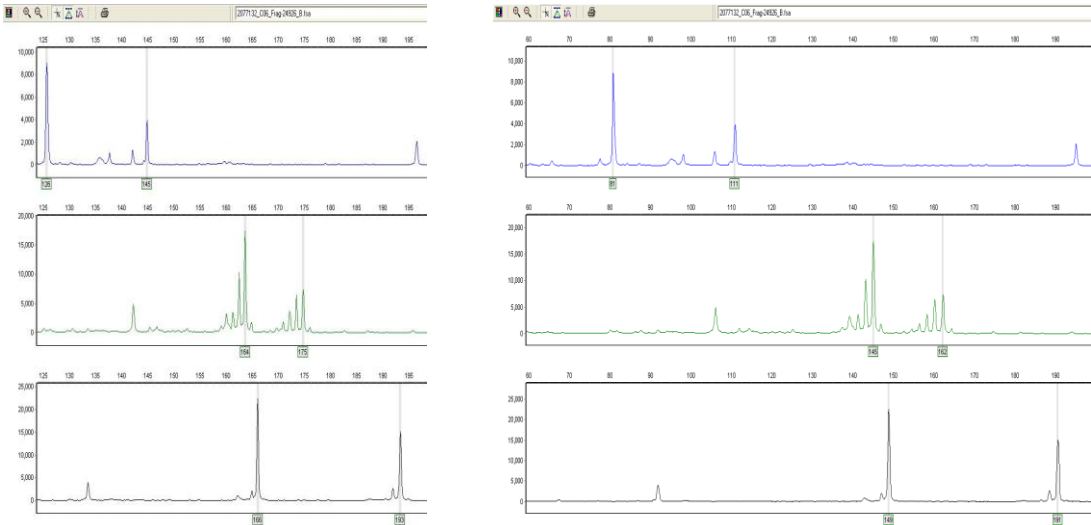
Why does this happen? Usually this is due to saturated dye labeled products drowning out the signal from the orange size standard peaks, (ROX size standard is RED). Here is the same sample with ALL dyes shown. Notice how much “blue” there is overlapping with the orange? This creates orange spectral “pull up” beneath the blue peak and the software incorrectly scores this tall orange peak as the 139 peak. This needs to be corrected manually.



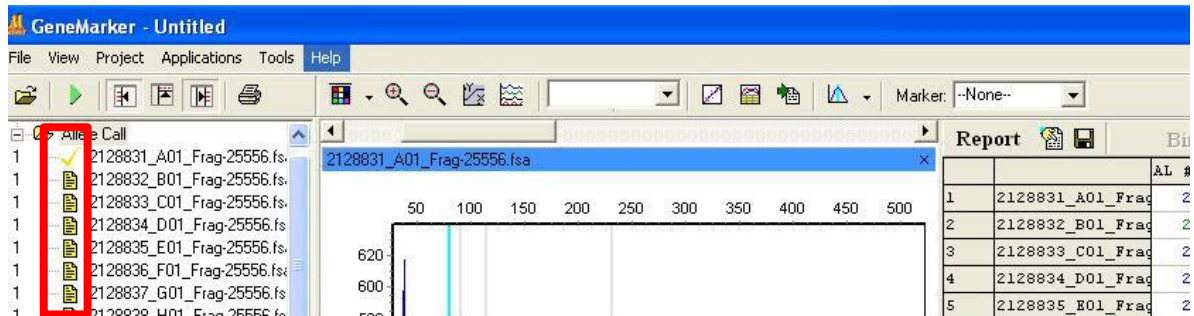
Silly computer! We humans have to fix this. I've edited the one peak at 139, but right click on the correct peak and adding the correct size and deleting the incorrect peak. The remaining peaks to the left correct themselves. GeneMarker is good like that.



Below is a sample, the left size standard peaks (not shown) were incorrectly assigned by a not so smart computer algorithm, the size standard peaks of the sample on the right have been corrected by a smart human. Notice the extreme shift in allele sizes for all the dye labeled peaks.



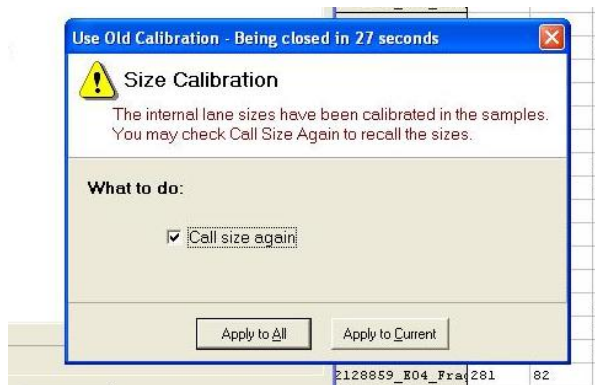
An immediate indication an incorrect Size Standard was used to calibrate data will be when the icon to the left of the sample in the left navigation window is yellow. (NOTE: if peaks can not be properly scored for a sample and the icon stays yellow even after editing then the sample data should be ignored!)



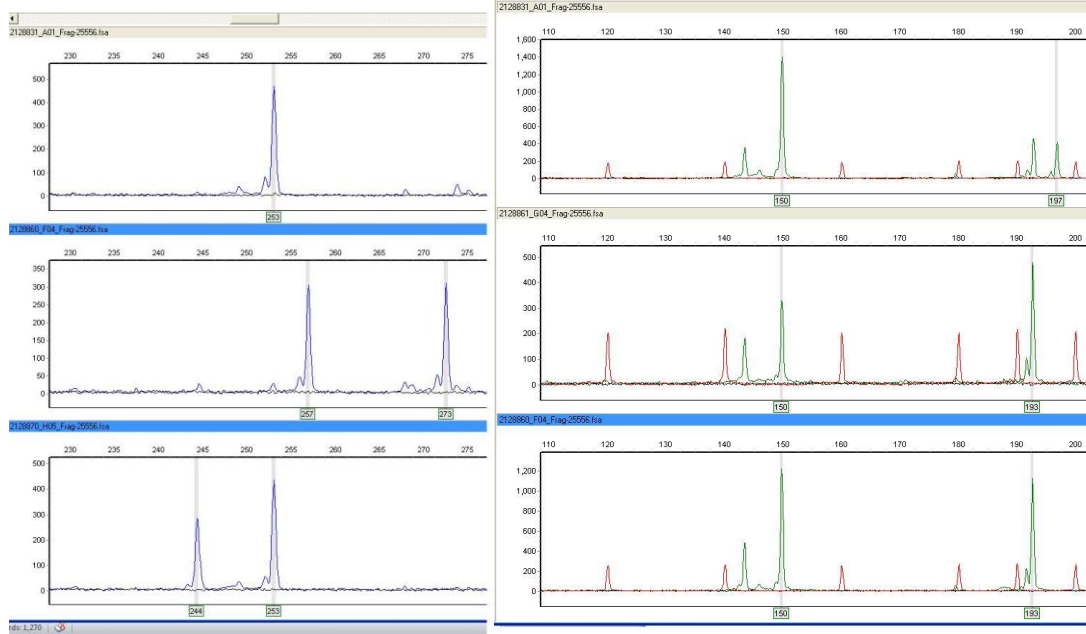
If this happens simply click on the green arrow in the top menu bar to re-run the analysis.




Select the correct size standard from the drop down list, Click **Next**, click **Next**, Click **OK**. Then **Check** Call size again and Click **Apply to All**.



7. Scoring alleles. At times this can be somewhat subjective. It's a topic all it's own. Developing the skill takes considerable time and effort. Here are some beautiful alleles. The likes of which are the result of well designed primers amplifying a perfect repeat microsatellite, highly optimized PCR reactions and thermal cycler conditions, and considerable effort to determine proper concentrations to run through the capillary. Let's end on a happy note! If this is what you're seeing, the sun is probably shining out your window and you should buy a lottery ticket!



8. To Export allele calls refer to the GeneMarker user manual for specific details, or play around with the settings, found by clicking on the icon on the right most report screen.



The screenshot shows the GeneMarker software interface. On the left is a small electropherogram with peaks labeled at 140, 164, 174, 199, 206, and 2261. On the right is a 'Report' table with columns for Bin, AL #1, AL #2, and AL. The table contains data for various samples, with row 240 highlighted. Below the table is an 'Allele Report Settings' dialog box with the following options:

- Report Style:
 - Allele List
 - Marker Table (Fragment)
 - Bin Table (AFLP/MLPA)
 - Peak Table
 - Allele Count
- Options:
 - Extend Diploid Homozygous
 - Show Allele Name
 - Show Size (0.1bps)
 - Show Height
 - Show Area
 - Show Score
- Other settings:
 - Sample Name: File Name
 - Orientation: Horizontal
 - Show " " when no allele call
 - Show Only Uncertain Alleles
 - Show Rejected Low Score Alleles
 - Hide Extra Sample Names

If you have any further questions and would like to review your data together we can schedule a half hour session. I have reviewed hundreds of thousands of files and I'm happy to offer my input.