# Analysis of diaPASEF data in Skyline

Note: This tutorial uses data from a Bruker timsTOF Pro instrument. If you prefer to use the data from a Thermo Q-Exactive Plus or Sciex TripleTOF, see their respective versions.

In this tutorial you will learn how to use Skyline to perform targeted post-acquisition analysis for peptide and inferred protein detection and quantification using a SWATH-MS dataset acquired on a QqTOF instrument with IMS separation (timsTOF Pro, Bruker), using a 24 fixed width window precursor isolation scheme and a 17-minute gradient.

The data are from samples replicating the LFQBench study where quantitative benchmarking samples were created by mixing proteomes of 3 organisms in defined ratios (figure).

Initially, you will set all the parameters in the Skyline session required to work with data-independent datasets and then you will proceed to extract the quantification information from the raw data files. You will import DDA search results to create a spectral library in order to generate peptide query parameters to analyze the DIA data.
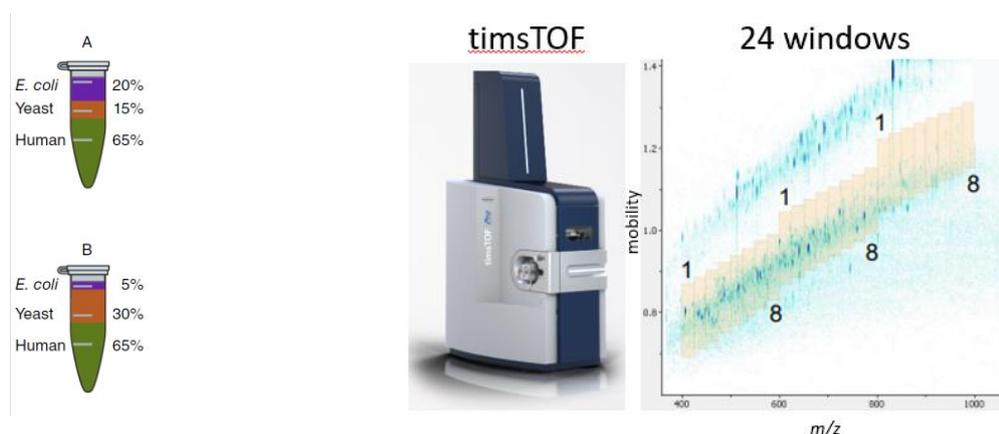
[figure adapted from Navarro, P. et al. A multicenter study benchmarks software tools for label-free proteome quantification. *Nature Biotech* 34, 1130–1136 (2016)]

Getting Started

There are two dataset options for this tutorial. The "full" dataset is ~19Gb. The "small" dataset is a 4-minute window extracted from the full dataset (~5Gb). The tutorial should run considerably faster on it, and it is also more likely to run on a laptop. The screenshots will be from the small dataset; because the results shown here are from a short 4-minute window, they look a bit sparse. The results from the full dataset are much better. To start this tutorial, download one of the following datasets:

https://skyline.ms/tutorials/DIA-PASEF-small.zip

https://skyline.ms/tutorials/DIA-PASEF-full.zip

Extract the files in it to a folder on your computer, like:

C:\Users\brendanx\Documents

This will create a new folder (corresponding to the ZIP file name) like:

C:\Users\brendanx\Documents\DIA-PASEF-small

If you have been using Skyline prior to starting this tutorial, it's a good idea to revert Skyline to its default settings. To do so:

- Start Skyline.
- On the **Start Page,** click **Blank Document** which looks like this:
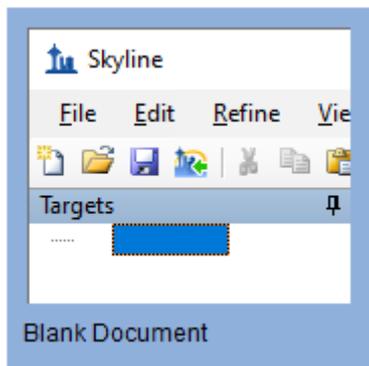
- On the **Settings** menu, click **Default.**
- Click **No** on the form asking to save the current settings.

The settings in your current instance of Skyline have now been reset to the default.

Since this tutorial covers a proteomics topic, ensure that the user interface is set to the "Proteomics interface"

- Click the user interface button in the upper right hand corner on the Skyline toolbar and select **Proteomics Interface** which looks like this:
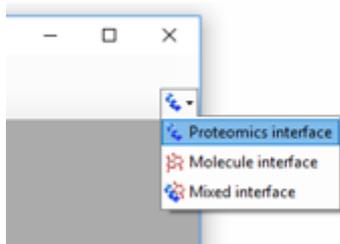
Skyline is operating in Proteomics mode which is displayed by the protein icon in the upper right-hand corner of Skyline.

For DIA analysis, you will want to force Skyline to integrate peak areas for all extracted quantitative transitions. To do that:

- On the **Settings** menu, click **Integrate all.**

Using the Import Peptide Search wizard

If you know where to look, you could perform all of the following steps through various menu options on Skyline. To simplify the normal workflow for DIA, you will instead use a "wizard" which will step you through the critical choices required for Skyline to run the analysis. You will start with building a spectral library from DDA peptide spectrum match results, followed by specifying a set of

DIA runs to extract chromatograms from, and then specifying various settings, and finally the targets themselves, which you are interested in querying.

You will build the library from the analysis of one DDA run each of the A and B samples. The DDA search has already been performed using the MSFragger search engine and post processed using PeptideProphet – see this link for more DDA search engines supported by Skyline https://skyline.ms/wiki/home/software/Skyline/page.view?name=building_spectral_libraries. You will start from the interact.pep.xml file which is the output of PeptideProphet and contains the database search results from both DDA files.

To get started do the following:

- On the **File** menu, choose **Import**, and click **Peptide Search**.

Skyline tells you that you need to save the current document so that it will know where you want to perform the analysis on your computer disk drive.

- Click the **OK** button.
- Navigate to the DIA-PASEF-small folder created earlier.
- In the **File name** field, type "DIA-PASEF-tutorial".
- Click the **Save** button.

The **Import Peptide Search** wizard should appear looking like this:

- Click the **Add Files** button.
- Navigate to the DIA-PASEF-small folder created earlier and into its DDA_search subfolder.
- Double-click on the "interact.pep.xml" file.
- In the **iRT standard peptides** dropdown list, click "Automatic".
- In the **Workflow** box, click the **DIA** option.

The **Spectral Library** page of the wizard should now look like this:

Note: 0.95 is the threshold applied to the PeptideProphet probability computed for every peptide spectrum match in the DDA database search – in this particular data set this corresponds to a PSM false discovery rate of 0.2% but this will differ among data sets so a score threshold to achieve the FDR you want to use should be entered here.

- Click the **Next** button.

Skyline should begin building the library, showing progress in a form titled **Build Peptide Search Library**. When the library build is complete, Skyline will attempt to find peptides consistently present in all runs to use as iRT landmark peptides. It will show a form asking how many of these peptides to use for adjusting RTs for other peptides:
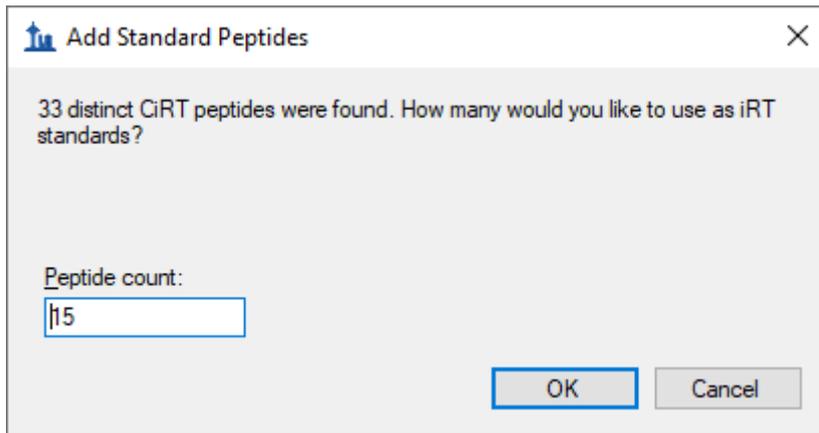
- In the **Peptide count** box, enter "15".
- Click the **OK** button.

Skyline will then present a summary form like the one below:

File	Points	Equation	R	Result
DIA-PASEF-tutorial.blib	15	iRT = 8.765 * Measured RT - 54.858	0.992	Success
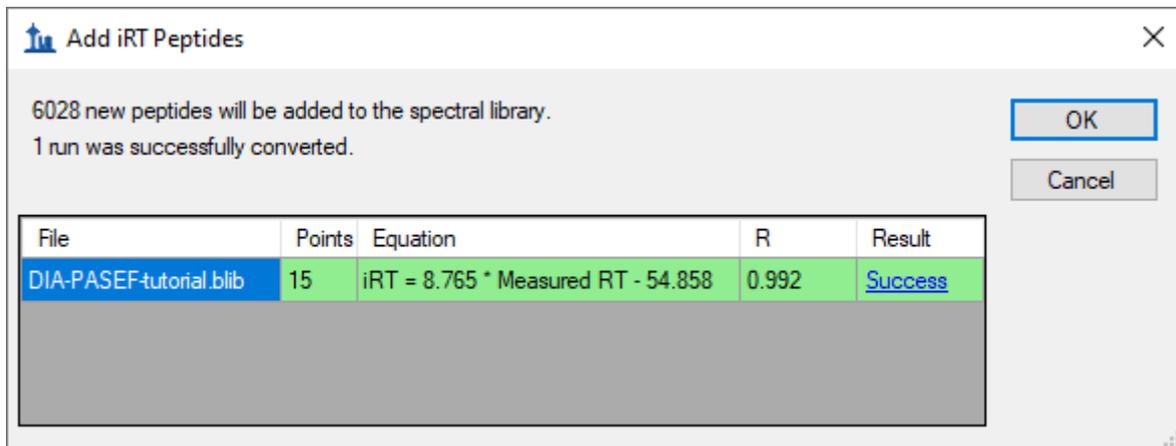
- Click the **Success** link (highlighted in blue) to see the linear regression.
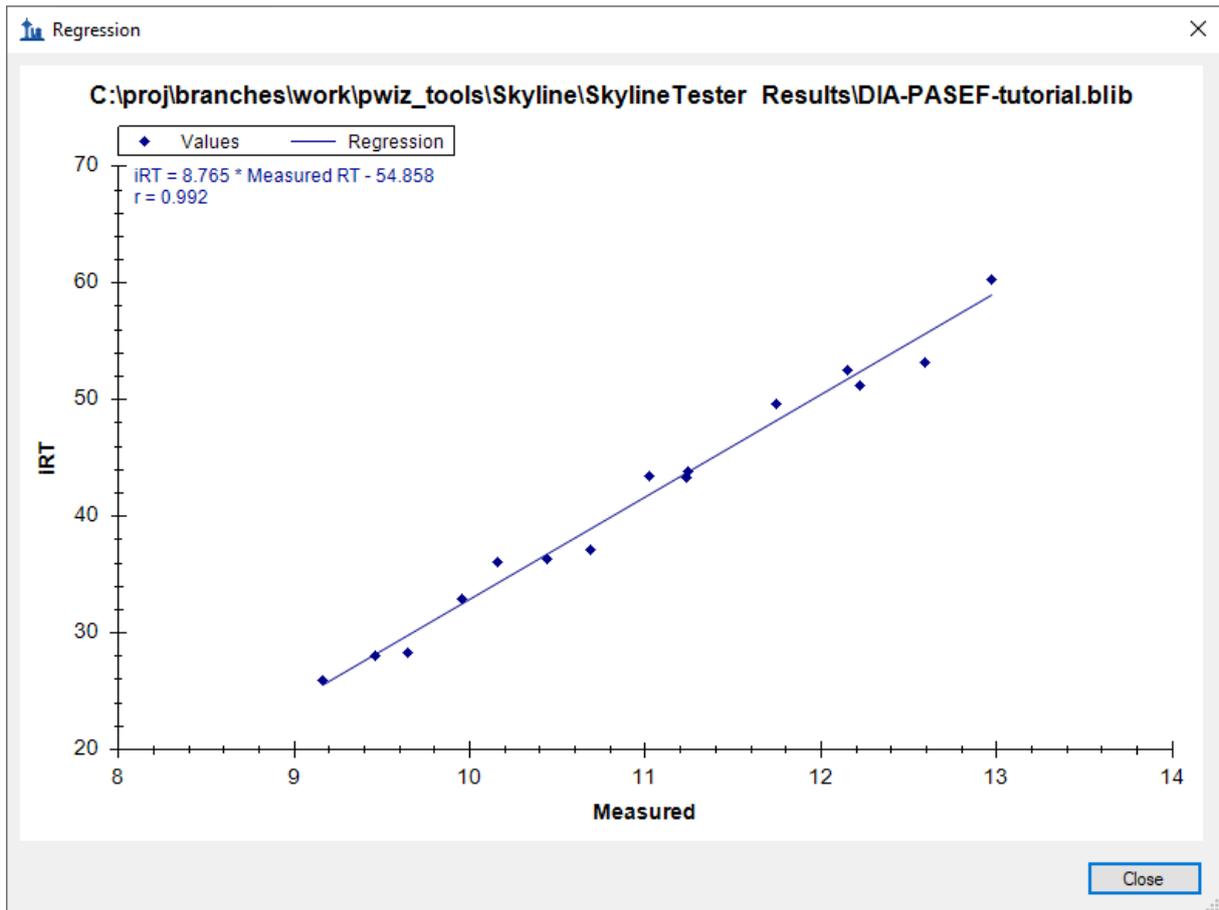
- Click the **Close** button.
- Click the **OK** button.

Skyline will ask if you want to recalibrate the iRT standard values onto the regression line calculated and shown in the graph above.

- Click the **No** button.

Skyline presents a long list of peptides with ambiguous matches. In other words, the only evidence for these peptides came from spectra which were matched to multiple peptides, making those spectra and matching peptides not ideal for subsequent targeted analysis.

- Click the **OK** button.

You are presented the **Extract Chromatograms** page which allows you to tell Skyline where to find the DIA data files it will use for chromatogram extraction, peak detection and peak area calculation.

- Click the **Browse** button.
- In the DIA-TTOF folder, double-click the DIA subfolder.
- Select all 6 files in this folder.

The **Browse for Results Files** should look like this:
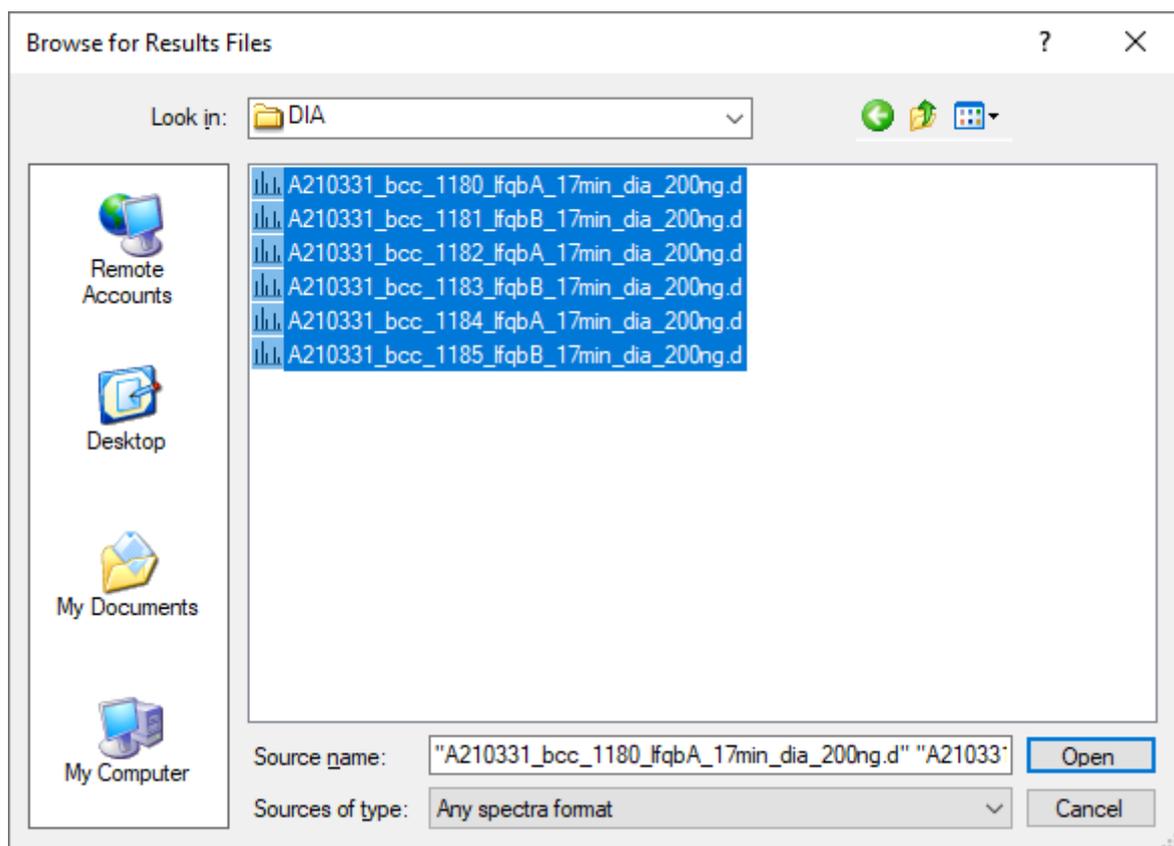
- Click the **Open** button.

The **Import Peptide Search** form should look like this:
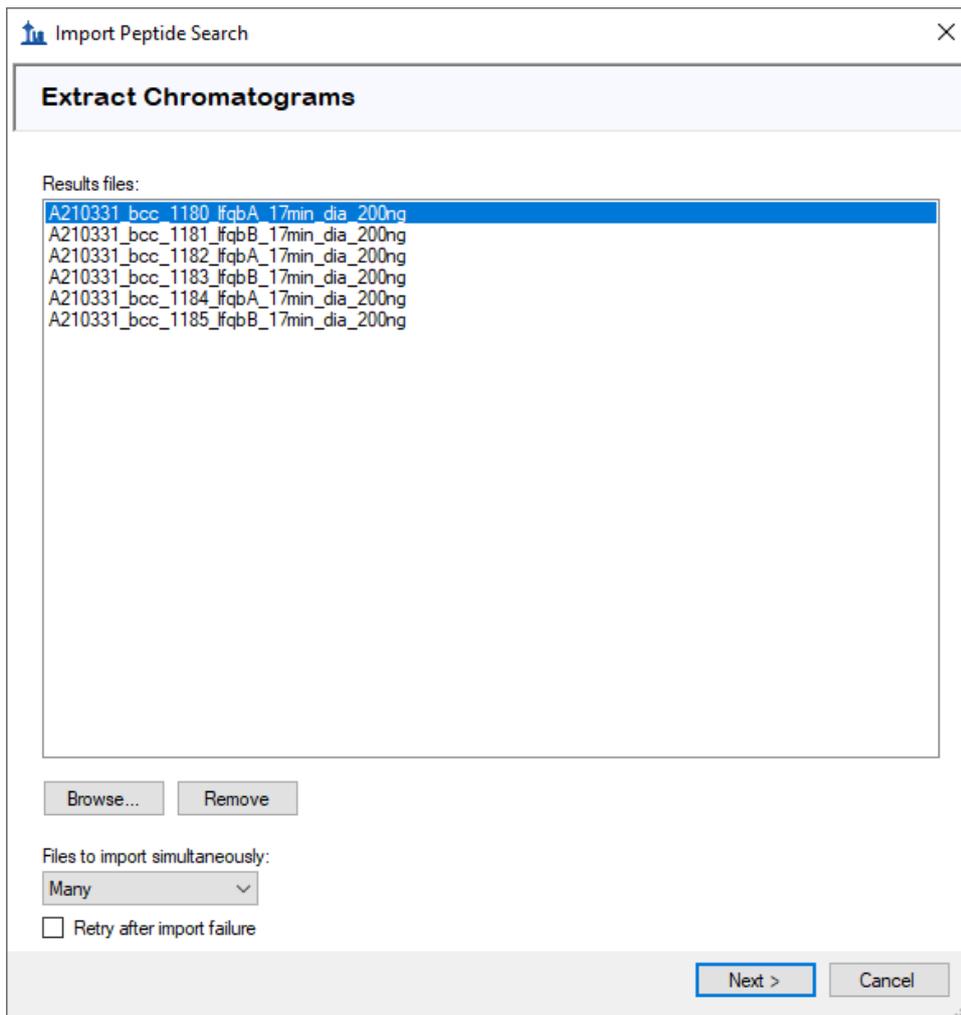
- Click the **Next** button.

Skyline presents a form asking if you want to remove the common prefix and suffix which makes the replicate names shown in the Skyline interface shorter and generally easier to work with.

- Click the **OK** button.

In the **Configure Transition Settings** page, make the following changes from the default values:

- Set **Precursor charges** to "2, 3, 4".
- Set **Ion types** to "y, b".

NOTE: You could leave the ion type "p" which stands for precursor, and Skyline would extract chromatograms covering the precursor isotope distribution from the MS1 spectra in the DIA data files. However, to simplify this tutorial and reduce processing time and output file size, the description below deals only with fragment ion chromatograms extracted from the DIA MS/MS spectra.

- Check **Use DIA precursor window for exclusion**.

The **Import Peptide Search** form should look like:

- Click the **Next** button.

Skyline presents the **Configure Full-Scan Settings** page.

Note: These extraction settings depend on the type and settings of the instrument used for data acquisition. The optimal extraction settings may be slightly different for each dataset. PASEF data is always centroided so you should select 'Centroided' and specify a '20' ppm mass accuracy for extraction.

Make the following changes to the default values:

- In both **Mass accuracy** fields, use "20" ppm.

Now you need to define a new isolation scheme according to the parameters defined on the instrument for data-independent acquisition.

Note: In this example, we used 24 fixed width (25 m/z) windows that covered the range from 400 to 1000 m/z.

- In the **Isolation scheme** dropdown list, click "<Add…>".

This brings up the **Edit Isolation Scheme** form in which you should do the following.

- In the **Name** field, enter "diaPASEF (24 fixed)".
- Click the **Prespecified isolation windows** option.
- Click the **Import** button.
- Navigate to the DIA-PASEF-small folder created earlier and into its DIA subfolder.
- Double-click the file "A210331_bcc_1180_lfqbA_17min_dia_200ng.d".

The **Edit Isolation Scheme** form should now look like this:

Skyline has worked out the **Start** and **End** *m/z* values for the DIA isolation ranges measured on the mass spectrometer.

To visualize the isolation scheme:

- Click the **Graph** button.

You will see a form like the one below:
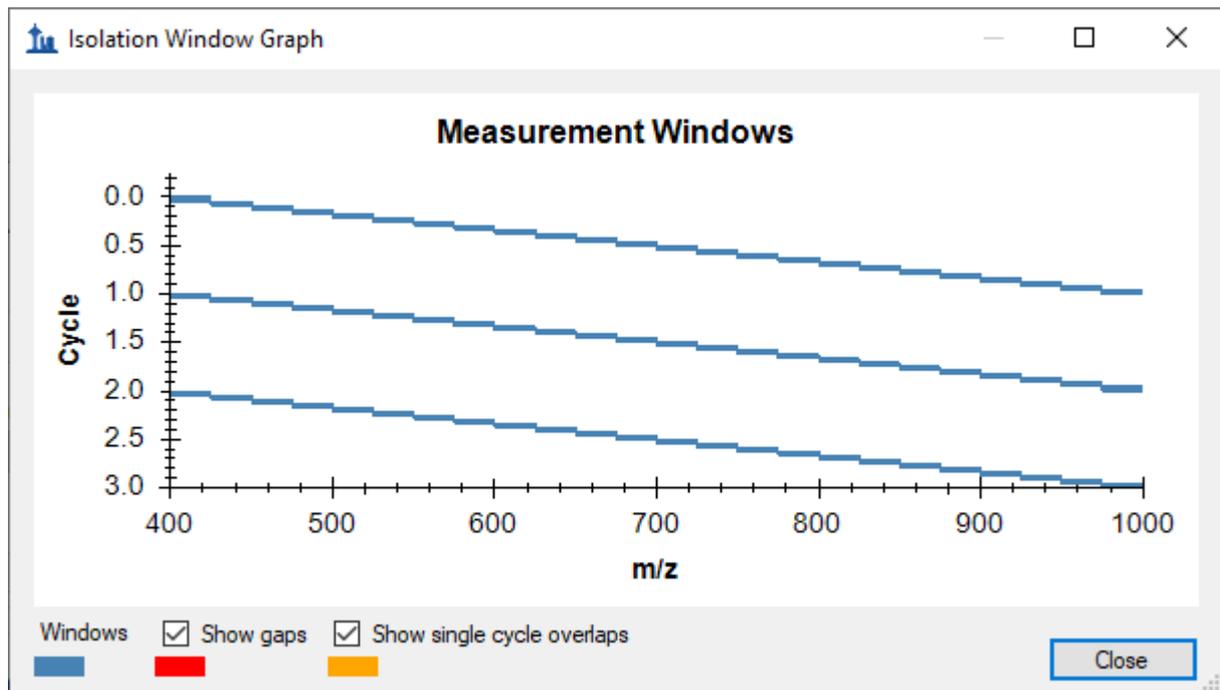
You can click-and-drag a rectangle to zoom in or use the mouse scroll-wheel. There is no red for **Show gaps** or yellow for **Show single cycle overlaps**, which are both errors in DIA isolation scheme design.

- Click the **Close** button.
- Click the **OK** button in the **Edit Isolation Scheme** form.

The **Import Peptide Search** wizard should now look like this:

- Click the **Next** button.

You will see the **Import FASTA** page, where you should do the following:

- Click the **Browse** button.
- Navigate to the DIA-PASEF-small folder created earlier and into its DIA subfolder.
- Double-click on the "target_protein_sequences.fasta" file.

NOTE: At this point, you could choose to perform a proteome-wide analysis by instead navigating to the DDA_search subfolder where you found the "interact.pep.xml" file and double-clicking the full FASTA sequence file "napedro_3mixed_human_yeast_ecoli_20140403_iRT_reverse.fasta" that was used in the peptide search. This would produce quite a lot more targets and take more time to process, but is still feasible on most modern laptops.

- In the **Decoy generation method** dropdown list, click "Shuffle Sequence".
- Check **Automatically train mProphet model**.

The **Import Peptide Search** wizard should now look like this:

- Click the **Finish** button.

You should be presented with a form describing the targets calculated based on your settings and the FASTA sequence text provided that looks like this:

Notice that with a **Min peptides per protein** filter of "1" the 12 proteins found in the FASTA file produce 10 "Mapped" proteins. This is because 2 "Unmapped" proteins have no peptides found in the spectral library.

Note: The final document will include an "iRT standards" peptide list, for the iRT peptides, and a "Decoys" peptide list, for the decoys you have specified. This will also add 15 iRT peptides and 30 decoy peptides. In the form above, the precursors and transitions count reflect these additions while the proteins and peptides counts do not.

- Click the **OK** button.

Skyline begins extracting chromatograms which should look like this on a standard i7 4-core processor:
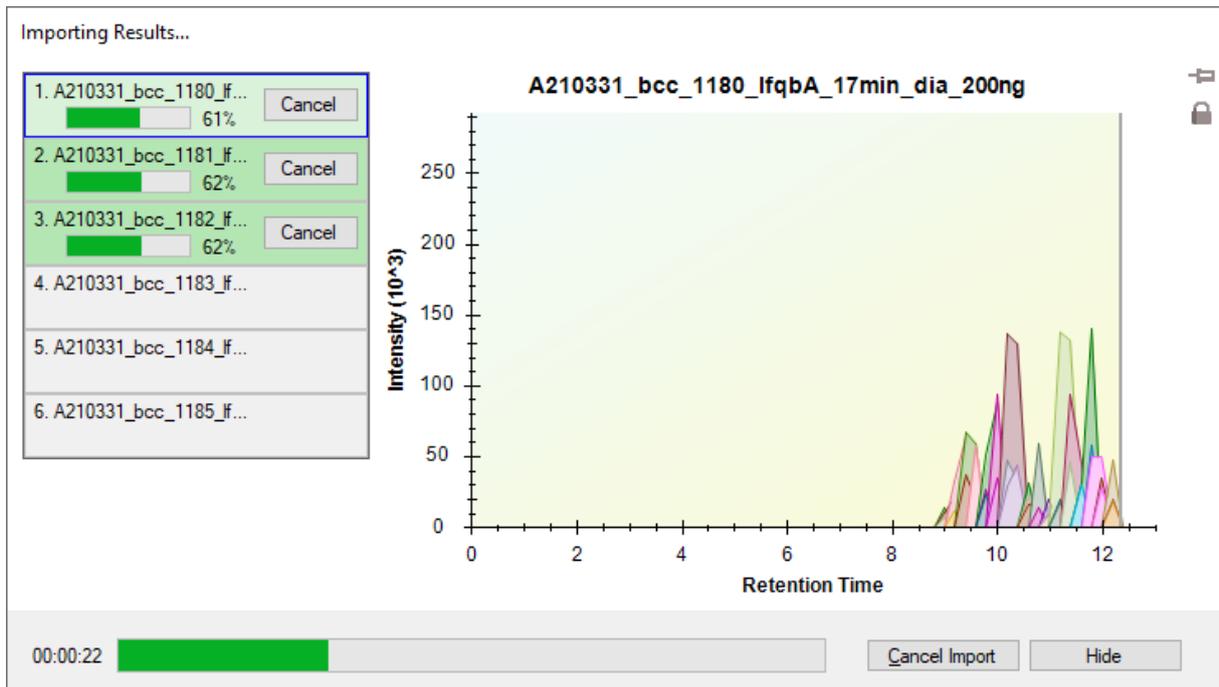
With 6-cores and above all files will be processed in parallel, and on most laptops with 2-cores the import will process 2 files at a time.

After the import is completed, Skyline will show you the mProphet model you requested on the **Import FASTA** page of the **Import Peptide Search** wizard. It should look like this:
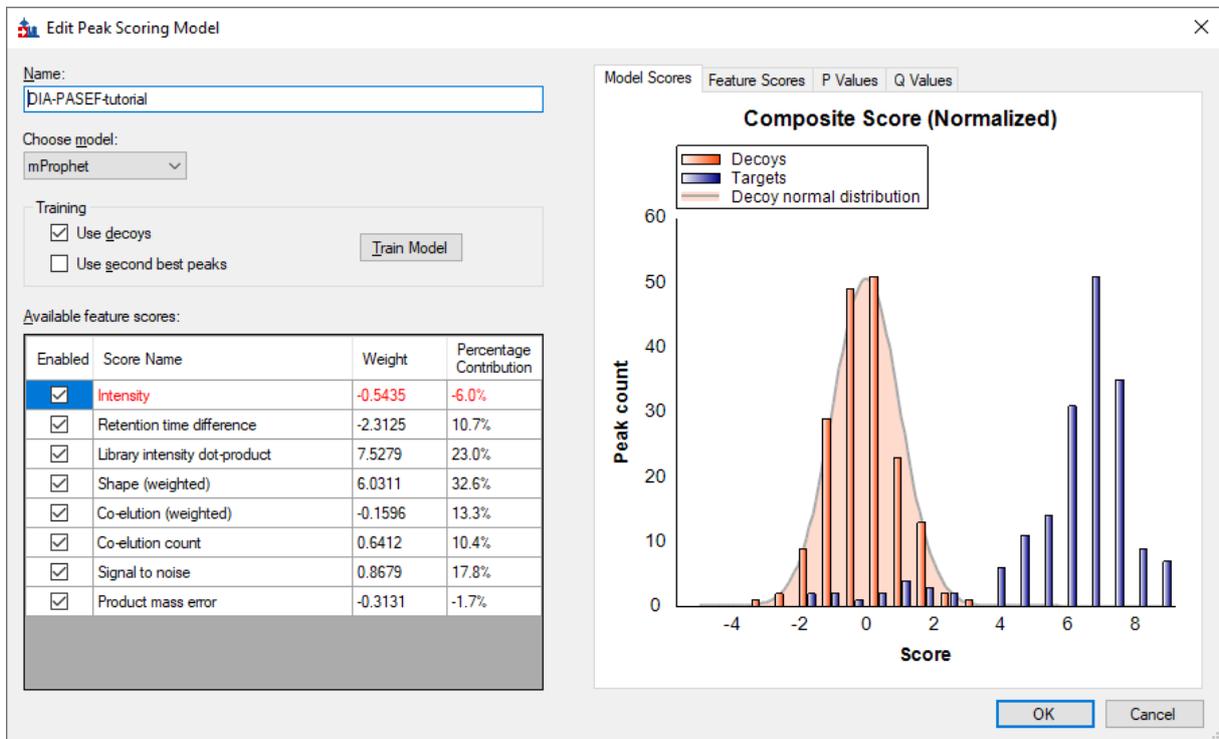
It has already been used to re-score the 10 best peaks Skyline found in each set of chromatograms, using its default score, and re-choose the peaks to be the ones with the best mProphet scores. These

mProphet scores (given the name **Detection Z Score** in Skyline) are scaled such that 1.0 is 1 standard deviation from the mean mProphet score for the distribution of similar best peaks for the sequence shuffled decoy peptides you requested. Each has also been assigned a q value (given the name **Detection Q Value** in Skyline).

Sample Annotation

Next you need to define which samples belong to which experimental group**:**

- On the **Settings** menu, click **Document Settings.**
- Click the **Annotations** tab if it is not active.
- Click the **Add** button to see the **Define Annotation** window.
- In the **Name** field, enter "Condition".
- In the **Type** field, choose "Value List".
- In the list under **Applies To,** check **Replicates.**

In this experiment there are two conditions: Condition A are samples that have a proteome composition of E.coli 20%, yeast 15% and human 65%, and condition B have a composition of E.coli 5%, yeast 30% and human 65%.

The **Define Annotation** window should like this:
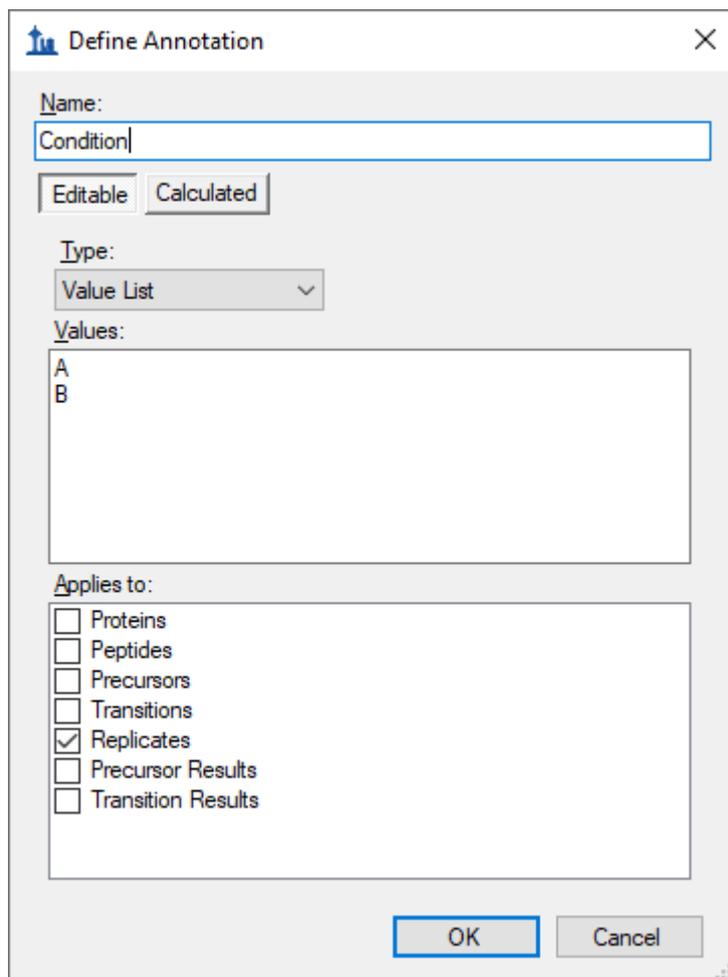
- Click the **OK** button.

- Follow the steps above and create a second annotation named "BioReplicate."
- In the **Type** dropdown list select "Text" for the BioReplicate annotation.
- In the list under **Applies To,** check **Replicates.**
- Click the **OK** button.

This will return you to the **Document Settings** window.

- Make sure the checkboxes for **Condition** and **BioReplicate** are checked.
- Click the **OK** button.

You are now ready to annotate the replicates you have imported:

- On the **View** menu, choose **Document Grid** (Alt-3).

This brings up the **Document Grid** window, which will show the **Proteins** report if you have never used it before:

- In the **Document Grid: Proteins** window, click the **Reports** menu and then click **Replicates.**

Annotate the samples as shown in the screen capture below using keyboard keys for A, B, 1, 2, 3, Enter and arrow keys to navigate:

- Close the **Document Grid: Replicates** window by clicking on the red X in the upper right hand corner of the form.

For easy viewing you can now split the data by condition into 2 panes in the Skyline window.

- On the **View** menu, choose **Arrange Graphs**, and click **Grouped**.
- In the **Arrange Graph Grouped** form, set **Group panes** to "2".
- Choose the option **Distribute graphs among groups**.
- In the **Display** field, choose "Tiled".
- In the **Sort order** field, choose "Document".
- Click the **OK** button.

Note: If the data is still importing at this stage you will need to wait until this is finished to save the document.

- Save your skyline document.

Inspect the data manually

Next you will manually inspect some of the chromatography and underlying spectra. To do that, you will want the **Retention Times** and **Peak Areas** views are showing:

- On the **View** menu, choose **Retention Times**, and click **Replicate Comparison** (F8).
- On the **View** menu, choose **Peak Areas**, and click **Replicate Comparison** (F7).

Now both views should be visible. You can configure them optimally on the screen by doing the following:

- If these new windows are floating you can dock them to the main window by clicking on the top border of the floating window, holding the left mouse button down, and dragging this window.

You can perform the same operations shown below with the **Library Match** window.

Note: In order to dock a window like this, the mouse arrow cursor must be placed inside one of the docking icons, in this case, at the top. Once you begin dragging with the left-mouse button held down, Skyline will show the docking icons and a transparent blue rectangle where the window will be placed if you were to release the left-mouse button at that moment.
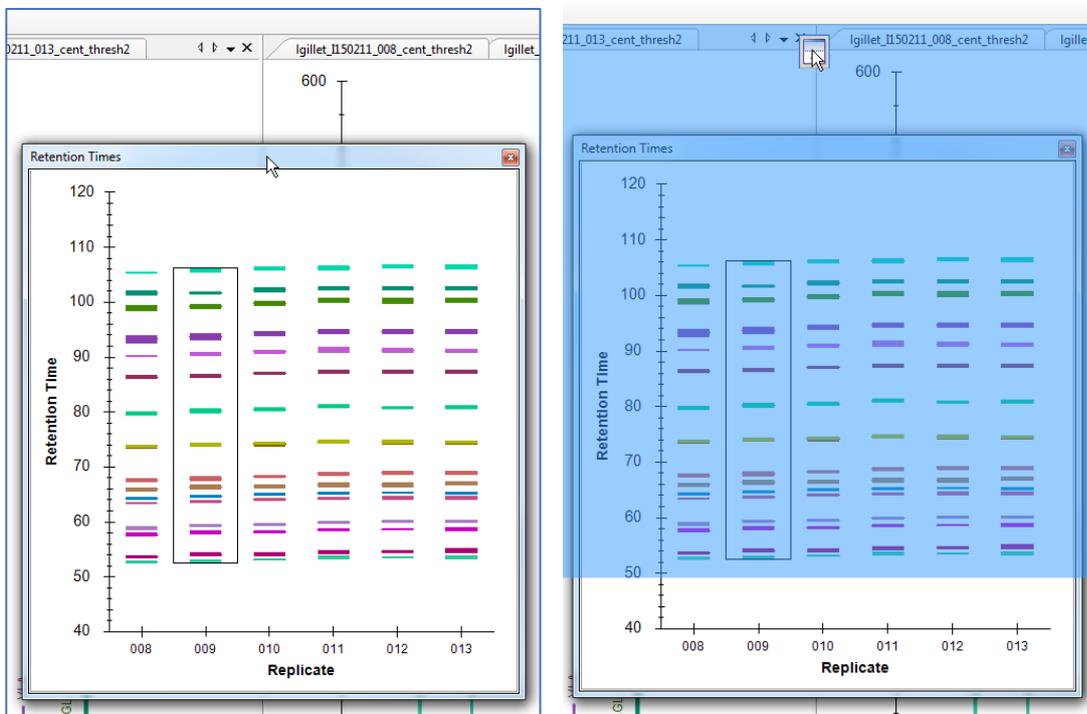
- Similarly, dock the **Peak Area: Replicate Comparison** and **Library Match** windows so that all information is easily viewable, as below:
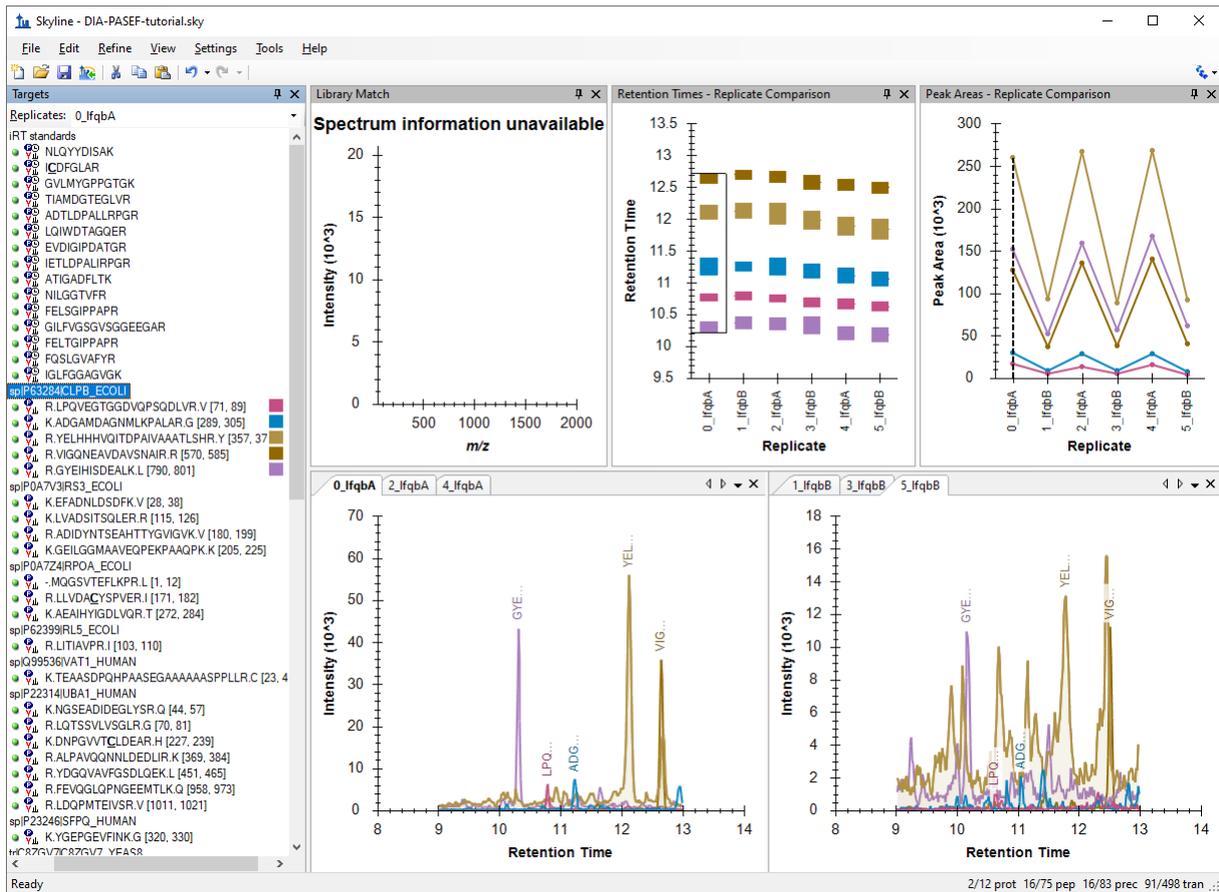
- Click on the first E. coli protein: "sp|P63284|CLPB_ECOLI".

You should see all of the peptides for this protein shown on the various plots (chromatogram, peak areas, and retention time replicate graphs). The screenshot above is an example of when one protein is selected and all of the peptides for this protein are summarized in each of these views (except library match window where nothing is shown).

Based on what you see in the **Peak Area - Replicate Comparison** plot, does this protein appear to be differentially regulated? Recall that the expected fold-change ratio between A:B is 4:1 for E. coli.

- Select the peptide LPQVEGTGGDVQPSQDLVR in this protein.

You get specific information for this peptide in all of the views, as shown below:
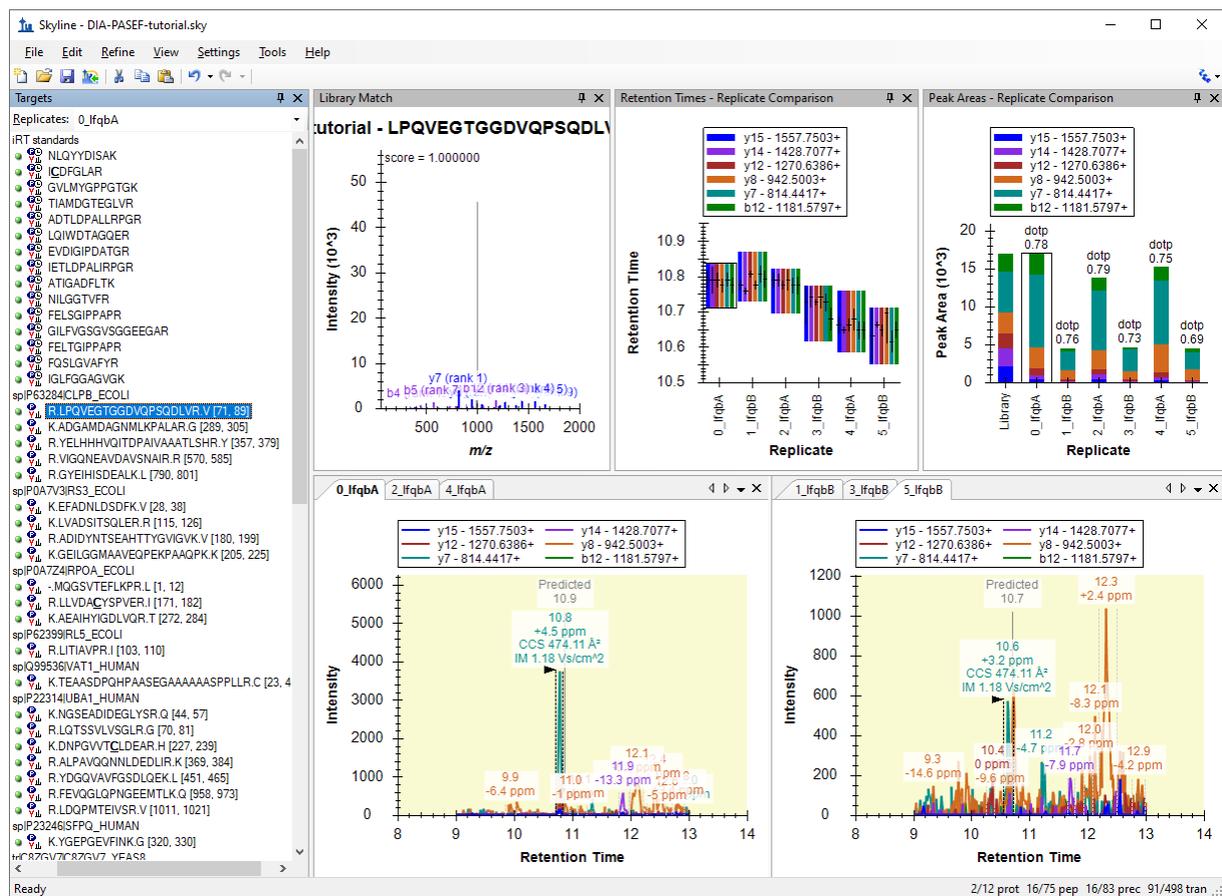
Examine the peak area patterns for the rest of the peptides belonging to this protein. Is the quantitative pattern for the peptides from this protein consistent with the expected differential regulation pattern (4:1)?

NOTE: If there is more than one precursor charge state for a given peptide sequence these are extracted and scored separately. You can look at these by clicking the **+** next to the peptide sequence and clicking on the individual charge states.

- Click on some of the human proteins in the document.

Examine the replicate peak areas from the protein level view and the peptide level view. Are the peak areas consistent with the expected ratio (1:1)? What about the yeast proteins (1:2)?

- Click again on the LPQVEGTGGDVQPSQDLVR peptide in the first protein again.
- On the **Views** menu, choose **Auto-Zoom**, and click **Best Peak** (F11).

Notice that if you hover the cursor over a chromatogram curve a circle appears on the curve with the same color as the curve.
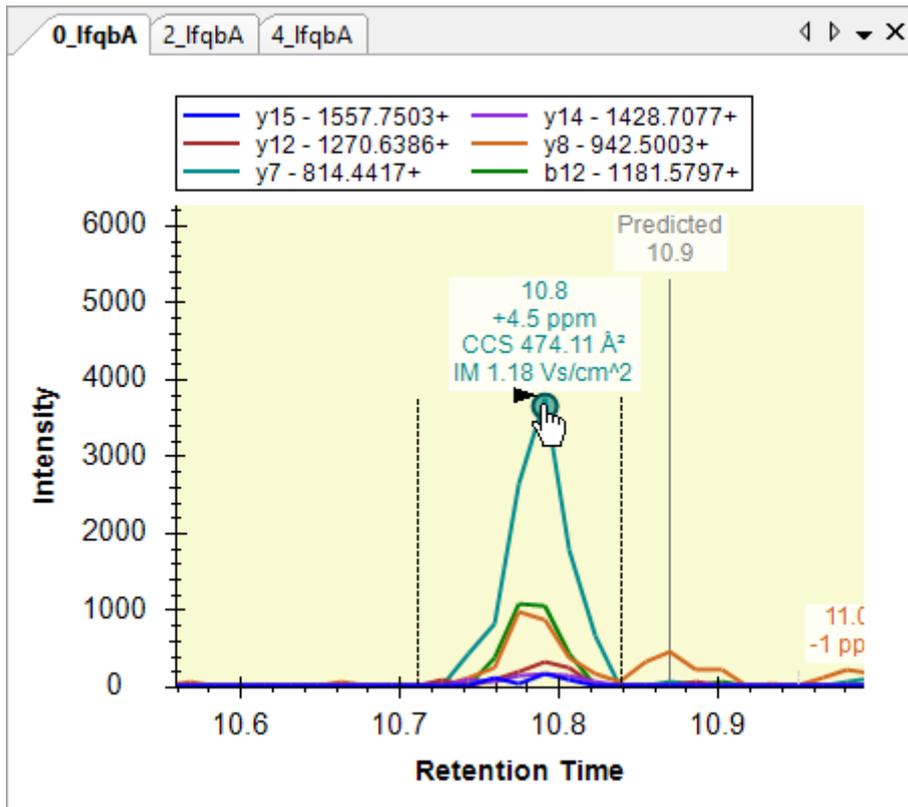
- Click the circle on a chromatogram point.

Skyline will open the **Full-Scan** view on the spectrum from which the chromatogram point was extracted zoomed on the ion that you selected.
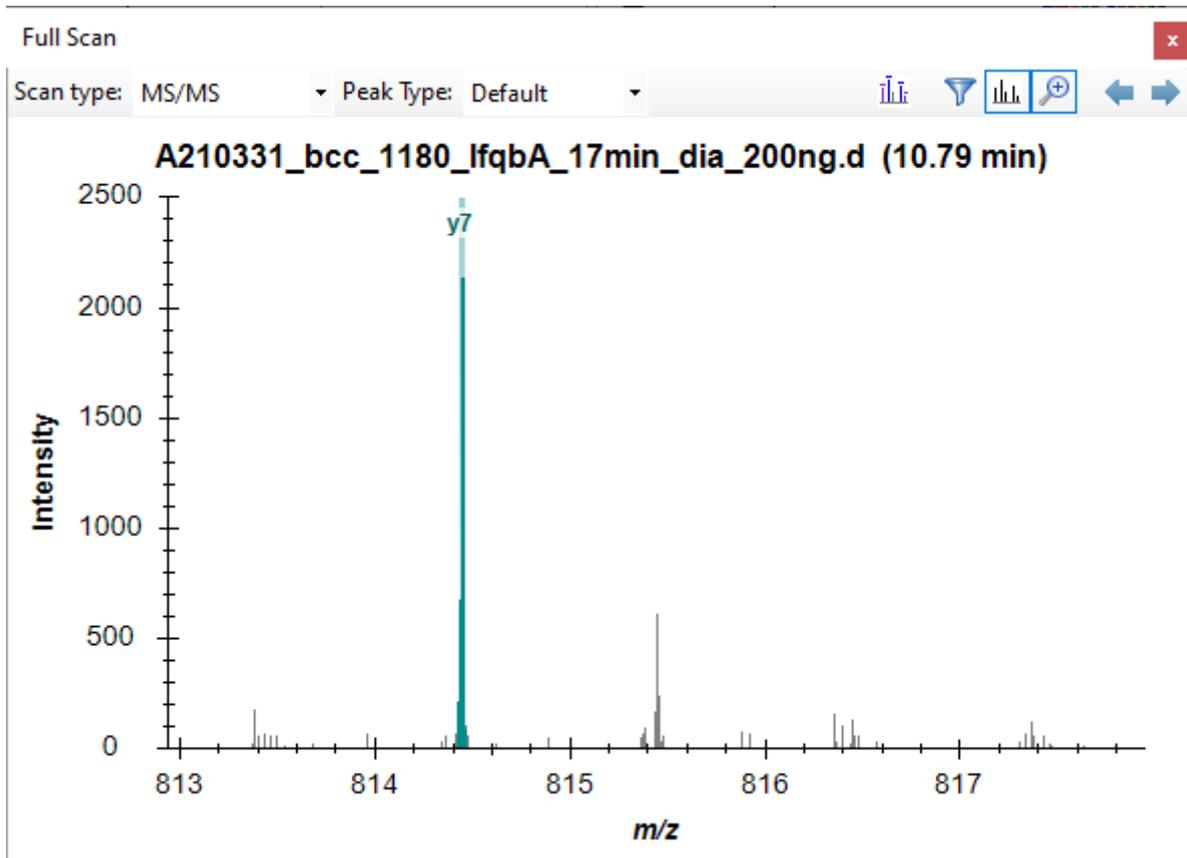
- Click the black spectrum button in the **Full-Scan** toolbar to view the mobility dimension.

Skyline will show the zoomed MS/MS spectrum with the mobility dimension on the y-axis:
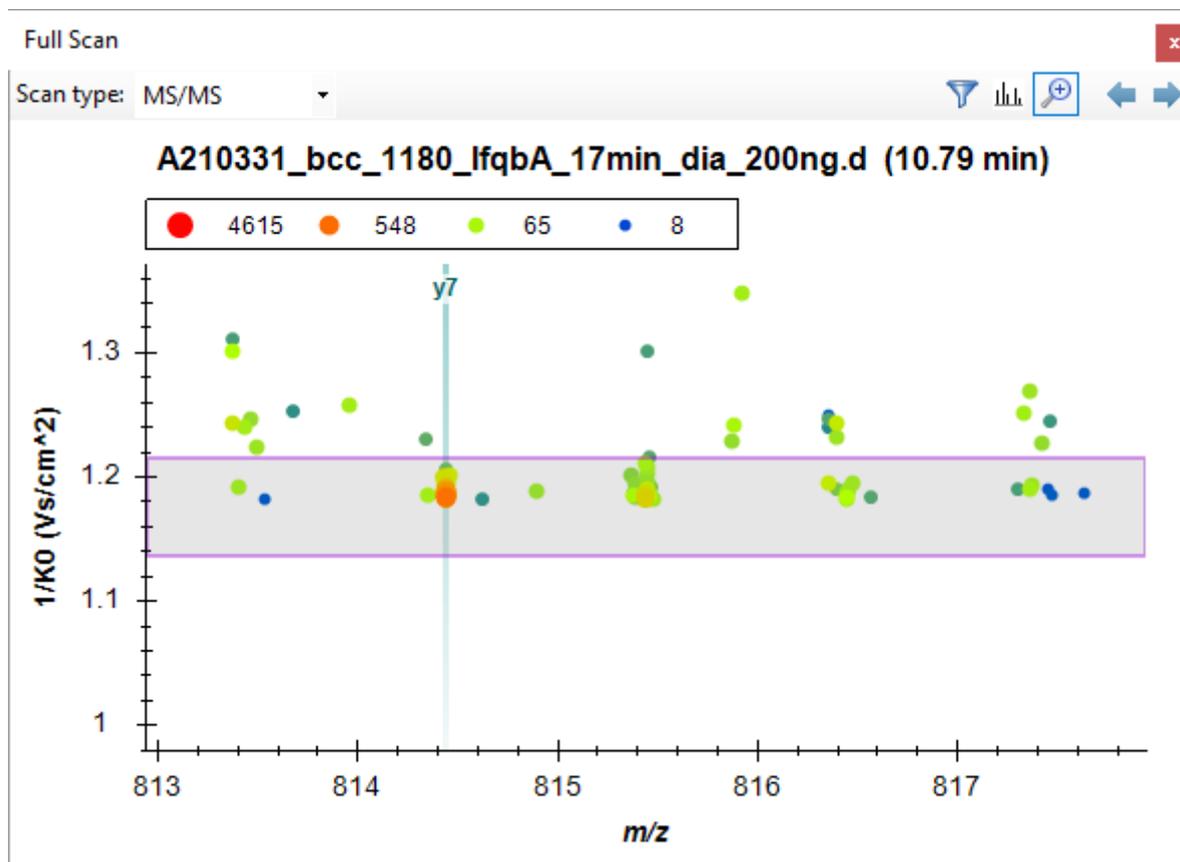
- Click the magnifying glass button in the **Full-Scan** toolbar to view the full spectrum.
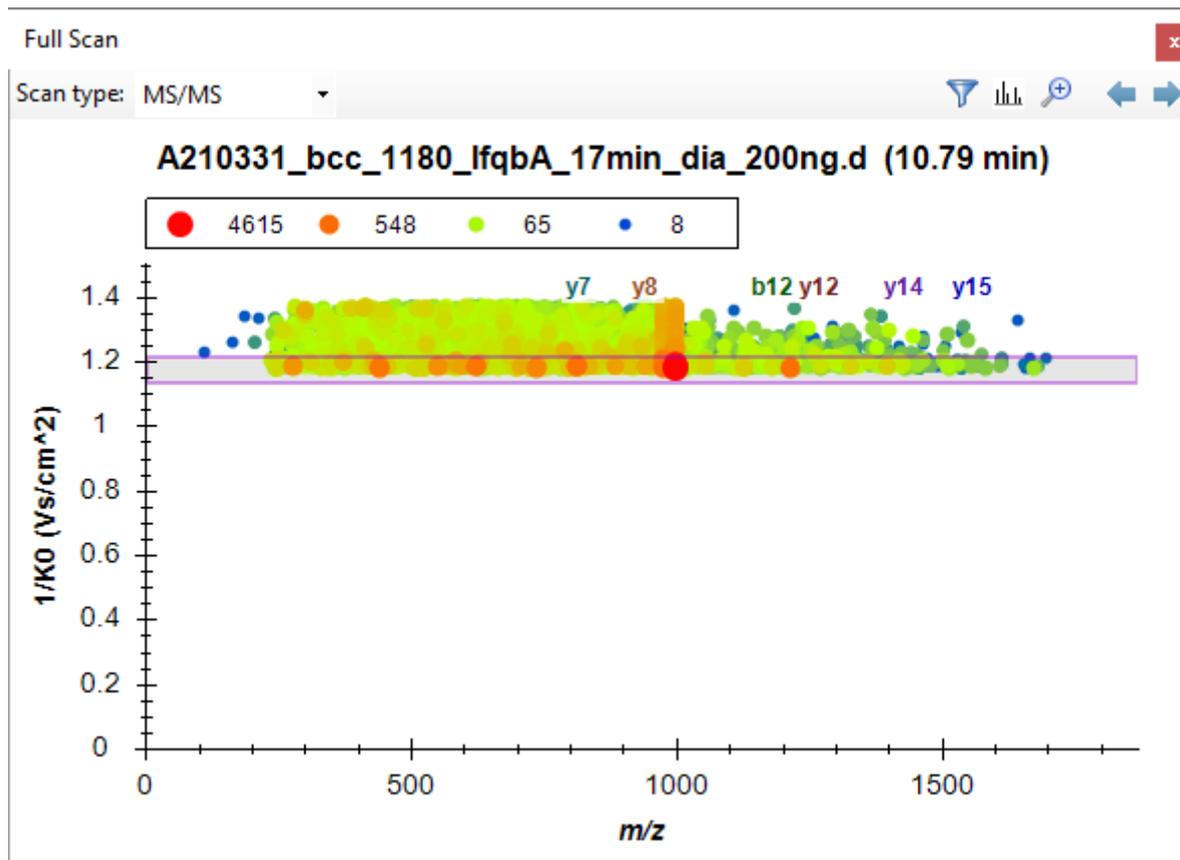
Explore the data further manually (including some decoys).

Mass error and retention time deviation

You can examine the mass accuracy and retention time prediction accuracy to determine whether the optimal extraction parameters have been used or whether some adjustment may improve the results.

- On the **View** menu, choose **Mass Errors**, and click **Histogram**.
- Right-click the graph that appears, choose **Points**, and click **Targets at 1% FDR**.

This will show the distribution of mass errors over the data set. Could the extraction window (±20 ppm) have been further optimized? Generally, the mean error +/- 3 standard deviations is sufficient. If the histogram appears cut off, however, with high counts at the extremes then you may want to widen your extraction windows.

Mass Errors - Histogram

Mean = 3.9,
Standard Deviation = 1.9
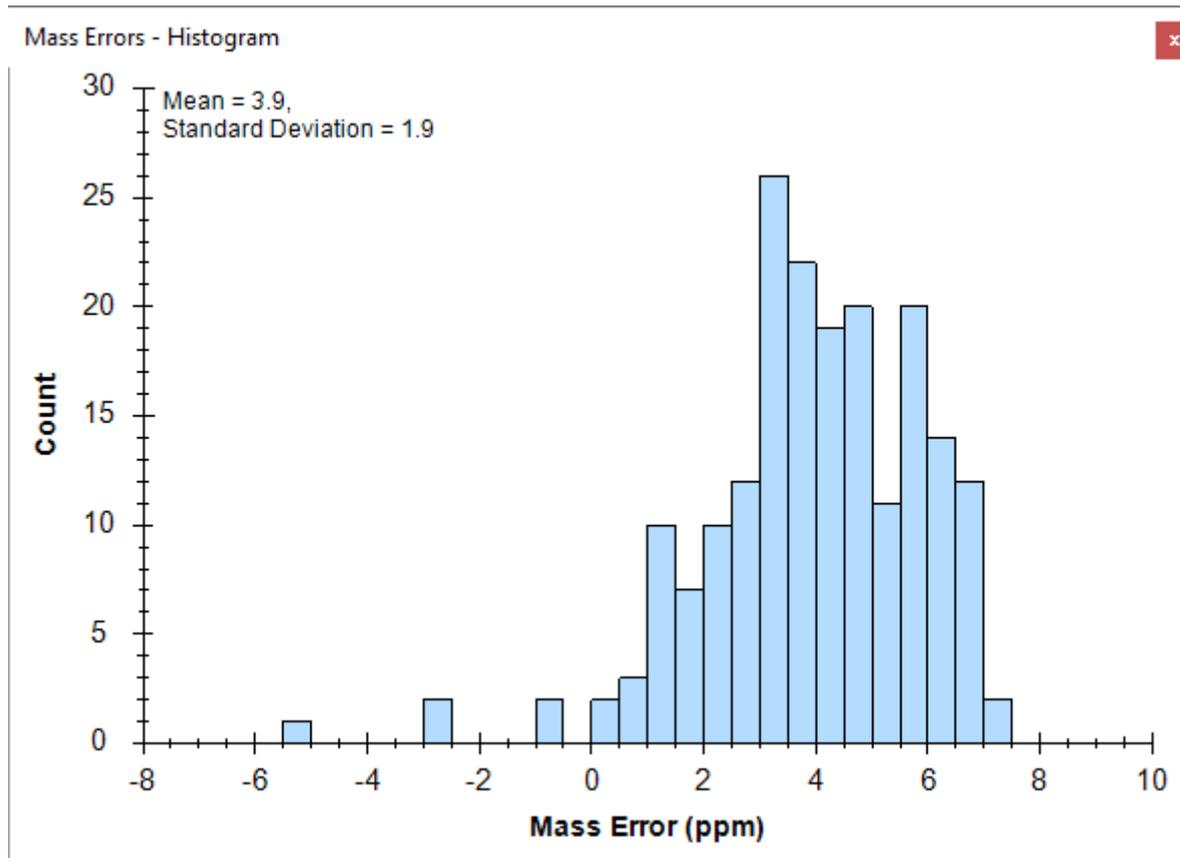
Note that the distribution appears to have two peaks, one around 3 and one around 6. To understand this better, it helps to know that these are the mass error values across all 6 runs.

- Right-click the graph, choose **Replicates**, and click **Single**.
- Click in the **Replicates** dropdown list at the top of the **Targets** view.
- Use the up and down arrow keys on your keyboard to review the mass error values in all 6 replicates.

You will see that the mean value ranges from 3.4 to 4.1, while the standard deviation value ranges from 1.8 to 3.4. Using the simple calculation Mean + 3 * SD = 4.1 + 3 * 3.4 = 14.3 PPM at the most extreme, this seems to indicate that a 20 PPM tolerance is a pretty good approximation of what these data require.

To view the mass errors for the decoys also:

- Right-click the graph, choose **Points**, and click **Decoys**.
- Right-click the graph, choose **Replicates**, and click **All**.
  (and then switch back to **Targets**)

To the see the linear regression used to predict the target peptide retention times based on the iRT peptides and library iRT values from the target peptides:

- On the **View** menu, choose **Retention Times**, then **Regression**, and click **Score to Run**.
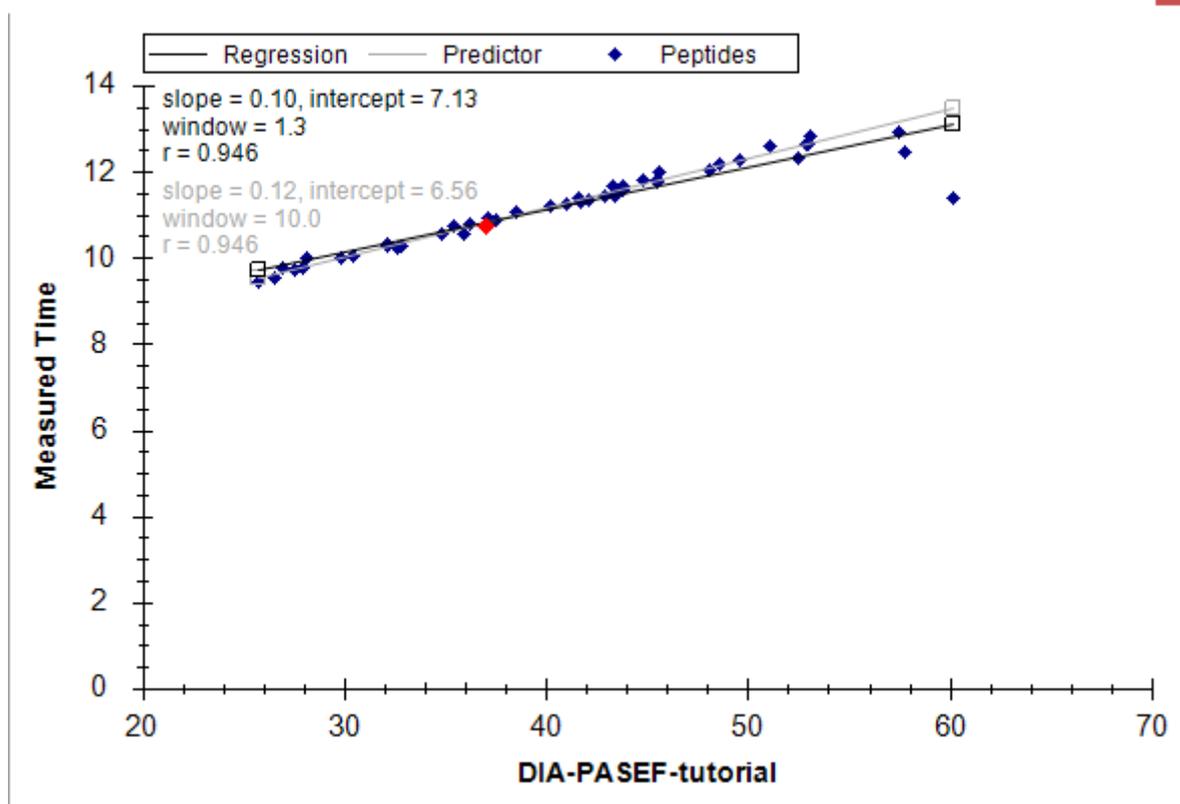
- Right-click the graph, choose **Plot**, and click **Residuals**.

This will show the deviations from the predicted retention times in this data set. Could the extraction window (± 5 minutes) have been further optimized for this analysis? Check the decoys as well.
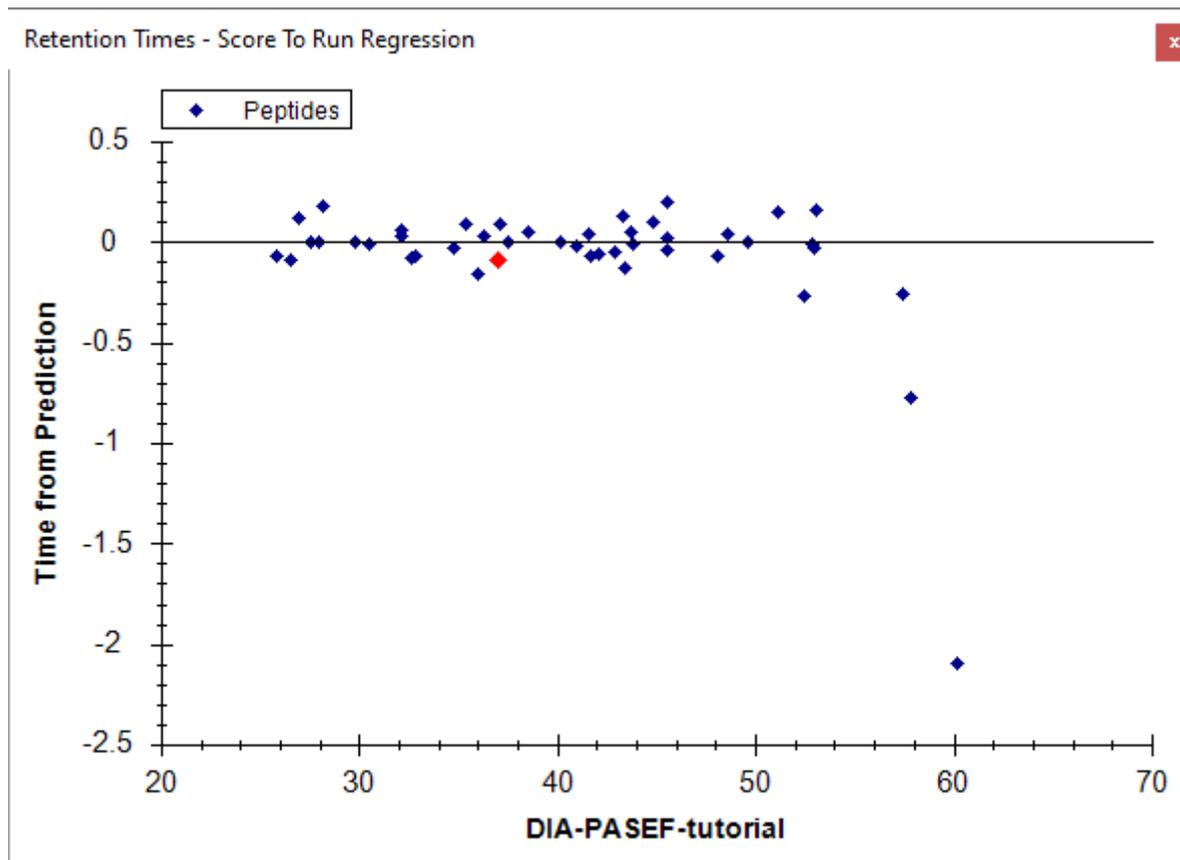
Note: As the spectral library for this analysis was constructed from a side-by-side analysis of the same samples. As such, the accuracy of the retention time predictions are very good. Retention times from external spectral libraries acquired on different instruments, at different times, from different samples would lead to larger errors in these predictions.

Quantitative comparison

You have performed some general validation that the data processing with Skyline has no serious flaws. All available features scores were included in the mProphet scoring model, and the model achieved reasonable separation between the targets and the decoys, which are used to simulate random undetectable targets. You may feel that the RT extraction range could be tightened from +/- 5 to 3. But, these settings will clearly work and the resulting mProphet model looks acceptable.

This type of analysis of even hundreds of peptides generally makes manual analysis of every single target time consuming and even error prone itself. The current analysis could be extended to the entire 3-organisms FASTA file, known as proteome-wide analysis, which would include tens of thousands of peptides. Rather than consider each individually, researchers more typically perform some type of grouped comparison and then follow up on peptides or proteins which appear to be changing in interesting ways.

To perform a simple pairwise group comparison inside Skyline do the following:

- On the **View** menu, choose **Other Grids**, then **Group Comparison**, and click **Add**.
- For the **Name** field, enter "By Condition".
- For the **Control group annotation** field, choose "Condition".

27

- For the **Control group value** field, choose "A".
- For the **Value to compare against** field, choose "B".
- Click the **Advanced** button.
- For the **Q-value cutoff** field, enter "0.01".

The **Edit Group Comparison** form should look like this.

- Click the **OK** button.

To see the group comparison you have just created:

- On the **View** menu, choose **Other Grids**, then **Group Comparison**, and click **By Condition**.

A table should appear that shows the peptide level fold-change and adjusted p-value (an estimate of false discovery rate - FDR) for the comparison between the A and B sample mixtures.

- Expand the width of the **Protein** column header so that you can see the full protein names – including the corresponding species name.
- Click on the header of the **Fold Change Result** column and click on **Sort Ascending**.

Inspect the fold changes estimated for some of the peptides in the table keeping in mind which species they are from and the expected ratios (human 1:1, yeast 1:2, E. coli 4:1). Consider the adjusted p values and what you might expect of them.


	Protein	Peptide	Fold Change Result	Adjusted P-Value
	iRT standards	GILFVGSGVSG...	1.32 (95% CI:1.01 to 1.72)	0.0786
	iRT standards	FELTGIPPAPR	0.87 (95% CI:0.08 to 9.8)	0.7242
	iRT standards	FQSLGVAFYR	2.12 (95% CI:0.34 to 13.36)	0.4020
	iRT standards	IGLFGGAGVGK	1.16 (95% CI:1.04 to 1.3)	0.0350
	sp|P63284|CLPB_ECOLI	LPQVEGTGGD...	0.3 (95% CI:0.25 to 0.36)	0.0002
	sp|P63284|CLPB_ECOLI	ADGAMDAGNM...	0.29 (95% CI:0.25 to 0.33)	0.0001
	sp|P63284|CLPB_ECOLI	YELHHHVQITD...	0.34 (95% CI:0.33 to 0.36)	0.0000
	sp|P63284|CLPB_ECOLI	VIGQNEAVDAV...	0.29 (95% CI:0.26 to 0.32)	0.0000
	sp|P63284|CLPB_ECOLI	GYEIHISDEALK	0.36 (95% CI:0.31 to 0.41)	0.0002
	sp|P0A7V3|RS3_ECOLI	EFADNLDSDFK	0.44 (95% CI:0.32 to 0.6)	0.0043
	sp|P0A7V3|RS3_ECOLI	LVADSITSQLER	0.32 (95% CI:0.3 to 0.33)	0.0000
	sp|P0A7V3|RS3_ECOLI	ADIDYNTSEAH...	0.25 (95% CI:0.2 to 0.31)	0.0002
	sp|P0A7V3|RS3_ECOLI	GEILGGMAAVE...	0.36 (95% CI:0.28 to 0.48)	0.0015
	sp|P0A7Z4|RPOA_ECOLI	MQGSVTEFLKPR	0.34 (95% CI:0.31 to 0.38)	0.0000

- Click the **Volcano Plot** button in the top left corner of the grid window.
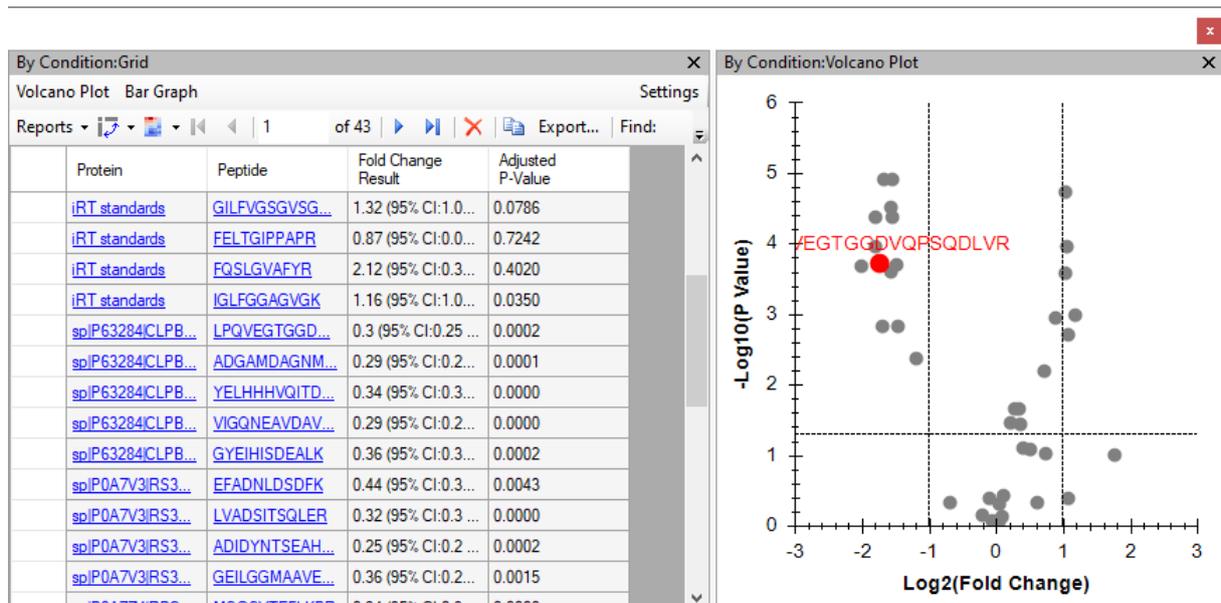
You can already see clusters of points around the expected ratios and above the 5% FDR horizontal line for most of the changing peptides. To get a better understanding of which points belong to which organism perform the following actions:

- Right-click the **Volcano Plot**, and click **Formatting**.
- Click the "…" button in the first row, after the **Expression** column.
- In the **Create Match Expression** form, set the **Match** dropdown list to "Protein Name".
- In the **Regular Expression** field, enter "ECOLI".

The form should show you the list of peptides with matching protein names like this:

- Click the **OK** button.
- Click the "…" button in the first row, after the **RGB** column.
- Choose a purple color and click the **OK** button.
- Repeat the above process for "YEAS" with orange.
- Repeat the above process for "HUMAN" with green.

Which should leave the **Volcano Plot Formatting** form looking like this:
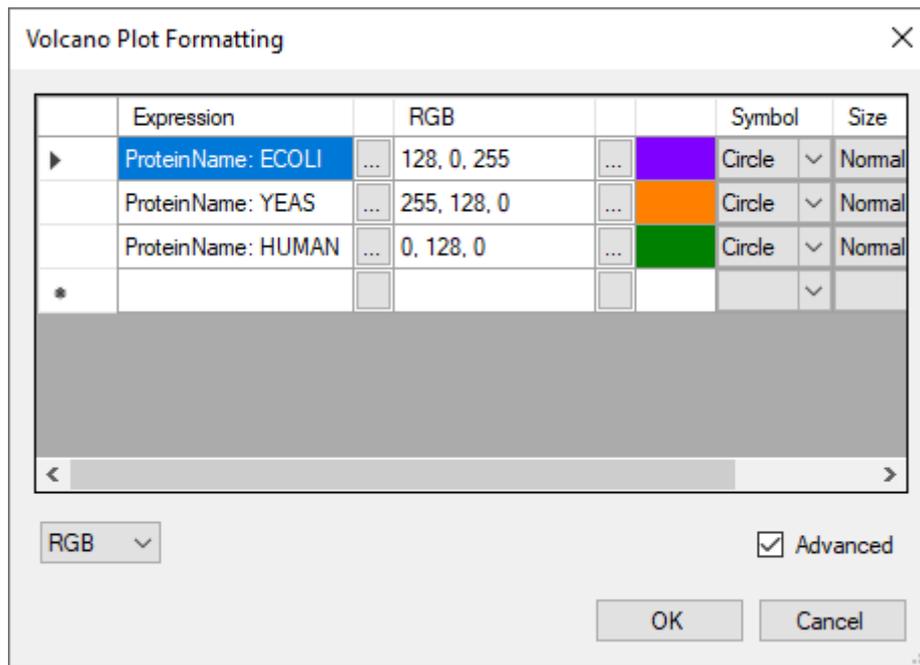
- Click the **OK** button.

Which should leave the volcano plot looking like this:
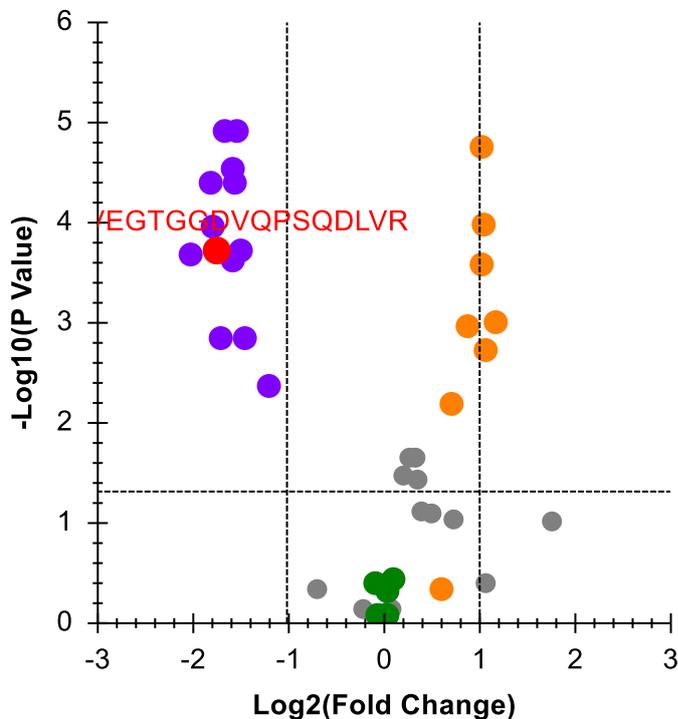
Notice that there are still some gray dots. You can click on any of the dots to cause them to be selected in the **Targets** view and have their display test shown on the graph in red. You can also hover the mouse cursor over any point to get more information on it. In this way you can see that the gray dots belong to the iRT standard peptides. You can filter them from the grid and the plot at the same time by doing the following:

31

- Click the **Protein** column header, and click **Filter**.
- In the **Filter type** dropdown list, choose "Does Not Contain".
- Enter in the field below "standard".
- Click the **OK** button.

This will remove the gray dots in the volcano plot. You might try limiting this plot to just one species using a different type of filter on the protein names.

Skyline also provides a bar plot view on the fold-change values which incorporates error bars for the (unadjusted) confidence intervals. Though, they are unadjusted for the multiple hypotheses tested, they still give you some insight into the variance in the measurements. Do the following to review the fold-change values in the bar plot:

- Click the **Bar Graph** button in the top left corner of the grid window.

The **Bar Graph** will appear on top of the **Volcano Plot**. To give it more space for viewing do the following:

- Click in the **By Condition:Bar Graph** tab, hold, and drag to a new location and release.
- Resize the now free-floating **By Condition:Bar Graph** window.
- In the **By Condition:Grid**, click the **Fold Change Result** column header, and click **Sort Ascending**.

The graph should now look like this:
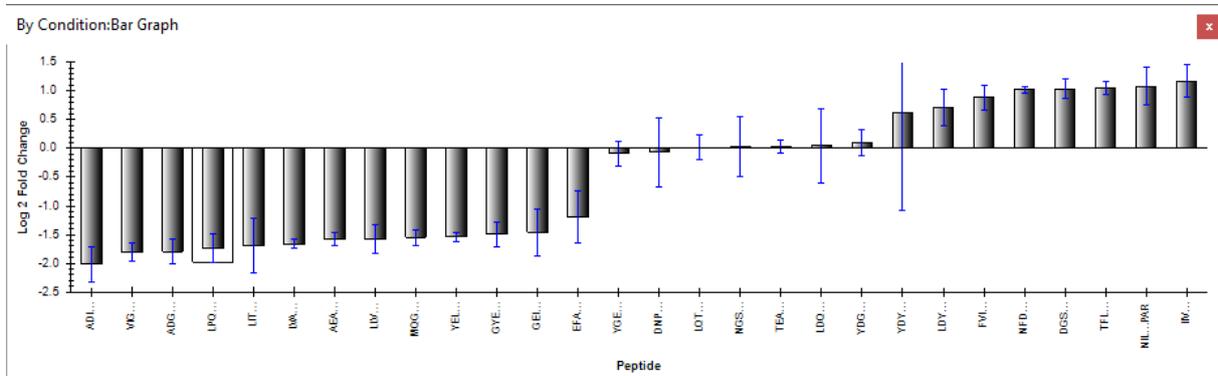
As with the volcano plot, you can click on the individual bars to select them in the **Targets** view (so that you can confirm which species various peptides in the graph are from).

To view the fold-change results at the protein-level for the 11 targeted proteins (excluding iRT standards and decoys), do the following:

- Click the **Settings** button in the upper-right corner of the grid window to re-open the **Edit Group Comparison** form.
- Change the **Scope** option from **Peptide** to **Protein**.

Skyline should adjust the **Bar Graph** and **Volcano Plot** immediately. You will see in the **Bar Graph** that there are just 9 bars and not the 10 you might have been expecting. This is because the current settings give Skyline no way to deal with missing values (due to 0.01 q value cut-off) other than

dropping targets without consistent enough measurement to support a T-test. To fix this, do the following:

- In the **Edit Group Comparison** form, change the **Summary method** to "Tukey's Median Polish".

The **Bar Graph** should now show fold-change estimates for all 10 proteins that looks like this:
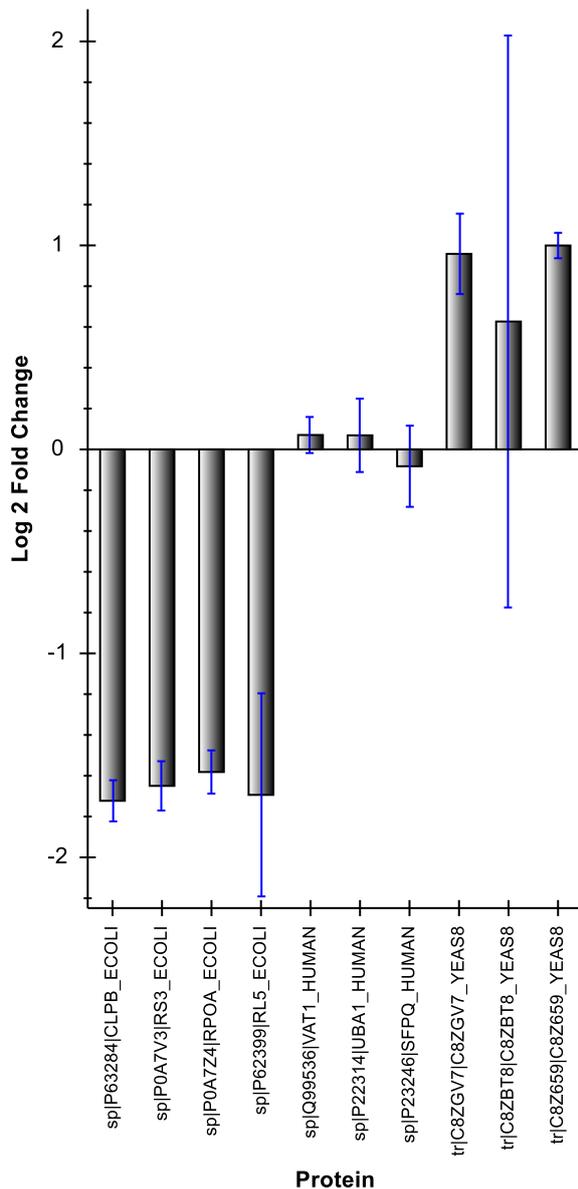
- Close the **Edit Group Comparison** by clicking the X square in the upper-right corner.
- Save the document.

You may also want to consider using a more sophisticated statistical tool like MSstats for your comparisons. To do this, you can install MSstats from the Skyline Tool Store (using **Tools** > **Store** in the Skyline menus or by clicking the Tool Store button on the Skyline installation web page). Then you can either run MSstats directly from the Skyline **Tools** menu or do the following to prepare to run MSstats in R or RStudio later:

- On the **File** menu, choose **Export** and click **Report**.
- From the **External Tools** folder, select **MSstats Input**.
- Save as "DIA_PASEF_ MSstats_Input.csv".

Conclusion

Congratulations! You have completed your first differential proteomics comparison using Skyline with diaPASEF data. You have learned how to use the **Import Peptide Search** wizard to streamline the initial data processing from building a spectral library out of ddaPASEF peptide search data to chromatogram extraction from quantitative diaPASEF runs with IMS filtering, creating and applying an mProphet statistical peak detection model to improve peak selection and assign q values which can be used to control the false discovery rate among the peak areas you used in your group comparison. You learned to first assess the DIA data quality with the mass error and retention time regression plots. And finally, you learned to perform a simple pairwise comparison between two groups of samples directly within Skyline, which gave you the interactive group comparison grid, volcano plot and bar graph to help you understand and interact with the results. Having exported the MSstats Input report, you could now use the MSstats external tool to perform further statistical tests.

Now consider retrying these steps with the proteome-wide FASTA file, the DIA-PASEF-full dataset, and/or including precursor ions to see how this works with a broader target set and/or another dimension (precursor ions in MS1 spectra). Or move on to trying this with your own data.