

Analyzing T-RFLP Files Using Peak Scanner

Background:

Once upon a time, and a very good time it was, there was a simple way to analyze T-RFLP data. First one used GeneScan to generate the peak files, then one followed with a couple of wonderful scripts written by Dr. Zaid Abdo at the University of Idaho which prepared the data for downstream clustering analysis. However, as time went on, GeneScan got older and ABI ceased to support it. This would still have been fine had ABI also made sure that chromatogram files generated on their newer machines could not be read by GeneScan. Let me emphasize that last point again since I've wasted an obscene amount of time trying to search for the backward compatibility that should be there: Chromatogram files generated on any ABI sequencer that is less than 4 or 5 years cannot be read by GeneScan. It is worth noting; however, that there is a conversion program that was made by ABI, but they will not give it out. Of course even if you had the conversion program GeneScan does not play nicely with XP. Fortunately, ABI has just released a free program called Peak Scanner which fills in many of the gaps left by not supporting GeneScan.

What you will need:

First you will need to go the ABI website and download Peak Scanner (type in Peak Scanner into the search if needed). Peak Scanner is a Windows only program (known to work under XP SP2, I don't know about Vista). If you are using an Intel Mac Peak Scanner runs beautifully in Parallels.

Peak Scanner Setup:

Once you have downloaded Peak Scanner you have got to set it up to be able to analyze the T-RFLP data.

- 1) Start Peak Scanner and you will be presented with the interface pictured in Figure 1.

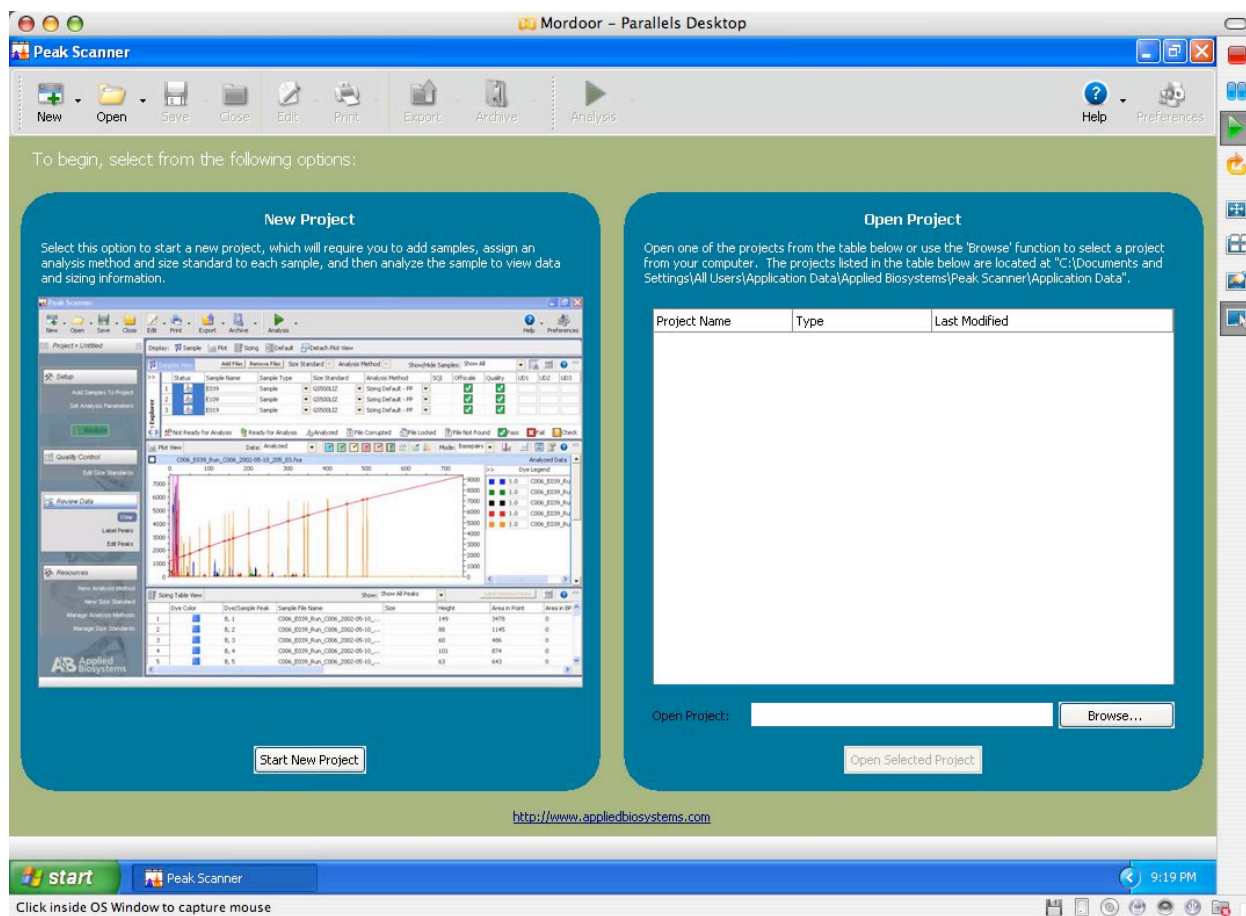


Figure 1

2) Next click on the New button in the upper left hand corner and select New Size Standard. You should see a window pop up as depicted in Figure 2. Name the size standard and set the fluorophor color. Then enter in the sizes (for ROX-1000 it should be 50, 75, 100, 125, 150, 200, 250, 300, 350, 400, 450, 475, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000). Click on Save or Save As.

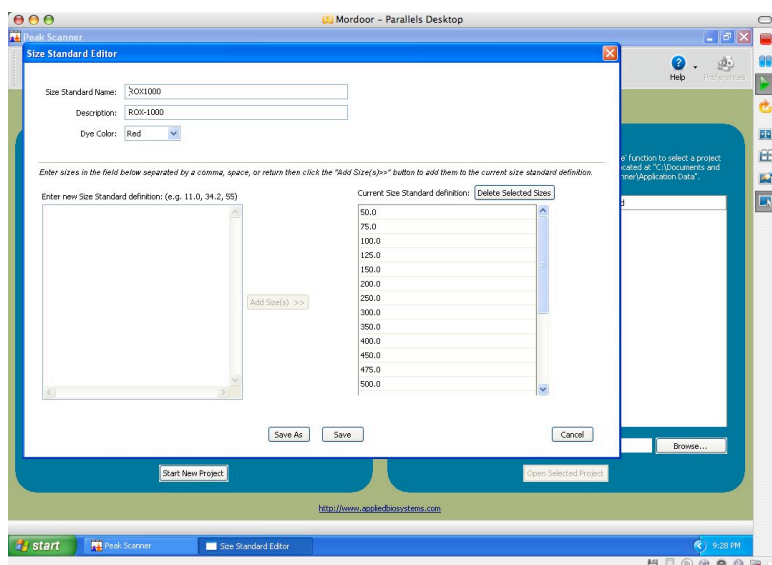


Figure 2

3) Select New again only this time select New Analysis Method. You should see a window open up like in Figure 3. Name the method (T-RFLP Analysis or something to that effect). Next click on Peak

Smoothing, Range, Peak Characteristics, Size Calling Method, Baseline-ing, and Minimum Peak Heights and fill in the values as shown in Figures 4-9 respectively.

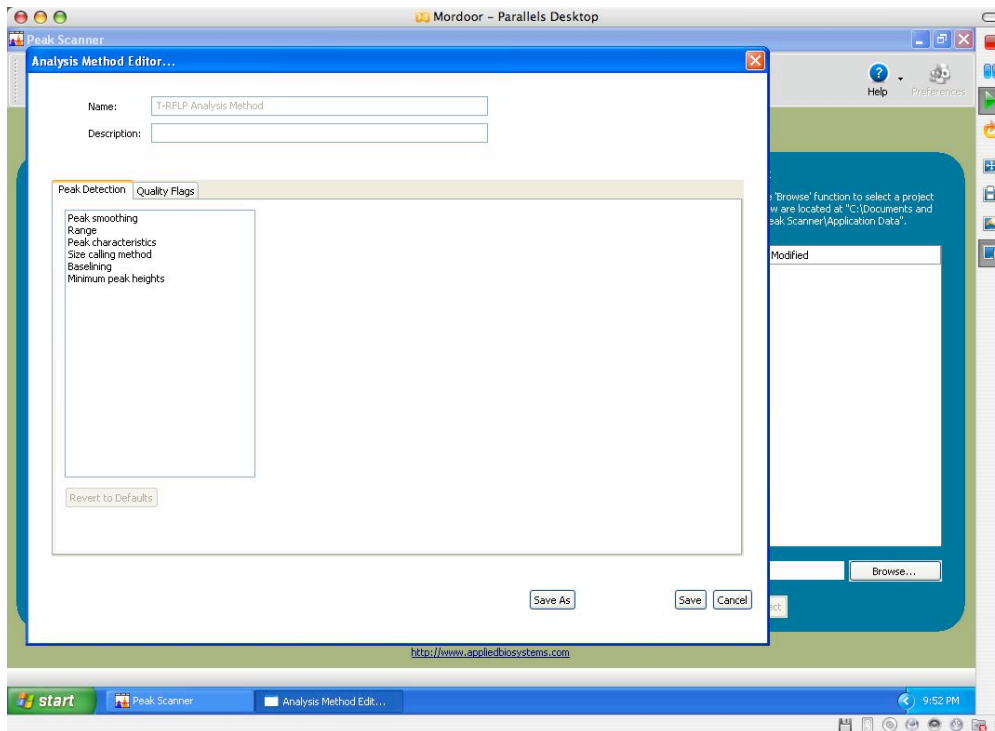


Figure 3

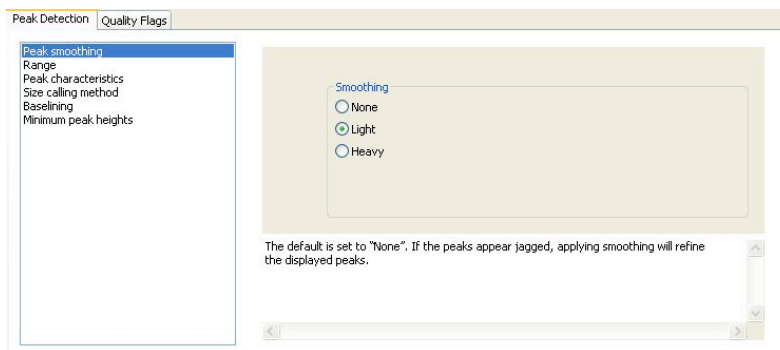


Figure 4

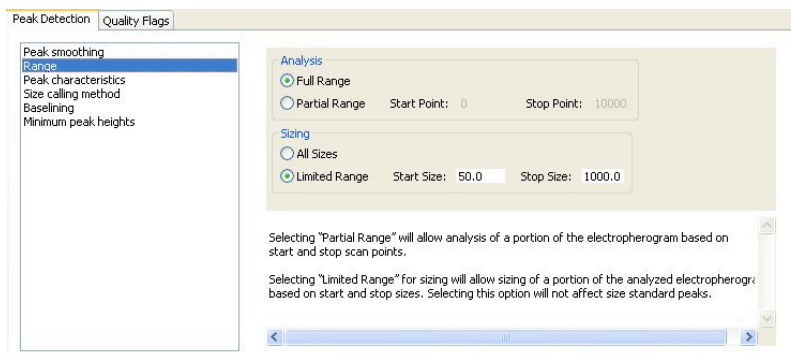


Figure 5

Peak Detection
Quality Flags

Peak smoothing
Range
Peak characteristics
Size calling method
Baselining
Minimum peak heights

Name	Blue	Green	Yellow	Red	Orange
1 Minimum Peak Half Width	1	1	1	1	1
2 Polynomial Degree	3	3	3	3	3
3 Peak Window Size	15	15	15	15	15
4 Peak Slope Start	0.0	0.0	0.0	0.0	0.0
5 Peak Slope Stop	0.0	0.0	0.0	0.0	0.0

These settings determine peak characteristics.

Minimum Peak Half Width: any peak narrower than this setting (in scans) will not be called as a peak.

Polynomial: changing this number will increase/decrease sensitivity of peak morphology, respectively.

Peak Window Size: changing this number will increase/decrease sensitivity of peak detection,

Figure 6

Peak Detection
Quality Flags

Peak smoothing
Range
Peak characteristics
Size calling method
Baselining
Minimum peak heights

☐ Global Southern
☐ Least Square
☒ Local Southern
☐ Cubic Spline

Select a size calling method. Local Southern is the recommended method.

Figure 7

Peak Detection
Quality Flags

Peak smoothing
Range
Peak characteristics
Size calling method
Baselining
Minimum peak heights

Baseline Window Size: 51

Baseline Window Size determines the size of the window to be used for baselining. Increasing the window size will decrease the baselining effect.

Figure 8

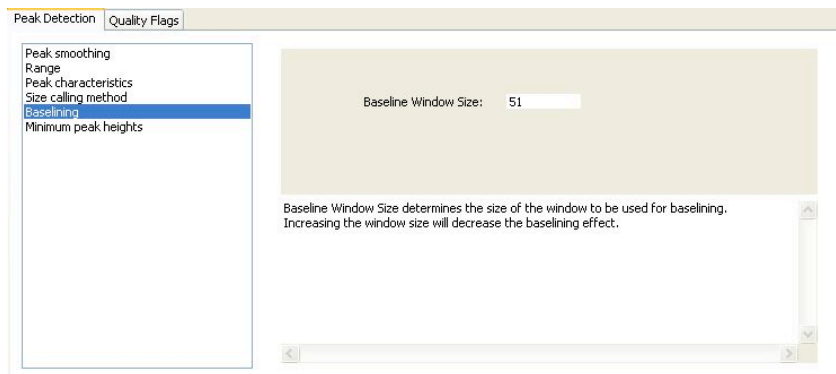


Figure 9

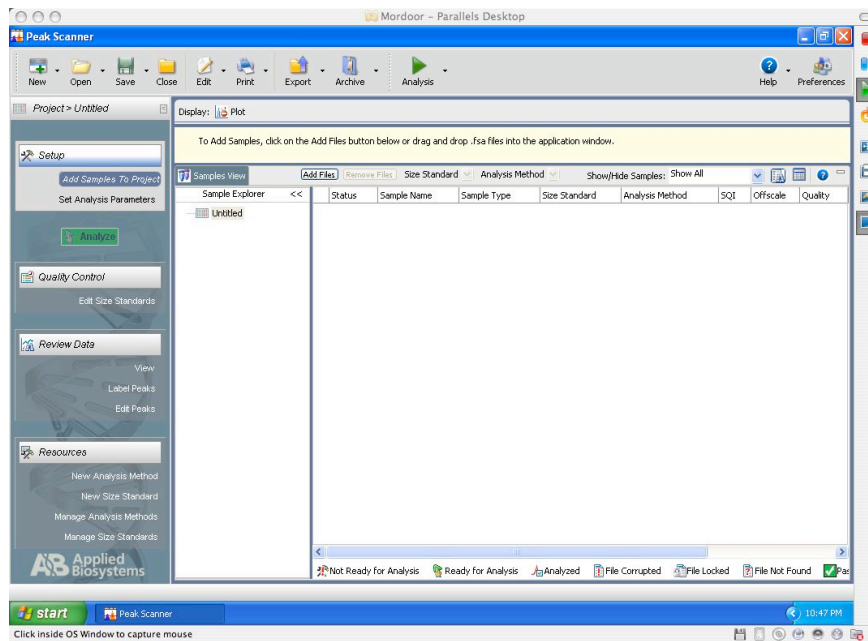
Next click on the Quality Flags tab and select Sizing Quality and fill in the values as shown in Figure 10.



Figure 10

And Save the method. Now you are ready to analyze some data.

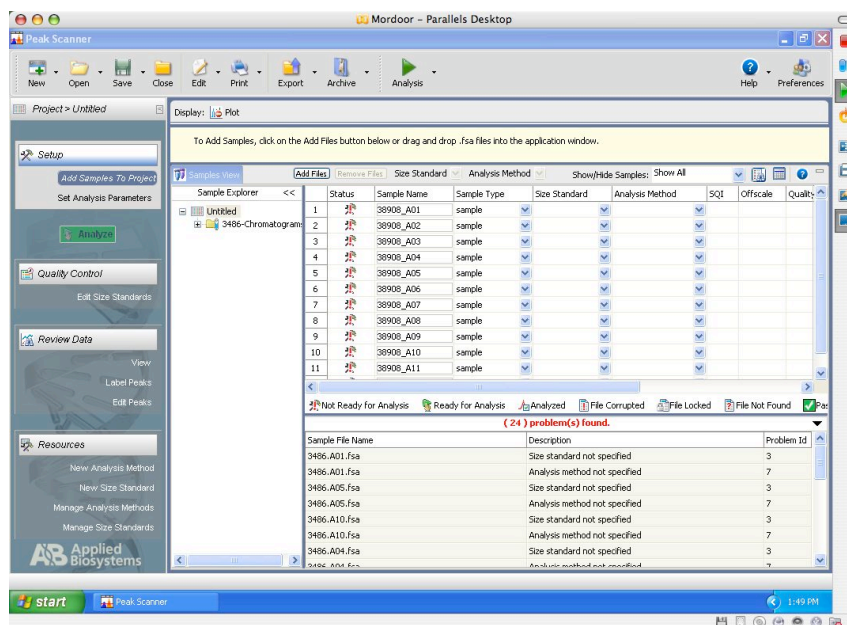
4) Click on the Start New Project button and you will be presented with a Project Window as in Figure 11.



Click on the Add Files button and navigate to where the chromatogram files are located. If you are using Parallels and the files were downloaded in the Mac environment don't worry, just drag the folder that contains the files to the Windows desktop and the folder will be copied over.

Figure 11

Now select the Chromatogram files you want to analyze, click on the Add Selected Files button then click on OK. The project window should now look something like what is



shown in Figure 12. Now don't worry about the error messages they will be fixed shortly.

Go to the Size Standard column of the first row and click on the tab. Select the Rox-1000 that you created earlier (or other desired size standard). Now click on the Size Standard column so that the whole column is selected and type Ctrl-d.

Figure 12

This should fill down the column with the same selection. If it doesn't work deselect the column, reselect it and try again. This happens sometimes and is part of this being a free (and a bit buggy) program. Next click on the tab in the first row of the Analysis Method column and select the T-RFLP method that you created. Select the whole column and fill down as before. Now everything should be set and no more errors should be reported. Finally, click on the Analyze Files.

5) Getting your peak files. You should see a window that looks something like what is shown in Figure 13.

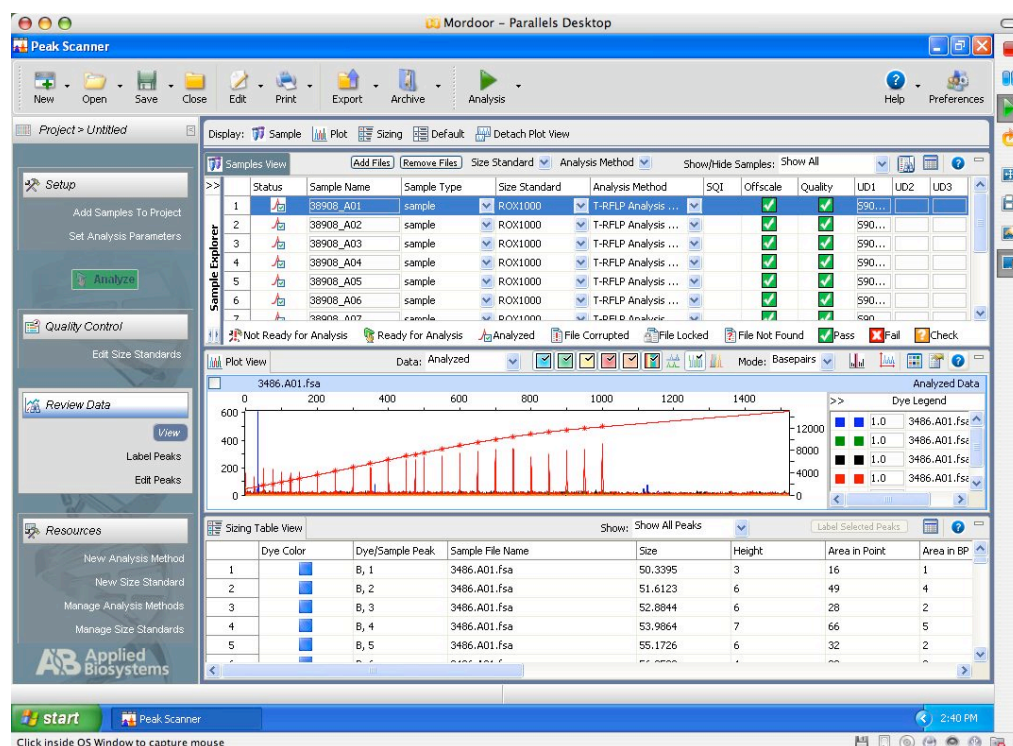


Figure 13

From here on in the order in which you do things can make a huge difference. Also which subwindow you have selected when you go to do a function makes a big difference so be aware of where you are clicking. So in order do the follow:

1) in the graph subwindow click on the rainbow square so that all of the colors are deselected. Next click on the blue square so that the blue channel is selected (Figure 14)

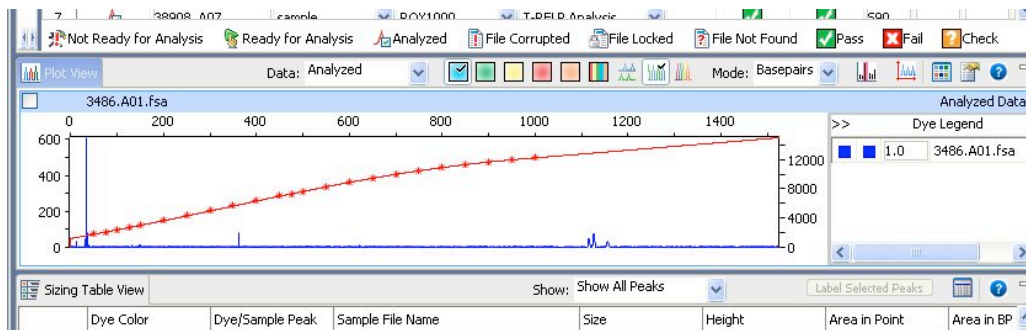


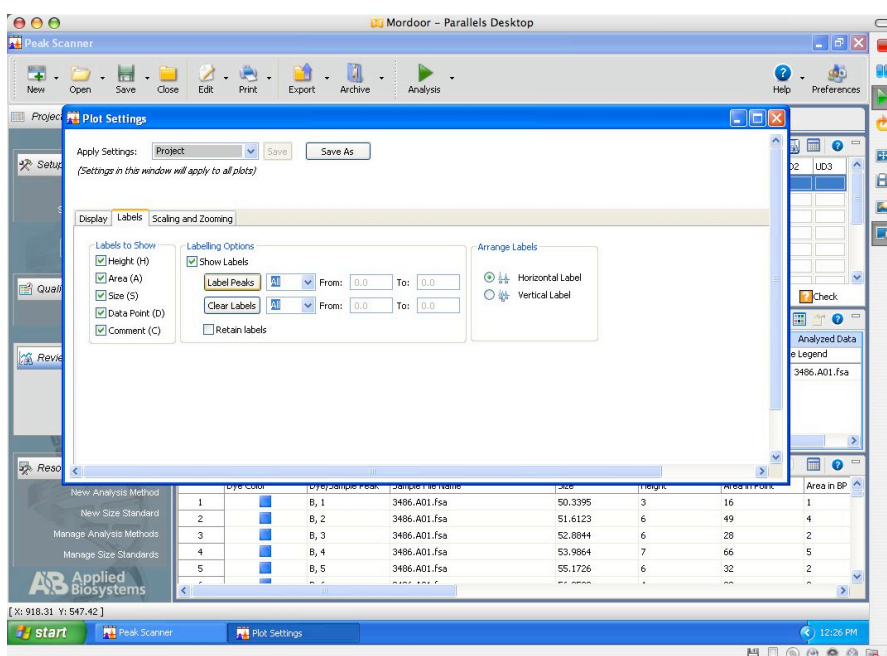
Figure 14

2) Select all of the files from the file list subwindow (Ctrl-a)

3) From the graph subwindow click on the Edit Plot Settings button




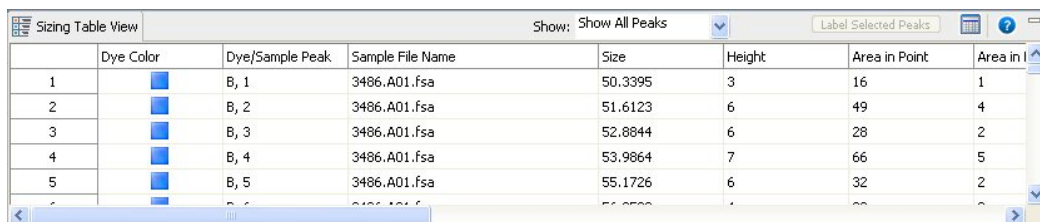
This will bring up a new window although for some reason it almost always brings it up behind the main window. So look to the task bar and you should see that one of them is titled Plot Settings, click on that to bring the settings window to the front. In the settings window click on the Labels Tab and then click on the Label Peaks button as shown in Figure 15.



Now you can close the settings window.

Figure 15

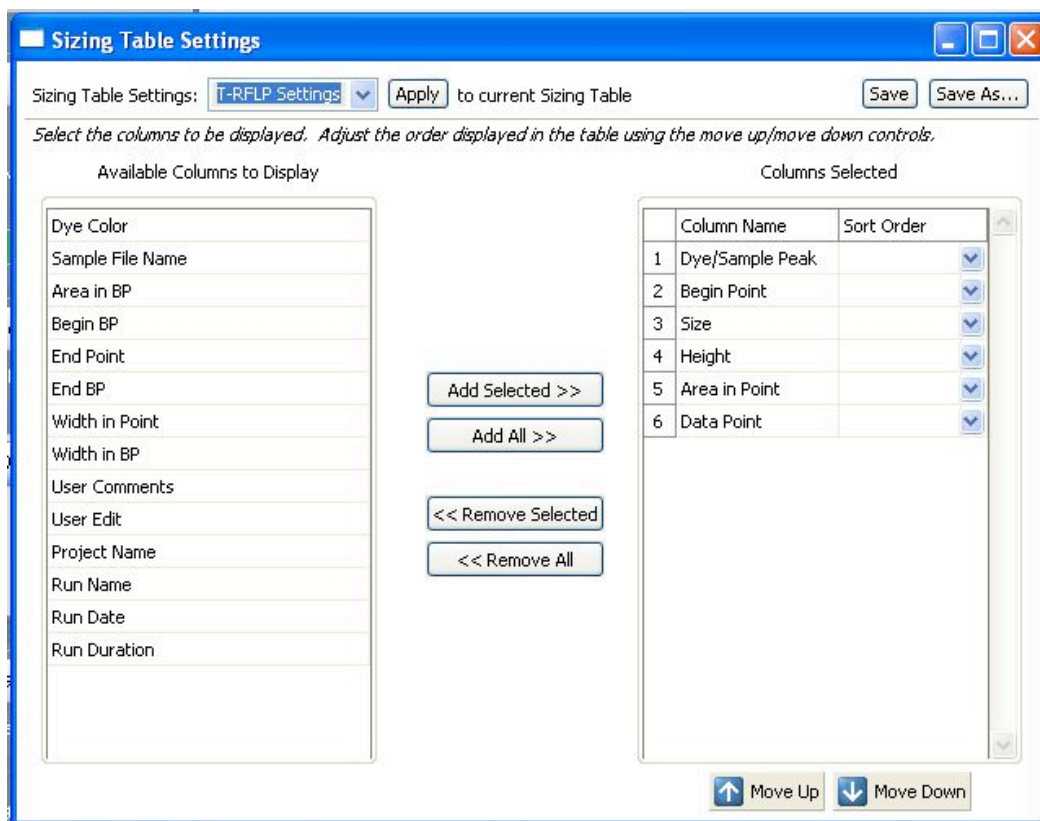
4) Next click on the Edit Table Settings button  in the top left of the bottom subwindow as shown in Figure 16.



	Dye Color	Dye/Sample Peak	Sample File Name	Size	Height	Area in Point	Area in I
1		B, 1	3486.A01.fsa	50.3395	3	16	1
2		B, 2	3486.A01.fsa	51.6123	6	49	4
3		B, 3	3486.A01.fsa	52.8844	6	28	2
4		B, 4	3486.A01.fsa	53.9864	7	66	5
5		B, 5	3486.A01.fsa	55.1726	6	32	2

Figure 16

This will bring up the Table Settings window. Make the Columns Selected on the right hand side of the window look like Figure 17. Click on Apply to apply this formatting. Before closing the window click on Save As to save this formatting. In the future you will just be able to select this method from the drop down menu and then hit Apply to format the table. Keep in mind that in order for downstream data analysis to work it is vitally important that this formatting be exactly as you see it.



Sizing Table Settings: **T-RFLP Settings** Apply to current Sizing Table Save Save As...

Select the columns to be displayed. Adjust the order displayed in the table using the move up/move down controls.

Available Columns to Display

- Dye Color
- Sample File Name
- Area in BP
- Begin BP
- End Point
- End BP
- Width in Point
- Width in BP
- User Comments
- User Edit
- Project Name
- Run Name
- Run Date
- Run Duration

Add Selected >>

Add All >>

<< Remove Selected

<< Remove All

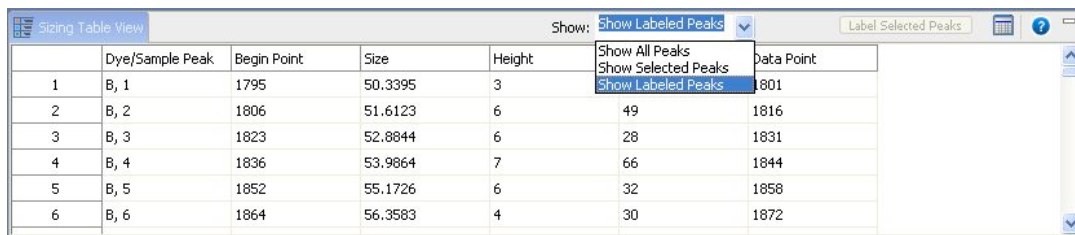
Columns Selected

	Column Name	Sort Order
1	Dye/Sample Peak	
2	Begin Point	
3	Size	
4	Height	
5	Area in Point	
6	Data Point	

Move Up Move Down

Figure 17

6) In the bottom subwindow click on the Show drop down menu and select Show Labeled Peaks (Figure 18). This will confine the data in the tables to just the blue channel.



	Dye/Sample Peak	Begin Point	Size	Height		Data Point
1	B, 1	1795	50.3395	3		
2	B, 2	1806	51.6123	6	49	1816
3	B, 3	1823	52.8844	6	28	1831
4	B, 4	1836	53.9864	7	66	1844
5	B, 5	1852	55.1726	6	32	1858
6	B, 6	1864	56.3583	4	30	1872

Figure 18

Set up for file export is now complete.

7) Select the first file in the file list. Now click on the Sizing Table View tab of the bottom subwindow. Now click on the Export button at the top of the window and select Export. Save the files in a new folder. You should name the files sequentially and something simple like A1 or M1. Repeat this step with every file. Remember the order to this because if you forget to click on the Sizing Table View tab you will export the wrong thing. When you have exported all of the files just quit Peak Scanner. Don't bother saving the project because yet another bug of the program is that you can never get a saved project to work properly. Remember it's free though.

8) Now drag the file that has all of the Peak Files onto the Mac Desktop or wherever for downstream analysis.

Troubleshooting:

I didn't write the program, but I've used this enough to field some questions. So if you have questions about this tutorial or basic Peak Scanner problems that are not covered here, then I can be reached at jre@umich.edu.