#### Supplementary Note for:

#### Scribe: next generation library searching for DDA experiments

Brian C. Searle<sup>1, 2, 3, 4\*</sup>, Ariana E. Shannon<sup>1, 2, 3</sup>, and Damien Beau Wilburn<sup>1, 2, 3</sup>

<sup>1</sup>Department of Biomedical Informatics, The Ohio State University Medical Center, Columbus, Ohio 43210, United States.

<sup>2</sup>Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States. <sup>3</sup>Pelotonia Institute for Immuno-Oncology, The Ohio State University Comprehensive Cancer Center, Columbus, Ohio 43210, United States.

<sup>4</sup>Proteome Software Inc., Portland, OR

\*Corresponding Address: Brian C. Searle, Ph.D., brian.searle@osumc.edu

# **Peptide detection with Scribe**

#### Tutorials based on Scribe version 2.11.2, last update on December 22, 2022

Written by Ariana Shannon (<u>shannon.225@buckeyemail.osu.edu</u>) and Brian Searle (<u>brian.searle@osumc.edu</u>)

Spectrum library searching is a powerful alternative to database searching for data dependent acquisition (DDA) experiments. Scribe is a new library search engine designed to leverage deep learning fragmentation prediction software such as Prosit. Scribe embeds Percolator for false discovery rate (FDR) correction and an interference-tolerant label-free quantification integrator to enable an end-to-end proteomics workflow.

### TABLE OF CONTENTS

OVERVIEW	2
1. PREREQUISITES AND INSTALLING SCRIBE	3
2. GENERATING MZMLS	4
3. GENERATING LIBRARIES USING PROSIT AND SCRIBE	8
4. SCRIBE SEARCH OPTIONS	15
5. INTERPRETATION OF SCRIBE OUTPUTS	18
6. VISUALIZING DATA WITH BUILT-IN FEATURES	21
7. WALKTHROUGH WITH DEMO DATA	40
8. USING SCRIBE TO BUILD LIBRARIES FOR DIA SEARCHES	43
9. FAQ	46

### **OVERVIEW**

Scribe comes with a user-friendly GUI interface. The upper left pane contains search options, while the right pane contains a process queue. The bottom left contains a console that provides specific information about the process EncyclopeDIA is running.

le View Convert Data Help					
🌹 EncyclopeDIA 🗽 Thesaurus 🍓 Walnut 🚟 Scribe		Jobs: 🤳 Add MZML	Save Chromatogram Library	Save Quant Reports	La Save BLIE
Scribe: Spectrum-Centric Library Search	ching 🛛	File Read 2022_09_01_HeLa_500r		ogress 2 to 416.7 m/z (780 MS/MS, 13903	total entries). Jobs
for Data-Dependent Acquisition (DDA) Data Scribe extracts peptide fragmentation spectra from MZML files, mat to spectra in libraries, and calculates various scoring features. Thesare interpreted by Percolator to identify peptides.	tches them				
Parameters:					
ibrary:uniprot_human-reference_reviewed_2022mar02.prosit_input.trypsin_nce29_hcd2020.dlib	Edit				
Background:uniprot_human-reference_reviewed_2022mar02.fasta	Edit				
arget/Decoy Approach: Normal Target/Decoy	~				
Enzyme: Trypsin	~				
ragmentation: CID/HCD (B/Y)	~				
Precursor Mass Tolerance: 10.0 PPM	~				
ragment Mass Tolerance: 10.0 PPM	~				
ibrary Mass Tolerance: 10.0 PPM	~				
Percolator Version: v3-01	~				
lumber of Quantitative lons:	5 🜩				
Minimum Number of Quantitative Ions:	3 🗘				
lumber of Cores:	15 🌩				
Additonal Command Line Options:					
Processing precursors scans Found 229 total ranges Preparing to maintain at most 26 jobs for 17 threads	^				
Constructing writer for D. Aviana ¥2022_09_01_HeLa_5000,g_10maxt_DDA_NCE_27 mzML scribe, features.txt. Nonking on 411.2 to 411.7 m/z (732 MS/MS, 12728 total entries). Jobs in waiting: 212 Nonking on 414.2 to 414.8 m/z (1011 MS/MS, 17921 total entries). Jobs in waiting: 212 Nonking on 414.2 to 414.8 m/z (812 MS/MS, 11974 total entries). Jobs in waiting: 212	unsorted				
volkning on 41 + 6 to 41.0 ± m/z (4+3 kidenins), besoto total entries), Jobs in waiting; 21 2 Morking on 41 + 17 to 41 2.7 m/z (91 + 11 MS/MS, 14320 total entries), Jobs in waiting; 21 2 Morking on 40 × 8 to 40 8 m/z (69 T MS/MS, 8850 total entries), Jobs in waiting; 21 2 Morking on 40 × 1 to 40 × 7 m/z (554 MS/MS, 7855 total entries), Jobs in waiting; 21 2 Morking on 40 × 1 to 40 × 7 m/z (71 MS/MS, 1532 total entries), Jobs in waiting; 21 2					
Noking on 415 to 416 2 m2 (11 Nokins), tosse total emines), soss in waring, 212 Noking on 415 5 to 416 2 m2 (841 MS/MS, 13100 total entries), Jobs in waiting; 212 Noking on 415 2 to 416 6 m/z (840 MS/MS, 12402 total entries), Jobs in waiting; 212 Noking on 407.3 to 408.8 m/z (1173 MS/MS, 19034 total entries), Jobs in waiting; 212 Noking on 410 2 to 410.7 m2 (708 MS/MS, 13903 total entries), Jobs in waiting; 212					
Noking on 406 2 to 407.3 m² (160 m3emis), 1535 data entres), 505 m waiting; 212 Noking on 408 2 to 407.3 m² (1832 MS/MS, 19944 total entries). Jobs in waiting; 212 Noking on 408 9 to 411.2 m² (1832 MS/MS, 23959 total entries). Jobs in waiting; 212 Noking on 403.7 to 408.2 m² (1447 MS/MS, 23421 total entries). Jobs in waiting; 212	v				

This manual outlines how to use Scribe, what outputs are given by Scribe, and how to use the outputs to interpret data.

# **1. PREREQUISITES AND INSTALLING SCRIBE**

Software requirements needed to process data using Scribe.

Scribe is a cross-platform Java application that has been tested for Windows, Macintosh, and Linux. Scribe requires 64-bit Java 1.8. While it is possible to use higher versions of Java, many versions are untested. Using untested versions of Java may result in unknown errors. In particular, Java versions 17 and 18 are known to cause stability issues in the current release.

A. If you don't already have Java, install Java 1.8 on your computer. if you are using Windows, you can download the Windows "x64 Installer" from:

<u>https://www.oracle.com/java/technologies/downloads/#java8</u>. Other operating system options are available at this URL as well.

B. Scribe is folded into the EncyclopeDIA software package. After you have 64-bit Java 1.8, go to EncyclopeDIA's bitbucket page (<u>https://bitbucket.org/searleb/encyclopedia/wiki/Home</u>) and download the most recent stable version. Once downloaded, double-click on the EncyclopeDIA .JAR file to launch the GUI interface. If you are using a Macintosh, you may need to right-click on the EncyclopeDIA .JAR and select "Open" to execute it for the first time with the proper permissions. Click on the tab named Scribe at the top to search DDA data.

C. We recommend using Proteowizard to create mzML files from your RAW files. You can freely download Proteowizard from here: <u>https://proteowizard.sourceforge.io/download.html</u>.

### 2. GENERATING MZMLS

How to use Proteowizard to generate vendor-neutral .mzML files from vendor-specific RAW files.

A. Before searching files in Scribe, RAW files must be converted to .mzML files using Proteowizard. To do so, open Proteowizard. Remove the "titleMakers" filter by selecting the filter in the parameters box, and clicking remove.

🛃 MSConvertGUI (64-bit)		– 🗆 ×
List of Files      File of file names  File:  Add Remove	Browse network resource V	About MSConvert
Output Directory:	Filters       Subset         MS levels:       -         Scan number:       -         Scan number:       -         Number of data points:         Scan time (seconds):       -         Scan event:       -         Activation type:         Scan polarity:       Any         Analyzer type:	
Options Output format: mzML v Extension:	Add Remove	uter Elevitore Rathell Maria
Binary encoding precision: <ul> <li>64-bit</li> <li>32-bit</li> </ul> Write index: <ul> <li>Use zlib compression:</li> <li>TPP compatibility:</li> <li>Package in gzip:</li> <li>Use numpress linear compression:</li> <li>Use numpress short logged float compression:</li> <li>Use numpress short logged float compression:</li> <li>Use numpress positive integer compression:</li> <li>Combine ion mobility scans:</li> <li>SIM as spectra:</li> <li>SRM as spectra:</li> </ul>	ttleMaker <pre><runid>.<scannumber>.<scannumber>.<chargest< pre=""></chargest<></scannumber></scannumber></runid></pre>	ate> File:" <sourcepath>", Nati</sourcepath>
Presets: Generic Defaults	Save Preset      Files to convert in p	oarallel: 5 🗢 Start

B. Add "Peak Picking" by selecting the peak picking option under filters, then click add. This should be the only filter in the box.

Add Remove	Browse network r	resource V About MSConvert
And Leitore	Filters	Peak Picking ~
		Algorithm:
		Vendor (does not work for UNIFI, and it MUST be the first filter!) ${\scriptstyle\checkmark}$
		MS Levels:         Min SNR:         Min peak spacing:           1         -         0.1         0.1
utput Directory: Browse		
browse		
Options		Add Remove
	Filter	Add Remove Parameters
Options	Filter	
Output format: mzML V Extension:	Filter	
Dptions Output format: mzML → Extension: Binary encoding precision:	Filter	
Dptions Output format: mzML → Extension: Binary encoding precision:	Filter	
Dptions Output format: mzML	Filter	
Options Output format: mzML ∨ Extension: Binary encoding precision:	Filter	
Options Output format: mzML ✓ Extension: Binary encoding precision:	Filter	

C. Select your files by clicking "browse", and locate the desired DDA files in your directory. Click "add" on the left-hand side of the screen. The output directory will automatically populate where your files were selected from.

e: Browse	Browse network re	About MSConvert
Add         Remove           0:\Ariana\2022.11.14_searleb_e3b9c5d4\raw\2022_0         0:\Ariana\2022.11.14_searleb_e3b9c5d4\raw\2022_0           0:\Ariana\2022.11.14_searleb_e3b9c5d4\raw\2022_0         0:\Ariana\2022.11.14_searleb_e3b9c5d4\raw\2022_0           0:\Ariana\2022.11.14_searleb_e3b9c5d4\raw\2022_0         0:\Ariana\2022.11.14_searleb_e3b9c5d4\raw\2022_0	Filters	Peak Picking V
		Algorithm: Vendor (does not work for UNIFI, and it MUST be the first filter!)
< >>		MS Levels:         Min SNR:         Min peak spacing:           1         -         0.1         0.1
Dutput Directory:		
D:\Ariana\2022-11-14_searleb_e3b9c5d4\ Browse		Add Remove
Options Output format: mzML V Extension:	Filter	Parameters
Binary encoding precision:   Binary encoding precision:  B	peakPicking	vendor msLevel=1-
Write index: Vuse zlib compression:		
TPP compatibility: Package in gzip:		
Use numpress linear compression:		
Use numpress short logged float compression:		
Use numpress short logged float compression:		

D. Under the options box, the output format should be mzML. You want "write index," "use zlib compression", "TPP compatibility, and "SIM as spectra" selected. Your settings should match the screenshot below.

🖶 MSConvertGUI (64-bit)		$ \Box$ $\times$
List of Files File of file names  File:      Remove  D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\2022_0 D:\Ariana\2022-11-14_searleb	Browse network resource  Filters Peak Picking Algorithm: Vendor (does not work for UNIFI, and it MUST be the first filter!) MS Levels: Min SNR: Min peak spacin 1 - 0.1 0.1	
Output Directory:         D:\Anana\2022-11-14_searleb_e3b9c5d4\         Options         Output format:       mzML          Binary encoding precision: <ul> <li>Betansion:</li> <li>Binary encoding precision:</li> <li>64-bit</li> <li>32-bit</li> <li>Write index:</li> <li>Use zlib compression:</li> <li>TPP compatibility:</li> <li>Package in gzip:</li> <li>Use numpress linear compression:</li> <li>Use numpress short logged float compression:</li> <li>Use numpress positive integer compression:</li> <li>Combine ion mobility scans:</li> <li>SIM as spectra:</li> <li>SRM as spectra:</li> </ul>	Add     Remove       Filter     Parameters       peak Picking     vendor msLevel=1-	
Presets: Generic Defaults	Save Preset	: 5 🚖 Start

E. Click "Start" in the lower right hand corner to start the file conversion.

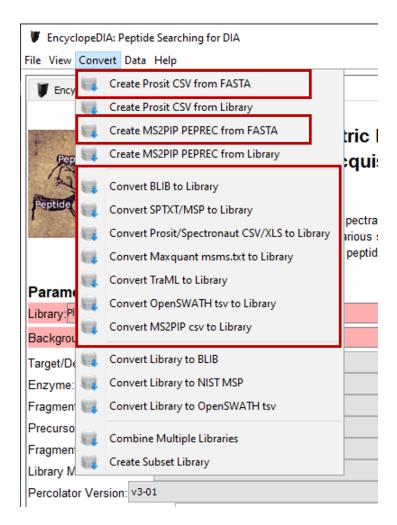
🖳 MSConvertGUI (64-bit)		- 0	$\times$
● List of Files ○ File of file names			
File: Browse	Browse network resource $\lor$	About MSC	nvert
Add Remove			
D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\2022_0 D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\2022_0 D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\2022_0 D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\2022_0 D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\2022_0	Filters	ak Picking ~	
	Vendor (does not	work for UNIFI, and it MUST be the first filter!)	
< >>	MS Levels:	Min SNR: Min peak spacing:	
Output Directory:			
D:\Ariana\2022-11-14_searleb_e3b9c5d4\ Browse		Add Remove	
Options Output format: mzML V Extension:	Filter Parameters		
Binary encoding precision:	peak Picking vendor msLe	vel=1-	
Write index:  Use zlib compression:  Propagation gzip:  Package in gzip:			
Use numpress linear compression:			
Use numpress short logged float compression:			
Use numpress positive integer compression:			
Combine ion mobility scans:			
SIM as spectra: 🗹 SRM as spectra: 🗌			
SIM as spectra:  SRM as spectra:	✓ Save Preset ▼	Files to convert in parallel: 5 🜩	Start
Presets: Generic Defaults	✓ Save Preset ▼		Start -
		Progress	Start
Presets: Generic Defaults Name	2022_09_01_HeLa_500ng_16mzst_DDA	Progress converting spectra: 60535/126636	Start
Presets: Generic Defaults Name D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\	2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA	Progress Converting spectra: 60535/126636 Converting spectra: 48932/126419	Start
Presets: Generic Defaults Name D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\ D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\	2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA	Progress Converting spectra: 60535/126636 Converting spectra: 48932/126419 Converting spectra: 44721/127590	Start
Presets: Generic Defaults           Name           D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw           D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw           D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw	x2022_09_01_HeLa_500ng_16mzst_DDA x2022_09_01_HeLa_500ng_16mzst_DDA x2022_09_01_HeLa_500ng_16mzst_DDA x2022_09_01_HeLa_500ng_16mzst_DDA	Progress Converting spectra: 60535/126636 Converting spectra: 48932/126419 Converting spectra: 44721/127590 Converting spectra: 51851/126608	Start
Presets: Generic Defaults           Name           D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\         D:\Ariana\2022-11-14_searleb_se3b9c5	x2022_09_01_HeLa_500ng_16mzst_DDA x2022_09_01_HeLa_500ng_16mzst_DDA x2022_09_01_HeLa_500ng_16mzst_DDA x2022_09_01_HeLa_500ng_16mzst_DDA	Progress Converting spectra: 60535/126636 Converting spectra: 48932/126419 Converting spectra: 44721/127590 Converting spectra: 51851/126608	Start
Presets: Generic Defaults           Name           D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\         D:\Ariana\2022-11-14_searleb_se3b9c5	x2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA	Progress converting spectra: 60535/126636 converting spectra: 48932/126419 converting spectra: 44721/127590 converting spectra: 51851/126608 converting spectra: 56956/126518	Start
Presets: Generic Defaults           Name           D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\         D:\Ariana\2022-11-14_searleb_se3b9c5	2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA 2022_09_01_HeLa_500ng_16mzst_DDA	Progress converting spectra: 60535/126636 converting spectra: 48932/126419 converting spectra: 44721/127590 converting spectra: 51851/126608 converting spectra: 56956/126518	Start

F. Once complete, the window will look like this. You can exit out of MSConvert once the conversion is done. Your files will be in the location previously specified.

Conversion Progress	- 🗆 ×
Name	Progress
D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\2022_09_01_HeLa_500ng_16mzst_DDA	Finished
D:\Ariana\2022-11-14_searleb_e3b9c5d4\raw\2022_09_01_HeLa_500ng_16mzst_DDA_NC	E_42.raw

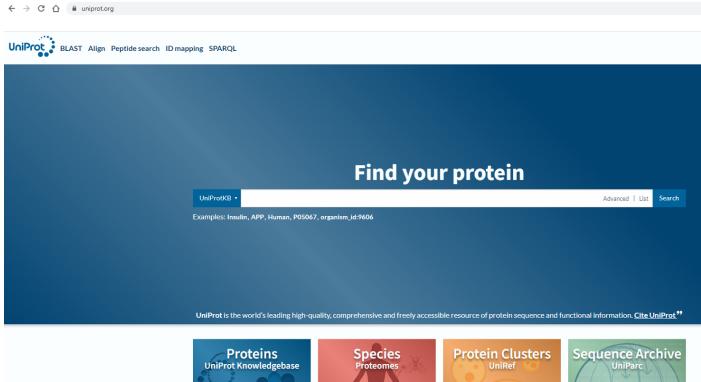
# **3. GENERATING LIBRARIES USING PROSIT AND SCRIBE**

How to acquire a FASTA file and then use Prosit and Scribe to generate a spectral library file from it.



The primary library format Scribe uses is .DLIB. Spectrum libraries from Skyline, NIST, TraML and other formats can be converted to .DLIB using the "Convert" menu. In particular, Scribe can search fully predicted libraries generated by either Prosit or MS2PIP, and produce input files for both of those tools. The following tutorial is given on how to obtain a FASTA file and use Prosit to generate a predicted .DLIB.

A. To generate a Prosit-predicted spectral library (.DLIB) for Scribe, start by obtaining a FASTA file for your organism of interest. Given the computational expense of generating libraries, in most cases we recommend using only canonical protein sequences. For example, if you require a FASTA of proteins sequences for Homo Sapiens, go to uniprot.org. Select "Reviewed."





#### Specify "Human (20,401)."

Status Reviewed (Swiss-Prot) (20,401)

UniProtKB 20,401 results

RLAST Align Map IDs 📩 Download 🖶 Add View: Cards 🔿 Table 🖲 🗹 Customize columns 👒 Share

Popular organisms	BLAST Align Map		🗠 Download 🔛 Add	view. Carus 🔿 Table 🖲 👱 Customize columns 👒 Share 🖲			
Human (20,401)	Entry 🔺		Entry Name 🔺	Protein Names 🔺	Gene Names 🔺	Organism 🔺	Length 🔺
	A0A0C5B5G6	3	MOTSC_HUMAN	Mitochondrial-derived peptide MOTS-c[]	MT-RNR1	Homo sapiens (Human)	16 AA
Taxonomy Filter by taxonomy	□ A0A1B0GTW7	8	CIROP_HUMAN	Ciliated left-right organizer metallopeptidase[]	CIROP, LMLN2	Homo sapiens (Human)	788 AA
	A0JNW5	8	UH1BL_HUMAN	UHRF1-binding protein 1-like[]	UHRF1BP1L, KIAA0701, SHIP164	Homo sapiens (Human)	1,464 AA
Proteins with 3D structure (7,557)	A0JP26	a	POTB3_HUMAN	POTE ankyrin domain family member B3	POTEB3	Homo sapiens (Human)	581 AA
Active site (2,278)	□ A0PK11	a	CLRN2_HUMAN	Clarin-2	CLRN2	Homo sapiens (Human)	232 AA
Activity regulation (1,527)	A1A4S6	a	RHG10_HUMAN	Rho GTPase-activating protein 10[]	ARHGAP10, GRAF2	Homo sapiens (Human)	786 AA
Allergen (6)	□ A1A519	8	F170A_HUMAN	Protein FAM170A[]	FAM170A, ZNFD	Homo sapiens (Human)	330 AA
Alternative products (isoforms) (10,634)	□ A1L190	a	SYCE3_HUMAN	Synaptonemal complex central element protein 3[]	SYCE3, C22orf41, THEG2	Homo sapiens (Human)	88 AA
More items	A1L3X0	8	ELOV7_HUMAN	Elongation of very long chain fatty acids protein 7[]	ELOVL7	Homo sapiens (Human)	281 AA
Protein existence	□ A1X283	8	SPD2B_HUMAN	SH3 and PX domain-containing protein 2B[]	SH3PXD2B, FAD49, KIAA1295, TKS4	Homo sapiens (Human)	911 AA
Protein level (16,585)	A2A2Y4	8	FRMD3_HUMAN	FERM domain-containing protein 3[]	FRMD3, EPB41L4O	Homo sapiens (Human)	597 AA
Transcript level (2,310)	□ A2RU14	a	TM218_HUMAN	Transmembrane protein 218	TMEM218	Homo sapiens (Human)	115 AA
Homology (754)	A2RUB6	8	CCD66_HUMAN	Coiled-coil domain-containing protein 66	CCDC66	Homo sapiens (Human)	948 AA
Uncertain (613)	A2RUC4	a	TYW5_HUMAN	tRNA wybutosine-synthesizing protein 5[]	TYW5, C2orf60	Homo sapiens (Human)	315 AA
Predicted (139)	A4D1B5	8	GSAP_HUMAN	Gamma-secretase-activating protein[]	GSAP, PION	Homo sapiens (Human)	854 AA
Annotation score	A4GXA9	a	EME2_HUMAN	Probable crossover junction endonuclease EME2[]	EME2	Homo sapiens (Human)	379 AA
(14,308)							

Select "Download." If you download the compressed version, you will get a .gz and have to uncompress the file. If you download the uncompressed version, the file may take longer, but it will be in a FASTA format for immediate use.

Download		×		Advanced   List Search	🕯 🔐 🖂 Не
Download selected (0)					
Download all (20,401)					
ormat					
FASTA (canonical & isoform)		🔻 stomize columns 🛛 📽 Share 🔹			
ompressed			Gene Names 🔺	Organism 🔺	Length 🔺
Yes		e MOTS-c[]	MT-RNR1	Homo sapiens (Human)	16 AA
No		etallopeptidase[]	CIROP, LMLN2	Homo sapiens (Human)	788 AA
	Generate URL for API Preview 10 Cancel Downl	oad []	UHRF1BP1L, KIAA0701, SHIP164	Homo sapiens (Human)	1,464 AA
		nember B3	POTEB3	Homo sapiens (Human)	581 AA
			CLRN2	Homo sapiens (Human)	232 AA
		n 10[]	ARHGAP10, GRAF2	Homo sapiens (Human)	786 AA
			FAM170A, ZNFD	Homo sapiens (Human)	330 AA
		l element protein 3[]	SYCE3, C22orf41, THEG2	Homo sapiens (Human)	88 AA
		atty acids protein 7[]	ELOVL7	Homo sapiens (Human)	281 AA
		g protein 2B[]	SH3PXD2B, FAD49, KIAA1295, TKS4	Homo sapiens (Human)	911 AA
		:ein 3[]	FRMD3, EPB41L4O	Homo sapiens (Human)	597 AA
			TMEM218	Homo sapiens (Human)	115 AA
		; protein 66	CCDC66	Homo sapiens (Human)	948 AA
		g protein 5[]	TYW5, C2orf60	Homo sapiens (Human)	315 AA
		protein[]	GSAP, PION	Homo sapiens (Human)	854 AA

B. Open Scribe to create a Prosit CSV from a FASTA file. Navigate to the "Convert." Select "Create Prosit CSV from FASTA." This will open a dialog window.

Fincyclopel	DIA: Peptide Searching for DIA			
File View Con	vert Data Help			
👿 Enc	Create Prosit CSV from FASTA			
	Create Prosit CSV from Library			
Martin 🖬	Create MS2PIP PEPREC from FASTA	tric I		
Per 🔍	Create MS2PIP PEPREC from Library	qui		
12	Convert BLIB to Library			
Peptide	Convert SPTXT/MSP to Library	pectra		
	Convert Prosit/Spectronaut CSV/XLS to Libra		Default Settings	
	Convert Maxquant msms.txt to Library	peptid	Convert FASTA to Prosit CSV	×
Param	Convert TraML to Library		- Parameters:	
Library:Pl	Convert OpenSWATH tsv to Library		This function will in silico digest peptides from your FASTA and	
Backgrou	Convert MS2PIP csv to Library		create an input file for Prosit. If you use this feature, please cit Searle et al. 2020.	e
Target/De	Convert Library to BLIB		FASTA Please select file	Edit
Enzyme:	Convert Library to NIST MSP		Charge range: 2 🔷 to 3 🗢	
Fragmen 🜉	Convert Library to OpenSWATH tsv		Maximum Missed Cleavage:	1 🔹
Precurso	Combine Multiple Libraries		m/z range: 396.4 🐳 to 1,002.7 🜩	
Fragmen	Create Subset Library		Default NCE:	33 🜩
Library M 🐸	Create Subset Library		Default Charge:	3 🖕
Percolator Ve	rsion: v3-01		OK Cancel	

Upload the FASTA you have downloaded. The Prosit CSV will be generated in the same folder where the FASTA is held. If you use the example from the FASTA file provided in the Scribe manuscript, you should load "uniprot\_human-reference\_Reviewed\_2022mar02.fasta."

Settings for	
"uniprot_human-reference_reviewed_2022mar02.prosit_input.trypsi	
Convert FASTA to Prosit CSV	×
Parameters: This function will <i>in silico</i> digest peptides from your FASTA and "cite" create an input file for Prosit. If you use this feature, please cite	

Parameters:				
"cite" This function will <i>in silico</i> digest peptides from your FASTA and create an input file for Prosit. If you use this feature, please cite <u>Searle et al. 2020.</u>				
FASTA:uniprot_human-reference_reviewed_2022mar02.fasta	Edit			
Charge range: 2 + to 3 +				
Maximum Missed Cleavage:	1 🛓			
m/z range: 396.4 🗘 to 1,002.7 😓				
Default NCE:	29 🗘			
Default Charge:	3 🜲			
OK Cancel				

C. Use Prosit to perform an *in silico* digestion, and obtain an .MSP/NIST format text file. Navigate to the Prosit website, <u>Prosit</u>. When you get to the page, there are three tabs available; "CE CALIBRATION," "SPECTRAL LIBRARY," and "RESCORING."

intensity model works with both CID and HCD fragmentation methods but you modification explicitly in your input files.	need to add fragmentation column to the input. We assum	ne all the sequences are fully labeled and you dor	r'l need to add the tml
Prosit 🗰	PREC	DICT LIBRARIES FAQ	STATUS 🞧 😯
Prosit offers high quality MS2 predicted spectra for any organism and protease as we synthetic dataset. When using Prosit is helpful for your research, please cite "Gessul			on the project's high quality
CE CALIBRATION	SPECTRAL LIBRARY	RESCORI	NG
<ul> <li>This task estimates the optimal collision energy (CE) based on a given search result Prosit will:         <ol> <li>Select a random subset of high-scoring PSMs</li> <li>Predict those in for each CE from 18 to 39.</li> <li>Calculate which CE achieves highest correlations with the experimental spectra Please note: Sequences with amino acid U or O are not supported. Modifications end Also note: Antivirus software may cancel large uploads - turn it off if you experience</li> </ol> </li> <li>Upload Files         msms bt and RAW file.     </li> </ul>	k xcept "M(ox)" are not supported. Each C is treated as Cys		on in MaxQuant).
msms.txt MaxQuant's msms.txt from a finished search. Note, amino acid U or O are r	nnt cumpartari		×
mendaden a mendular nom a menediadensi, nore, Billitti Bula U B O BIET			
RAW RAW file that was searched (restricted to Thermo Fisher HCD Orbitrap). Fil	ia sina ia limikad la 200		×
rour ine una vea searcheù (restricteù lo Theffito Pisitel PCU Orbitrap). Fil			
			NEXT >
Model     Select intensity and IRT model for prediction			

### Go to the "SPECTRAL LIBRARY" tab.

	nent intensities prediction (Prosit_TMT_intensity_2021) and the other is for iRT prediction (Prosit_TMT_irt_2021). The o add fragmentation column to the input. We assume all the sequences are fully labeled and you don't need to add the tmt
Prosit 🚻	PREDICT LIBRARIES FAQ STATUS 😱 ?
Prosit offers high quality MS2 predicted spectra for any organism and protease as well as IR synthetic dataset. When using Prosit is helpful for your research, please cite "Gessulat, Schn	T prediction. Prosit is part of the ProteomeTools ( <u>www.proteometools.org/</u> ) project and was trained on the project's high quality midt et al. 2019" <u>DOI 10.1038/s41592-019-0426-7</u> .
CE CALIBRATION	SPECTRAL LIBRARY RESCORING
This task generates a spectral library either by digesting a given FASTA file, or by predicting for prediction, please use "CE Calibration". When a FASTA file is provided, Prosit will: 1. Digest the FASTA, for the given parameters (i.e. protease). 2. Predict all spectra at the given collision energy. When a CSV with peptides is provided, Prosit will directly predict all spectra. Please note: Antivirus software may cancel large uploads - turn it off if you experience uplo	g a list of peptides given in a CSV file. You need to provide a collision energy (CE) for prediction. To estimate an optimal CE
Settings Indicate collision energy, the maximum number of missed cleavages, and number of oxidized methior How would you like to provide the list of peptides?	inines per peptide.
<ul> <li>CSV</li> <li>FASTA (comming soon)</li> </ul>	
CSV Format	
modified_sequence collision_energy precure M(ox)CSDSDGLAPPQHLIR 15	rsor_charge fragmentation 2 ⊢ ⋛ HCD

For the 1st step, CSV is already selected. You can click next to get to the 2nd step. Upload the Prosit CSV created in Scribe.

CSV Format					
modified_sequence collisi	ion_energy precurs	sor_charge		fragmentation	
M(ox)CSDSDGLAPPQHLIR	15	2	For TMT models Only	HCD	
EMPQSDPSVEPPLSQETFSDLWK	28	2	els o	HCD	
TCPVQLWVDSTPPPGTR	35	3	Po	CID	
QSQHM(ox)TEVVR			-		
Please provide all three columns below and use motified_sequence_Use upper case letters in the with carbamidomethylation. Prosit does not supp collision_energy_Use integer values from 10 al precursor_charge_Use integer values from 11 to I fragmentation_Either HCD or CID, Use upper co	e column and indicate oxio port U or O as amino acids ind 50. 6.		e with "M(o:	CID	the range of 7 to 30. Each C is treated as Cysteine
Please provide all three columns below and use modified sequence. Use upper case letters in the with carbamidomethylation. Prosit does not supp collision_energy. Use integer values from 10 ai precursor_charge. Use integer values from 1 to 0	as a separator. e column and indicate oxio port U or O as amino acids ind 50. 6.	dized Methionin	e with "M(o:		the range of 7 to 30. Each C is treated as Cysteine
Please provide all three columns below and use modified_sequence Use upper case letters in the with carbamidomethylation. Prosit does not supp collision_energy Use integer values from 10 al precursor_charge Use integer values from 1 to fragmentation Either HCD or CID, Use upper co	as a separator. e column and indicate oxio port U or O as amino acids ind 50. 6.	dized Methionin	e with "M(o		the range of 7 to 30. Each C is treated as Cysteine
Please provide all three columns below and use	as a separator. e column and indicate oxid orort U or O as armino acids nd 50. 6. ase letters.*	dized Methionin	e with "M(o		the range of 7 to 30. Each C is treated as Cysteine

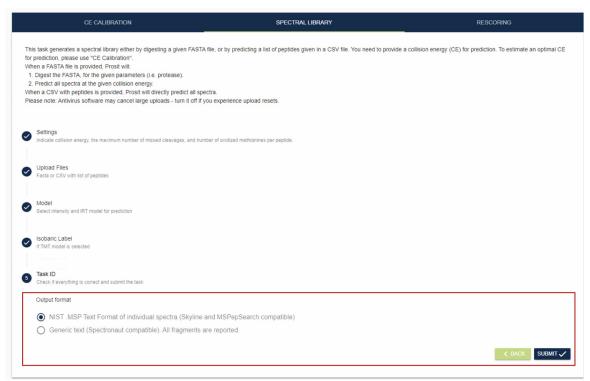
Once the CSV has once loaded, click next.

	modified_sequence collision	ion_energy precurs	sor_charge		fragmentation	
	M(ox)CSDSDGLAPPQHLIR	15	2	For TMT models Only	HCD	
EMP	QSDPSVEPPLSQETFSDLWK	28	2	dels	HCD	
	TCPVQLWVDSTPPPGTR	35	3	Fc	CID	
	QSQHM(ox)TEVVR	35	5		CID	
:	modified_sequence Use upper case letters in the with carbamidomethylation. Prosit does not support collision_energy. Use integer values from 10 an precursor_charge Use integer values from 1 to 6 fragmentation Either HCD or CiD, Use upper ca	ort U or O as amino acids nd 50. 6.		e with "M(ox	". Sequence length is	unceu io ure range of 7 to 30. Each C is treated as Cysteine
:	modified_sequence Use upper case letters in the ith carbamidomethylation. Prosit does not suppor collision_energy Use integer values from 10 an precursor_charge Use integer values from 1 to 6	e column and indicate oxid ort U or O as amino acids nd 50. 6.		e with "M(ox	". Sequence length is	unceu io ure range of 7 to 30. Each C is treated as Cysteine
:	adified_sequence Use upper case letters in the infl carbamidomethylation. Prosit does not suppor collision_energy. Use integer values from 10 an precursor_charge. Use integer values from 10 to 6 fragmentation_Either HCD or CID, Use upper ca fragmentation_Either HCD or CID, Use upper ca fragment	e column and indicate oxid ort U or O as amino acids nd 50. 6.		e with "M(ox	". Sequence length is	un, eou no une range of 7 to 30. Each C is treated as Cysteine
:	adified_sequence Use upper case letters in the infl carbamidomethylation. Prosit does not suppor collision_energy. Use integer values from 10 an precursor_charge. Use integer values from 10 to 6 fragmentation_Either HCD or CID, Use upper ca fragmentation_Either HCD or CID, Use upper ca fragment	e column and indicate oxid ort U or O as amino acids 6. ase letters.* 2mar02 fasta.trypsin z3	š.	e with "M(ox	". Sequence length is	un, eou no une range of 7 to 30. Each C is treated as Cysteine
*Onl	modified_sequence Use upper case letters in the tht carbamidomethylation. Prost does not supp collision_energy. Use integer values from 10 an precursor_charge Use integer values from 10 an precursor_charge Use integer values from 10 to 6 fregmentation Either HCD or CID, Use upper ca y for TMT model	e column and indicate oxid ort U or O as amino acids 6. ase letters.* 2mar02 fasta.trypsin z3	š.	e with "M(ox	". Sequence length is	un, eou no une range of 7 to 30. Each C is treated as Cysteine

Select the desired model. For this example, we want to use the "Prosit\_2020\_intensity\_hcd" for the Intensity prediction model, and the "Prosit\_2019\_irt" for the iRT prediction model. Click next.

Р	ease note: Antivirus software may cancel large uploads - turn it off if you experience upload resets.	
•	Settings Indicate collision energy, the maximum number of missed cleavages, and number of oxidized methionines per peptide.	
•	Upload Files Fasta or CSV with list of peptides	
3	Model Select intensity and IRT model for prediction	
Г	Intensity prediction model	
	<ul> <li>Prosit_2019_intensity_hcd</li> <li>Prosit_2020_intensity_preview</li> <li>Prosit_2020_intensity_cid</li> <li>Prosit_2020_intensity_2021</li> <li>IRT prediction model</li> <li>Prosit_2019_int</li> <li>Prosit_TMT_int_2021</li> </ul>	
4	Isobaric Label If TMT model is selected Task ID Check if everything is correct and submit the task	]

For Task ID Output format, select "NIST .MSP Text Format of individual spectra (Skyline and MSPepSerach compatible).



Submit the task. It is helpful to record the task number or bookmark the page shown after you submit the task to come back to once the job is complete.

	ed. One is for fragment intensities prediction (Prosit, TMT_intensity_2021) and the other is for IRT prediction (Prosit_TMT_int_2021). The ds but you need to add fragmentation column to the input. We assume all the sequences are fully labeled and you don't need to add the time in the interval of the input.
Prosit 🚻	PREDICT LIBRARIES FAQ STATUS 🔿 🤪
Prosit offers high quality IMS2 predicted spectra for any organism and prot synthetic dataset. When using Prosit is helpful for your research, please cl	tease as well as IRT prediction. Prosit is part of the ProteomeTools ( <u>www.proteometools.org(</u> ) project and was trained on the project's high quality the "Gessulat, Schmidt et al. 2019" <u>DOI 10.1038/s41592-019-0426-7</u> .
Task 63E59E6B03CB54EC9E46D2AA185FCB6F	
This task is in progress. Tasks may take several hours for full proteomes de completion. Resubmitting tasks will not lead to faster results.	epending on system load. Please note down your Task ID or save this URL to check back later. You can download the results here upon

Once the task is done, download the .MSP file.

We now offer two new Posit TMT models that will soon be published. One is for fragment intensities prediction (Prosit_ intensity model works with both CID and HCD fragmentation methods but you need to add fragmentation column to the modification explicitly in your input files.	
Prosit	PREDICT LIBRARIES FAQ STATUS 🗘 ?
Prosit offers high quality MS2 predicted spectra for any organism and protease as well as IRT prediction. Prosit is part of the Pr synthetic dataset. When using Prosit is helpful for your research, please cite "Gessulat, Schmidt et al. 2019" <u>DOI 10.1038/s415</u>	
Task 63E59E6B03CB54EC9E46D2AA185FCB6F	
Your files are ready.	
DOWNLOAD	

D. Convert the Prosit output (.MSP/NIST) to a Library (.DLIB) using Scribe. In Scribe, go to "Convert," then navigate to "Convert SPTXT/MSP to Library." This window will pop up.

Deladit	
Convert NIST SPTXT/MSP to Library	×
Parameters:	
SPTXT/MSP:Please select file	Edit
FASTA: Please select file	Edit
OK Cancel	

Default

Upload the downloaded .MSP file from Prosit, and the FASTA file used to generate the .MSP.

After Uploading Files	
Convert NIST SPTXT/MSP to Library	×
Parameters: SPTXT/MSP:myPrositLib.msp	Edit
FASTA:uniprot_human-reference_reviewed_2022mar02.fasta	Edit
OK Cancel	

Click okay. The dialog box will indicate the .MSP file is being converted to a .DLIB. Once it is complete, the dialog box will tell you.

U ACC	520 Counts	^
7 Acc	557 Counts	
8 Acc	394 Counts	
9 Acc	362 Counts	
10 Acc	254 Counts	
11 Acc	212 Counts	
12 Acc	144 Counts	
13 Acc	90 Counts	
14 Acc	75 Counts	
15 Acc	86 Counts	
16 Acc	47 Counts	
17 Acc	35 Counts	
18 Acc	33 Counts	
19 Acc	24 Counts	
20 Acc	214 Counts	
Writing libra	ry file myPrositLib.dlib	
9%		
19%		
29%		
39%		
49%		
59%		
69%		
79%		
89%		
99%		
100%		
	ading myPrositLib.msp	~

The resulting .DLIB should be analogous to

uniprot\_human-reference\_reviewed\_2022mar02.prosit\_input.trypsin\_nce29\_hcd2020.dlib.

### 4. SCRIBE SEARCH OPTIONS

Descriptions of the search options available in Scribe.

#### Here is a screenshot of the Scribe default search options:

EncyclopeDIA: Peptide Searching for DIA File View Convert Data Help	
Thesaurus 🚳 Walnut Scribe	
Scribe: Spectrum-Centric Library Searchin for Data-Dependent Acquisition (DDA) MS Data Scribe extracts peptide fragmentation spectra from MZML files, matches to spectra in libraries, and calculates various scoring features. These feat are interpreted by Percolator to identify peptides.	/MS
Parameters:	
Library:Please select file	Edit
Background:Please select file	Edit
Target/Decoy Approach: Normal Target/Decoy	~
Enzyme: Trypsin	$\sim$
Fragmentation: CID/HCD (B/Y)	~
Precursor Mass Tolerance: 10.0 PPM	~
Fragment Mass Tolerance: 10.0 PPM	~
Library Mass Tolerance: 10.0 PPM	~
Percolator Version: v3-01	~
Number of Quantitative lons:	5 🜩
Minimum Number of Quantitative Ions:	3 🗘
Number of Cores:	17 🔹
Additonal Command Line Options:	

A. Scribe has several options for searching files. Before you can start loading data, you need to specify both a .DLIB or .ELIB library to search as well as a background FASTA. These will be shaded red until they are properly specified. Libraries can be either in the .ELIB (chromatogram library) or .DLIB (spectrum library) format.

B. Scribe has several other search settings. As a general rule, we recommend using the default search parameters first. Other settings are defined below:

**Target/Decoy Approach:** In some circumstances, it may be necessary to add additional decoys to improve statistical analysis. However, as a general rule, this should be left at "Normal Target/Decoy".

Enzyme: Several common digestion enzymes are supported.

Fragmentation: In general, we recommend using CID/HCD (B/Y) fragmentation for most CID or HCD experiments. However, if your library is particularly large or messy you may get improved results with "HCD (y- only)".

**Precursor/Fragment/Library Mass Tolerance:** Tolerances can be specified in PPM, AMU, or resolution.

Percolator Version: Percolator 3.1 is recommended for most experiments.

**Number of Cores:** This is the number of CPU cores you allow EncyclopeDIA to use. The maximum value you should set this to is one less than the number of cores your computer has. You need to leave at least one core for background processes.

C. Additional command line options can be specified in the command line options box. For example:

Additonal Command Line Options: -variable M=15.9949	
---	--

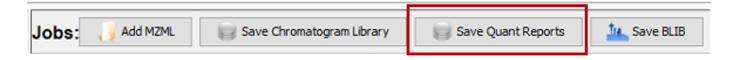
D. Use the "Add MZML" button to add RAW files for library searching. These will be automatically placed on the queue and executed in order using the current settings. If an MZML has been previously analyzed, Scribe will remember where it left off and try to not process it a second time.

Jobs: 🤳 Add MZML	Save Chromatogram Library	Save Quant Reports	kave BLIB
File	Progres	55	

Here is a screenshot of the Scribe search queue:

File	Progress
Read 2022_09_01_HeLa_500ng_16mzst_DDA_NCE_22.mzML	Converting files
Read 2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27.mzML	
Read 2022_09_01_HeLa_500ng_16mzst_DDA_NCE_32.mzML	
Read 2022_09_01_HeLa_500ng_16mzst_DDA_NCE_37.mzML	
Read 2022_09_01_HeLa_500ng_16mzst_DDA_NCE_42.mzML	

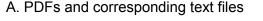
E. Search results can also be saved for downstream quantitative assessment using the "Save Quant Reports" button. RAW files between these experiments are expected to contain shared peptides so retention-time alignment is performed and match-between-runs quantification is calculated.

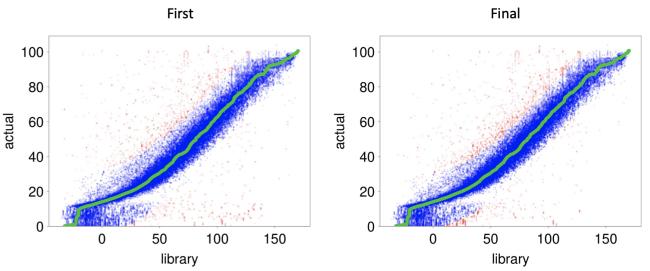


### **5. INTERPRETATION OF SCRIBE OUTPUTS**

Descriptions of the output files given from Scribe.

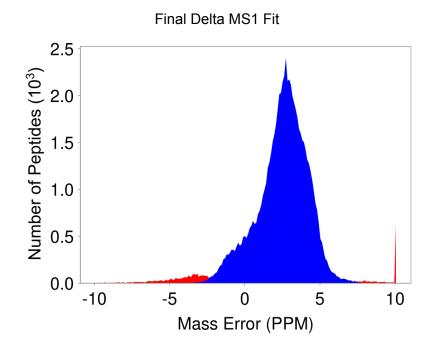
Scribe outputs figures as PDFs displaying fits for retention time (RT), MS1, MS2, delta MS1, and delta MS2. These files can be used to partially assess Scribe's performance in matching spectra to peptides. Additional metrics can be found using the Features viewer, described in part 6 of this manual.



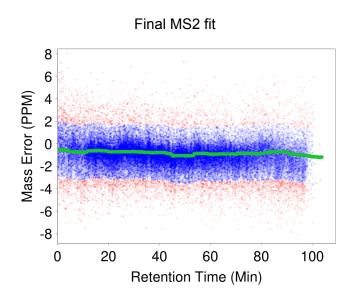


RT plots show the Prosit-predicted RT to the actual RT detected. The green line shows the trace of the kernel density estimation (KDE) RT fit. The blue dots indicate peptides that are matched within 3 standard deviations of the green fit. The red dots represent peptides that are rejected by the alignment, which are outside of 3 standard deviations of the median fit.

In the first fit, peptide FDRs are calculated using Percolator, then fed back into the RT algorithm for further refinement. The refined, or final retention time fit is represented in the second plot. Peptides that are sent through the second pass of the Percolator are a low fraction of the peptides that were fed into the first pass. As a result, the first and final traces only look subtly different. For the remaining plots, only the final fit will be shown.



Similarly to the RT fit, the delta MS1 fit displays the mass error of peptide detections. The error is calculated by comparing the expected mass of a given precursor to the detected mass. The blue portion of the peak represents peptides found within 3 standard deviations of this mass error fit, while the red represents peptides outside of this range.



Precursor m/z and fragment m/z have outputs that plot RT vs precursor mass error. We expect the acceptable error to be within the mass error that is selected during the peptide search. For this search, a mass tolerance of 10 PPM was selected for fragment ions, which gives the green line, or the median mass error fit. The blue dots represent peptides identified within 3 standard deviations of the median mass error, while the red dots are peptides that fall outside of this range. Although the MS1 and delta MS2 fit are not shown, both plots can be interpreted using the same principles described here.

D. Other files generated during sample searches and quantitation

#### Scribe.txt, decoy.txt, and features.txt

The ending in ".scribe.txt" contains 6 columns. The first column is the PMSid or sample name concatenated to intensity, peptide sequence, and charge. The next columns contain the scribe score, q-value, and posterior error probability. The Scribe score is a sum of squares error calculated from the actual intensity and the predicted intensity. This value reports the difference between the observed intensity and the predicted intensity. The last two columns contain peptides with protein matches. You can use this text file to asses the peptide spectral match for targets of interest

Files ending in .scribe.decoy.txt contain analogous scoring and information for the decoys generated for FDR estimation.

Features.txt files contain the complete 17 scores the Scribe algorithm uses to match peptide spectra to their target proteins using Prosit-predicted libraries.

#### Protein and Peptide.txt

After clicking "Save Quant Reports," a quantitative output is formatted in easy-to-use .txt files at both the protein and peptide levels. The protein and peptide reports use match-between-runs for quantification and therefore are only generated if you are searching more than one .mzML file. Quant report files are formatted as expression matrices that can be analyzed by hand in a program such as excel, or reformatted for use in a package using software like R or Python. Samples are in columns, while peptides or proteins are in rows. The number of peptides and the number of fragments are included as an additional column.

#### Diffacto inputs

Scribe also produces easy-to-outputs to use in Python with the differential expression tool Diffacto. This software, developed by Lukas Käll's research group, is capable of processing and normalizing data before performing a differential expression analysis using bayesian modeling to select only peptides that behave similarly to each other. Scribe writes the necessary .csv file for sample abundances, and the .lst file containing annotations. The .lst output contains sample annotations. The default lists all samples as separate sample groups. You will need to manually enter sample groupings before using Diffacto, if you wish for samples to be grouped by condition. To use Diffacto, consult its <u>Github</u> page.

Diffacto citation: Zhang, Bo et al. Covariation of Peptide Abundances Accurately Reflects Protein Concentration Differences Molecular & Cellular Proteomics, Volume 16, Issue 5, 936 - 948. <u>https://doi.org/10.1074/mcp.0117.067728</u>.

#### F. Library, .dia, and model files

When Scribe performs quantitation on individual files, or match-between-run quantitation of multiple files, an ELIB file is generated as a part of the output. ELIB files contain the same information present in DLIBs, with retention time information added. The ELIB output can be uploaded to Skyline or EncyclopeDIA to visualize peptides, or opened in a SQLite DB Browser to extract specific information using a query. Refer to part 6 for more information on how to visualize ELIBs and DLIBs.

Scribe writes .dia files as it is processing samples against predicted libraries. These files are read by EncyclopeDIA when visualizing .mzML files, and can also be read in Skyline.

MODEL files are necessary for visualizing data in EncyclopeDIA, and should not be deleted or moved.

### 6. VISUALIZING DATA WITH BUILT-IN FEATURES

Three ways to access data processed by Scribe,

Using the outputs you obtain from Scribe or EncyclopeDIA, you can visualize the data using three methods; through an SQLite database browser, using Scribe, or looking through the data manually within excel. This is not an all-inclusive list of ways to analyze or visualize the data.

#### A. Accessing the data through SQLite

ELIB and DLIB files are formatted as SQLite databases, allowing users to access the data through an SQLite viewer, such as <u>DB Browser for SQLite</u>. Below is a map of how DLIB databases are structured. SQL databases contain tables, which are represented as boxes in the map, the names of these tables are "entries," "metadata," "peptidequants," "peptidescores," "peptidetoprotein," and "proteinscores."Each box contains the column names within the tables and the type of data each

column contains. For example, "PeptideCharge" is formatted as "int," indicating the peptide charges are integers in the data table.

# Scribe .DLIB Database (DDA Library)

PrecursorCharge	int	⊢	PrecursorCharge	int	- P	recursorCharge	int	
PeptideModSeq	string	⊢	PeptideModSeq	string	P	eptideModSeq	string	
PeptideSeq	string		PeptideSeq	string	P	eptideSeq	string	
SourceFile	string		SourceFile	string	S	ourceFile	string	
PrecursorMz	double	1	RTInSecondsCenter	double	Q	Value	double	
Copies	int		RTInSecondsStart	double	P	osteriorErrorProbability	double	
RTInSeconds	double		RTInSecondsStop	double	ls	Decoy	boolean	
Score	double		TotalIntensity	double				
MassEncodedLength	int		NumberOfQuantIons	int				
MassArray	blob		QuantionMassLength	int		peptidetopro	tein	
IntensityEncodedLength	int		QuantionMassArray	blob		PeptideSeq	string	H
IntensityArray	blob		QuantIonIntensityLength	int		isDecoy	Boolean	
CorrelationEncodedLength	int		QuantIonIntensityArray	blob		ProteinAccession	string	
CorrelationArray	blob		BestFragmentCorrelation	double		metoincorre		
RTInSecondsStart	double		BestFragmentDeltaMassPPM	double		proteinscores		
RTInSecondsStop	double		MedianChromatogramEncodedLength	int		ProteinGroup ProteinAccession	int	
MedianChromatogramEncodedLength	int		MedianChromatogramArray	blob		SourceFile	string	L
MedianChromatogramArray	blob		MedianChromatogramRTEncodedLength	int		QValue	string double	
QuantifiedIonsArray	blob		MedianChromatogramRTArray	blob		MinimumPeptidePEP	double	L
metadata Key string			IdentifiedTICRatio	double		IsDecoy	boolean	

To open the browser, double-click the icon on the desktop, or find the program in the location it was downloaded.



Once the database browser is open, go to the File > Open Database menu option, then find the ELIB in the location where it is stored.

New Database	Ctrl+N
New In-Memory Database	
Open Database	Ctrl+O
Open Database Read Only	Ctrl+Shift+O
Attach Database	
Close Database	Ctrl+F4
Write Changes	Ctrl+S
Revert Changes	
Import	1
Export	,
Open Project	
Save Project	
Save Project As	
Save All	Ctrl+Shift+S
1D:\Beller\ysis compare dia\Lysis Buffer Compare_Chromatogram Library_newSeachEncylopeDIA_Nb.elib	Ctrl+1
2 D:\Beller\ysis compare dia\sqlite_comparison_lysis_buffers.db	Ctrl+2
Exit	Ctrl+O

On the first tab, you can visualize the database structure. For example, the first file in the database is the "entries" table, as you can see below. This matches the information seen in the map above. You can see there are 9 tables in total. Below the tables, are indices, which are lookups for finding information within tables.

tabase Structure Browse Data Greate Table Greate Table Tables (9) Tables (9) FrecursorMz PrecursorMz PreptoleSeq PeptoleSeq PeptoleSeq PeptoleSeq RTInSecondS Seq RTInSecondS Seq RTInSecondSstop RtinsSecondSstop	k Print	Execute SQL Type double int string string string double double int blob	Schema CREATE TABLE entries ( PrecursorMz double not null, PrecursorCharge int not null, PeptideModSeq string not null, PeptideSeq string not null, Copies int not null, RTInSeconds dou "PrecursorMcT double INOT NULL "PeptideModSeq" string NOT NULL "PeptideModSeq" string NOT NULL "Copies" int NOT NULL "Copies" int NOT NULL "Score" double NOT NULL "Score" double NOT NULL
Tobles (9)     Tobles (9)     Tobles (9)     PrecursorMz	gth .ength	double int string string int double double int	CREATE TABLE entries (PrecursorMz double not null, PrecursorCharge int not null, PeptideModSeq string not null, PeptideSeq string not null, Copies int not null, RTInSeconds dou "PrecursorCharge" int NOT NULL "PeptideSeq" string NOT NULL "PeptideSeq" string NOT NULL "Copies" int NOT NULL "RTInSeconds" double NOT NULL "Score" double NOT NULL
Tobes (9)           * entries           * PrecursorMz           * RTInSeconds           * CorrelationArray           * OtherChromatop           * MedianChromatop           * Predideozlazators           * peptideozlazators           * petideozlazators           * petideozlazators           * petideozlazators           * petideozlazators           * petideozlazators           * petideozlazators <td< th=""><th>.ength</th><th>double int string string int double double int</th><th>CREATE TABLE entries (PrecursorMz double not null, PrecursorCharge int not null, PeptideModSeq string not null, PeptideSeq string not null, Copies int not null, RTInSeconds dou "PrecursorCharge" int NOT NULL "PeptideSeq" string NOT NULL "PeptideSeq" string NOT NULL "Copies" int NOT NULL "RTInSeconds" double NOT NULL "Score" double NOT NULL</th></td<>	.ength	double int string string int double double int	CREATE TABLE entries (PrecursorMz double not null, PrecursorCharge int not null, PeptideModSeq string not null, PeptideSeq string not null, Copies int not null, RTInSeconds dou "PrecursorCharge" int NOT NULL "PeptideSeq" string NOT NULL "PeptideSeq" string NOT NULL "Copies" int NOT NULL "RTInSeconds" double NOT NULL "Score" double NOT NULL
Tobes (9)           * entries           * PrecursorMz           * RTInSeconds           * CorrelationArray           * OtherChromatop           * MedianChromatop           * Predideozlazators           * peptideozlazators           * petideozlazators           * petideozlazators           * petideozlazators           * petideozlazators           * petideozlazators           * petideozlazators <td< td=""><td>.ength</td><td>double int string string int double double int</td><td>CREATE TABLE entries (PrecursorMz double not null, PrecursorCharge int not null, PeptideModSeq string not null, PeptideSeq string not null, Copies int not null, RTInSeconds dou "PrecursorCharge" int NOT NULL "PeptideSeq" string NOT NULL "PeptideSeq" string NOT NULL "Copies" int NOT NULL "RTInSeconds" double NOT NULL "Score" double NOT NULL</td></td<>	.ength	double int string string int double double int	CREATE TABLE entries (PrecursorMz double not null, PrecursorCharge int not null, PeptideModSeq string not null, PeptideSeq string not null, Copies int not null, RTInSeconds dou "PrecursorCharge" int NOT NULL "PeptideSeq" string NOT NULL "PeptideSeq" string NOT NULL "Copies" int NOT NULL "RTInSeconds" double NOT NULL "Score" double NOT NULL
entries     entries     entries     Precursor/targe     PeptideModSeq     PeptideModSeq     PeptideModSeq     PeptideModSeq     PeptideModSeq     PeptideModSeq     PeptideModSeq     PeptideModSeq     CorrelationArray     GorelationArray     GorelatioArray     GorelatioArray     Go	.ength	int string string int double double int	"PrecursorMz" double NOT NULL "PrecursorMzge" int NOT NULL "PeptideMode" string NOT NULL "PeptideSeq" string NOT NULL "Copies" int NOT NULL "Strins-cond' double NOT NULL "Score" double NOT NULL
PrecursorMt     PrecursorMt     PrecursorMt     PredeModSeq     PeptdeModSeq     PeptdeModSeq     Corels     Sore     MassEncodedLeng     MassEncodedLeng     MassEncodedLeng     MassEncodedLeng     CorrelationEncode     CorrelationEncode     CorrelationEncode     CorrelationEncode     SoreFile     Fragmentquents     Peptdelocalizations	.ength	int string string int double double int	"PrecursorMz" double NOT NULL "PrecursorMzge" int NOT NULL "PeptideMode" string NOT NULL "PeptideSeq" string NOT NULL "Copies" int NOT NULL "Strins-cond' double NOT NULL "Score" double NOT NULL
PrecursorCharge     PepddeModSeq     PepddeModSeq     PepddeModSeq     Copies     KTinSeconds     Score     MassEncodedLeng     MassEncodedLeng     MassInreave     IntensityEncodedL     ThensityArray     CorrelationEncode     CorrelationEncode     CorrelationEncode     CorrelationEncode     ScoreEnl     MedianChromatog     MedianChromatog     SourceFile     SecureFile     S	.ength	int string string int double double int	*PrecursorCharge* Int NOT NULL *Peptide&qf string NOT NULL *Peptide&qf string NOT NULL *Copies* Int NOT NULL *CTIRSeconds* double NOT NULL *RTIRSeconds* double NOT NULL
Peptidesque     Peptidesque     Peptidesque     Copies     TTinSeconds     TrinSeconds     TrinSecondsrray     OrrelationArray     Orrelation	.ength	string string int double double int	"PeptideModSeq" string NOT NULL "PeptideSeq" string NOT NULL "Copies" int NOT NULL "KTINSeconds" double NOT NULL "Score" double NOT NULL
PepideSeq     Copies     Copies     Kassinorded.eng     Massinorded.eng     Massinorded.eng     Massinorded.eng     Massinorded.eng     Massinorded.eng     Massinorded.eng     TrinisecondStat     CorrelationArray     QuantifiedionsKrr     ThisecondStat     MedianChromatog     MedianChromatog     MedianChromatog     MedianChromatog     SourceFile     Frgmentquants     Pepidelocalizations     Pepidelocalitations     Pepidelocalizations     Pepidelocalizat	.ength	string int double double int	"PeptideSeq" string NOT NULL "Copies" Int NOT NULL "KTINSECORA' double NOT NULL "Score" double NOT NULL
Copies C	.ength	int double double int	"Copies" Int NOT NULL "RTInSeconds" double NOT NULL "Score" double NOT NULL
RTInSeconds     Score     MassEncodedLeng     MassEncodedLeng     MassEncodedLeng     MassEncodedLeng     MassEncodedLeng     CorrelationEncode     CorrelationEncode     CorrelationEncode     Score     RTInSecondSStor     MedianChromatog     MedianChromatog     SourceFie     Fragmentquants     peptidelocalizations     potentions     Sourcefie     Sourcefie     Sourcefie     peptidelocalizations     peptidelocalizations     potentions     potention	.ength	double double int	"RTInSeconds" double NOT NULL "Score" double NOT NULL
Score MassEncodedLeng MassFray MassEncodedLeng MassArray IntensityEncodedL TensityAray CorrelationArray QuantificiansArr ThiSecondStap MedianChromatog MedianChromatog MedianChromatog SurceFile FiltscondStap MedianChromatog Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocres Peptidelocres Pertentontmes Intens(16) Key,Metadata_index	.ength	double int	"Score" double NOT NULL
MassFrocdedum MassArcodedum IntensityArcodedum TrensityArcodedum CorrelationArcodedum CorrelationArcom CorrelationArcom MedianChromatog MedianChromatog SurceFile Peptidekoalizations peptidekoalizations peptidekoalizations peptidekoalizations peptidekoalizations peptidekoalizations peptidekoalizations peptidekoalizations peptidekoratism peptidekoratism peptidekoratism peptidekoratism peptidekoratism peptidekoratism proteinscores % Key,Metadata_index	.ength	int	
MassArray intensityEncoded. IntensityEncoded. IntensityEncoded. CorrelationArray QuantifiedIonsArr RTInSecondSStar MedianChromatog MedianChromatog MedianChromatog MedianChromatog Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocrotes Peptidelocrotes Peptidelocrotes Peptidelocrotes Peptidelocrotes Peptidelocrotes Peptidelocrotes Peptidelocrotes Peptidelocrotes Peptidelocrotes Pertentiontimes Mates (16) Netedata_index	.ength		"MassEncodedLength" int NOT NULL
IntensityEncodedL IntensityArray OrrelationEncode CorrelationEncode CorrelationArray QuantifiedIonsArra RTInSecondStata RTInSecondStata RtInSecondStata MedianChromatog GuarceFile GuarchCorelata Deptidelocalizations Dep		5100	Massenzoueuengun in HOL HOLL
Intensity-Array     CorrelationEncode     CorrelationEncode     CorrelationEncode     CorrelationArray     QuantifiedionsKrr     ThisecondStart     ThisecondStart     Medianc/tromatog     Medianc/tromatog     SourceFile     Fragmentquants     peptidelocalizations     potelinscores     fraceres     Matadata_index		int	"Intensify Good Hor HOLE
CorrelationEncode CorrelationEncode The CorrelationArray Quantified/JoneArray ThiSecondSStap MedianChromatog SurceFile MedianChromatog SurceFile metadata peptidelocalizations peptidelocalizations peptidelocres peptidelocres peptidelocres peptidelocres peptidelocres pertentiontimes retentiontimes Mates (16) % Key, Metadata_index		blob	Intensive root NUL
Correlation/ray Quantifiediona/rr RTInSecond/Stlart RTInSecond/Stlart RTInSecond/Stlart MedianChromatog SurceFie Fragmentquants Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Peptidelocalizations Profelioscores Relations Not Key, Metadata_Index	di oneth	int	anderskylan by older het het belander int
QuantifiedIonsArr RTInSecondStatu RTInSecondStatu RTInSecondStatu MedianChromatog MedianChromatog SourceFile SourceFile SourceFile peptidelocalizations peptidelocalizations peptidelocalizations peptidelocres peptidelocres peptidelocres peptidelocres peptidelocres peptidelocres petidelocres petidelocres petidelocres Mices (16) Key, Metadata_index	eutengui	blob	CorrelationEncodedLength Int
RTInSecondsStart     RTInSecondsStart     RTInSecondsStart     MedianChromatog     SourceFile     peptidelocalizations     potentions     potentions     potentions     potentions     potentions     potentions     potentions     potentions	-	blob	Contrelations area blob
RTInSecondsStop     MedianChromatog     MedianChromatog     SourceFile     SurceFile     peptidelocalizations     peptidelocalizations     peptidelocalizations     peptidelocalizations     peptidelocrotein     proteinscores     Indecs (16)     % Key,Metadata_index		double	Quantileanisation of the second start' double
MedianChromatog MedianChromatog SurceFile > fragmentquants > metadata > peptidelocalizations > peptidescores > peptidescores > peptidescores > petidescores > petidescores > retentiontimes > Index (16) > % Key, Metadata_index		double	R Insecondsstor double
MedianChromatog SourceFile SourceFile metadata peptidelocalizations peptidelocalizations peptidescores peptidescores peptidescores functional proteinscores functional functiona			MedianChromatogramBinodedLength* int
SourceFile		blob	MedianChromatogramArray' blob
<ul> <li>Fragmentquants</li> <li>metadata</li> <li>peptidelocalizations</li> <li>peptidelocalizations</li> <li>peptidescores</li> <li>peptidestoretein</li> <li>proteinscores</li> <li>retentiontimes</li> <li>Indices (16)</li> <li>% Key_Metadata_index</li> </ul>	gramArray		MeuanLinomaugramistray iloo
>         metadata           >         peptidelocalizations           >         peptidescores           >         peptidescores           >         peptidescores           >         proteinscores           >         retentiontimes           >         Indices (16)           >         % Key_Metadata_index		string	Sourcernie suing NOT NUCL. CREATE TABLE fragmentquants ( PrecursorCharge int not null, PeptideModSeg string not null, PeptideSeg string not null, SourceFile string not null, IonType string not null, IonType string not null, IonType string not null, SourceFile string not null, SourceFile string not null, SourceFile string not null, IonType string not null, SourceFile string not null, SourceFi
>     peptidelocalizations       >     peptidequants       >     peptidescores       >     peptidetoprotein       >     proteinscores       >     retentiontimes       Indices (16)       >     Key_Metadata_index			CREATE FABLE metadation (Key string on hull) value string not null.) + paudemoused string not null, sourceme string not nu
<ul> <li>peptidequants</li> <li>peptidescores</li> <li>peptidetoprotein</li> <li>proteinscores</li> <li>retentiontimes</li> <li>Indices (16)</li> <li>Key_Metadata_index</li> </ul>			CREATE I ABLE metadada (Key samg not many vane samg not man) CREATE TABLE peptideocalizations (PrecursorCharge int not null). PeptideModSeq string not null, PeptideSeq string not null, SourceFile string not null, LocalizationPeptideModSec
<ul> <li>&gt; peptidescores</li> <li>&gt; peptidetoprotein</li> <li>&gt; proteinscores</li> <li>&gt; retentiontimes</li> <li>&gt; Indices (16)</li> <li>&gt; Key_Metadata_index</li> </ul>			
<ul> <li>peptidetoprotein</li> <li>proteinscores</li> <li>retentiontimes</li> <li>Indices (16)</li> <li>Key_Metadata_index</li> </ul>			CREATE TABLE peptidequants (PrecursorCharge int not null, PeptideModSeq string not null, PeptideSeq string not null, SourceFile string not null, RTInSecondsCenter double not
<ul> <li>proteinscores</li> <li>retentiontimes</li> <li>Indices (16)</li> <li>Key_Metadata_index</li> </ul>			CREATE TABLE peptidescores ( PrecursorCharge int not null, PeptideModSeq string not null, PeptideSeq string not null, SourceFile string not null, QValue double not null, Posteric CREATE TABLE peptidetoprotein (PeptideSeq string not null.isDecov boolean, ProteinAccession string not null)
<ul> <li>iii retentiontimes</li> <li>Indices (16)</li> <li>Key_Metadata_index</li> </ul>			
<ul> <li>Indices (16)</li> <li>Key_Metadata_index</li> </ul>			CREATE TABLE proteinscores ( ProteinGroup int not null, ProteinAccession string not null, SourceFile string not null, QValue double not null, MinimumPeptidePEP double not null, CREATE TABLE retentiontimes (SourceFile string not null, Library float not null, Actual float not null, Predicted float not null, Delta float not null, Probability float not null, Deory bo
> 📎 Key_Metadata_index			CREATE TABLE retentiontimes (Sourcerile string not null, Library noat not null, Actual noat not null, Predicted noat not null, Deita noat not null, Probability noat not null, Decoy bo
			CDEATE RUDDY I/vou Materiales induction instruction (World ACC)
	Channel Co		CREATE INDEX 'Key_Metadata_index' on 'metadata' ('Key' ASC) CREATE INDEX 'PeptideModSeq_PrecursorCharge_SourceFile_Entries_index' on 'entries' ('PeptideModSeq' ASC, 'PrecursorCharge' ASC, 'SourceFile' ASC)
> > PeptideModSeq_Precip > > PeptideModSeq_Precipient Precipient Precipien			
> > PeptideModSeq_Preci > > PeptideModSeq_Preci			CREATE INDEX 'PeptideModSeq_PrecursorCharge_SourceFile_Fragments_index' on 'fragmentquants' ('PeptideModSeq' ASC, 'PrecursorCharge' ASC, 'SourceFile' ASC) CREATE INDEX 'PeptideModSeq_PrecursorCharge_SourceFile_Localizations_index' on 'peptidelocalizations' ('PeptideModSeq' ASC, 'PrecursorCharge' ASC, 'SourceFile' ASC)
· · · · · · · · · · · · · · · · · · ·			CREATE INDEX 'PeptideModSeq_PrecursorCharge_SourceFile_Peptides_Index' on 'peptidequants' ('PeptideModSeq' ASC, 'PrecursorCharge' ASC, 'SourceFile' ASC) CREATE INDEX 'Instituted for Breamscharge SourceFile (Asc) CREATE INDEX 'Instituted for Breamscharge (SourceFile) ASC)
· · · · · · · · · · · · · · · · · · ·		u	CREATE INDEX 'PeptideModSeq_PrecursorCharge_SourceFile_Scores_index' on 'peptidescores' ('PeptideModSeq' ASC, 'PrecursorCharge' ASC, 'SourceFile' ASC)
			CREATE INDEX 'PeptideSeq_Entries_index' on 'entries' ('PeptideSeq' ASC)
> PeptideSeq_Fragmen			CREATE INDEX 'PeptideSeq_Fragments_index' on 'fragmentquants' ('PeptideSeq' ASC)
			CREATE INDEX 'PeptideSeq_Localizations_index' on 'peptidelocalizations' ('PeptideSeq' ASC)
> > PeptideSeq_PeptideT > > PeptideSeq_PeptidESeq_Pe	ions_index		CREATE INDEX 'PeptideSeq_PeptideToProtein_index' on 'peptidetoprotein' ('PeptideSeq' ASC) CREATE INDEX 'PeptideSeq_Peptides_index' on 'peptidequants' ('PeptideSeq' ASC)

If you click on the next tab	, "Browse Data,"	you will be able to	look through tables.
<b>,</b>	,,	<b>J</b> · · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·

ie	w Database 🛛 🙀 Open Da	tabase 🚬 🛱 Write Changes 🔅 Revert Changes	Open Project Cave Project	🛃 Attach Database 🛛 🔀 Close Database		25 ~		
bi	ase Structure Browse Da	ta Edit Pragmas Execute SQL						
[	peptidequants v		Filter in any column					
	entries fragmentquants	PeptideModSeq	PeptideSeq	SourceFile	RTInSecondsCenter	RTInSecondsStart	RTInSecondsStop	TotalInten '
	metadata	ter	Filter	Filter	Filter	Filter	Filter	Filter
	peptidelocalizations	ETLDTIK	EEETLDTIK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	1407.86999511719	1316.27099609375	1415.08325195313	26323200
	peptidequants	DYLEGIR	LVDYLEGIR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	3211.31640625	3208.42700195313	3308.06030273438	2102088
	peptidescores peptidetoprotein	AEDQNSLLK	WAEDQNSLLK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	2271.26171875	2251.97631835937	2299.94995117187	1681471
	proteinscores	[+15.994915]QIQPNREDTETLIK	LMQIQPNREDTETLIK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	2674.57885742188	2668.18481445312	2680.48120117187	115922
l	retentiontimes	TAQAWDGTTDYQVEETSR	QITAQAWDGTTDYQVEETSR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	3260.9365234375	3248.53173828125	3272.85473632812	512221
	2 A	EEEIMK	AEEEIMK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	831.160217285156	824.255432128906	864.167541503906	5281887
	2 E	GLTSIEEVTK	EGLTSIEEVTK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	2389.99462890625	2375.47143554688	2407.390625	3753205
	2 V	/IQAGMFDQK	VIQAGMEDQK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	1939.56713867187	1929.17077636719	1946.10302734375	560567
	2 1	EINGNWISASSINEAR	NEINGNWISASSINEAR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	3394.36767578125	3383.1416015625	3406.775390625	744737
	2 A	AELIANSLATAGDGLIELR	AAELIANSLATAGDGLIELR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	5393.37451171875	5378.44482421875	5406.99072265625	1635833
	2 5	SPAIAAVGPIK	SPAIAAVGPIK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	1904.90075683594	1877.49133300781	1918.33996582031	2543427
	2 1	/IILGDSGVGK	VIILGDSGVGK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	2166.76489257812	2150.4384765625	2187.71362304687	5653970
	2 1	TLPVDFVTADK	ITLPVDFVTADK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	4155.4443359375	4137.9990234375	4180.34228515625	166000220
	2 6	SFNPESYELDK	ESFNPESYELDK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	2959.03198242187	2942.90209960938	2979.39526367187	1456231
	2 F	AESM[+15.994915]LQQADK	RAESMLQQADK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	879.356811523438	874.1005859375	885.236694335937	470069
	3 F	EEHVQSVDIAAFNKI	FEEHVQSVDIAAFNKI	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	3826.56884765625	3814.69653320312	3839.63818359375	777476
	2 0	LPWSC[+57.0214635]SADEVM[+15.994915]R	GLPWSCSADEVMR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	3549.35083007812	3540.51147460937	3555.07397460938	241983
	2 T	C[+57.0214635]AAQLVSYPGK	TCAAQLVSYPGK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	1997.228515625	1979.51184082031	2004.29040527344	1985949
	3 (	QGLETFKPDYSEQK	QGLETFKPDYSEQK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	1923.76342773438	1907.5341796875	1934.22705078125	398382
	2 H	IPDSSVNFAEFSK	HPDSSVNFAEFSK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	2077.16748046875	2060.81665039062	2090.49169921875	2811341
	4 1	FQVAQELSGEDMHQFHR	IFQVAQELSGEDMHQFHR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	3354.83984375	3342.6767578125	3372.05639648437	536543
	2 0	SILAADESTGSIAK	GILAADESTGSIAK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	1972.1494140625	1956.30200195313	1988.77624511719	74386897
	2 1	LEAWEMNEK	ILEAWEMNEK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	2801.77099609375	2787.81274414062	2816.12548828125	1240637
	2 A	EDEVQR	AEDEVQR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	310.843658447266	288.667419433594	315.162322998047	4854728
	2 0	QTPDENDQVVVK	DQTPDENDQVVVK	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	1320.55749511719	1309.63464355469	1332.82543945313	1805525
	2 A	YSSFGGGR	AYSSFGGGR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	981.891723632813	969.392333984375	987.909912109375	3170589
	3 1	TAPTHVPLQYIGPNQPR	ITAPTHVPLQYIGPNQPR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	2881.13427734375	2871.93823242188	2902.48974609375	648206
	2 (	VAQQEAQR	QVAQQEAQR	2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27	353.957092285156	344.921966552734	361.379364013672	2635777
		MLPLNTNIR	OMLPLINTNIR	2022 09 01 HeLa 500ng 16mzst DDA NCE 27	2651.43041992187	2634.31323242188		4155492
		WWDEDDD	INVECTOR	2022 00 01 Mol a 50000 16mmet DDA MCE 27	1550 5202724275	1503 46601004531	1557 675	euseau ,
ŝ	1 - 30 of 34398 🕨 🕅			Go to: 1				

You can use filters to obtain specific data you are interested in, or use regular expressions for a more powerful query. To export a table you are interested in specifically, click "File > Export > Table(s) as CSV."

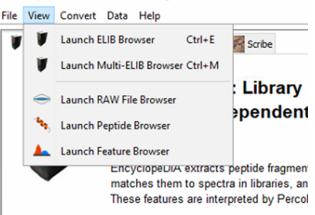
	Database In-Memory Database		Ctrl+N	h Database 🛛 🗙 Close Datab	ase		53	^							
	n Database		Ctrl+O												
Oper	Database Read Only		Ctrl+Shift+O												
Atta	ch Database			SourceFile		RTInSecondsCenter	PTInSocondeStart	RTInSecondsStop	TotalIntensity						
Close	e Database		Ctrl+F4	Sourcerne		Filter	RTInSecondsStart Filter	Filter	Filter Fi						
Write	e Changes		Ctrl+S	a_500ng_16mzst_DDA_NC	E 27	1407.86999511719		1415.08325195313	2632320000.0						
🖗 Reve	ert Changes			a 500ng 16mzst DDA NC		3211.31640625		3308.06030273438	210208896.0						
Impo	rt		•	a 500ng 16mzst DDA NC	-	2271.26171875		2299.94995117187	168147168.0						
Expo	rt		•	Database to SQL file	27	2674.57885742188		2680.48120117187	11592220.0						
Oper	Project			Table(s) as CSV file	27	3260.9365234375		3272.85473632812	51222120.0						
Save	Project			Table(s) to JSON [a_500ng_16mzst_DDA_NC		831.160217285156		864.167541503906	528188704.0						
C Save	Project As			a_500ng_16mzst_DDA_NC		2389.99462890625	2375.47143554688	2407.390625	375320576.0						
Save	All		Ctrl+Shift+S	a 500ng 16mzst DDA NC	-	1939.56713867187		1946.10302734375	56056740.0						
1 D:\	Shannon\Scribe_manual\2022-11-14_searleb_e3b9c5d4\raw\guantrepor	t_DDA_NCE22_trypsin.elib	Ctrl+1	a_500ng_16mzst_DDA_NC	-	3394.36767578125	3383.1416015625	3406.775390625	74473720.0						
2 \\h	ummonproteome2\d\\$hannon\\$cribe_manual\2022-11-14_searleb_e3b9c	5d4\raw\quantreport_DDA_NCE22_trypsin.elb	Ctrl+2	a_500ng_16mzst_DDA_NC		5393.37451171875		5406.99072265625	163583376.0						
3 D:\	Beller\ysis compare dia\Lysis Buffer Compare_Chromatogram Library_ne	wSeachEncylopeDIA_Nb.elib	Ctrl+3	a_500ng_16mzst_DDA_NC	-	1904.90075683594	1877.49133300781	1918.33996582031	254342752.0						
4D:\	Beller \ysis compare dia\sqlite_comparison_lysis_buffers.db		Ctrl+4	a 500ng 16mzst_DDA_NC		2166.76489257812		2187.71362304687	565397056.0						
Exit			Ctrl+Q	a 500ng 16mzst DDA NC	-	4155.4443359375	4137.9990234375	4180.34228515625	16600022016.0						
14	2 ESFNPESYELDK	ESFNPESYELDK	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	2959.03198242187	2942.90209960938	2979.39526367187	145623168.0						
15	2 RAESM[+15.994915]LQQADK	RAESMLQQADK	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	879.356811523438	874.1005859375	885.236694335937	47006944.0						
16	3 FEEHVQSVDIAAFNKI	FEEHVQSVDIAAFNKI	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	3826.56884765625	3814.69653320312	3839.63818359375	77747696.0						
17	2 GLPWSC[+57.0214635]SADEVM[+15.994915]R	GLPWSCSADEVMR	2022_09_01_H	La_500ng_16mzst_DDA_NCE_27		3549.35083007812	3540.51147460937	7 3555.07397460938	24198384.0						
18	2 TC[+57.0214635]AAQLVSYPGK	TCAAQLVSYPGK	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	1997.228515625	1979.51184082031	2004.29040527344	198594944.0						
19	3 QGLETFKPDYSEQK	QGLETFKPDYSEQK	2022_09_01_H	eLa_500ng_16mzst_DDA_NCE_27.		1923.76342773438	1907.5341796875	1934.22705078125	39838288.0						
20	2 HPDSSVNFAEFSK	HPDSSVNFAEFSK	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	2077.16748046875	2060.81665039062	2090.49169921875	281134144.0						
21	4 IFQVAQELSGEDMHQFHR	IFQVAQELSGEDMHQFHR	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	3354.83984375	3342.6767578125	3372.05639648437	53654320.0						
22	2 GILAADESTGSIAK	GILAADESTGSIAK	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	1972.1494140625	1956.30200195313	1988.77624511719	7438689792.0						
23	2 ILEAWEMNEK	ILEAWEMNEK	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	2801.77099609375	2787.81274414062	2816.12548828125	124063760.0						
24	2 AEDEVQR	AEDEVQR	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	310.843658447266	288.667419433594	315.162322998047	485472832.0						
25	2 DQTPDENDQVVVK	DQTPDENDQVVVK	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	1320.55749511719	1309.63464355469	1332.82543945313	180552560.0						
26	2 AYSSFGGGR	AYSSEGGR	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	981.891723632813	969.392333984375	987.909912109375	317058976.0						
27	3 ITAPTHVPLQYIGPNQPR	ITAPTHVPLQYIGPNQPR	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	2881.13427734375	2871.93823242188	2902.48974609375	64820648.0						
28	2 QVAQQEAQR	QVAQQEAQR	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	353.957092285156	344.921966552734	361.379364013672	263577792.0						
29	2 QMLPLNTNIR	QMLPLNTNIR	2022_09_01_H	eLa_500ng_16mzst_DDA_NC	E_27	2651.43041992187	2634.31323242188	2669.82592773437	415549280.0						
20 C	2 01/00000	NAVDEDDD	2022 00 01 U	ol a 500ng 16mzet DDA MC	C 77	1550 5202724275	1502 46601004521	1557 675	50052056 0						

For more information on using DB Browser for SQLite check out the <u>SQLite Browser Wiki Github Page</u>.

B. Looking at data through Scribe's built-in visualization tools.

EncyclopeDIA and Scribe have built-in tools to help you visualize raw data, in .mzML format, libraries, and quant reports. Use the View menu to access the following options for visualizing data.

EncyclopeDIA: Peptide Searching for DIA

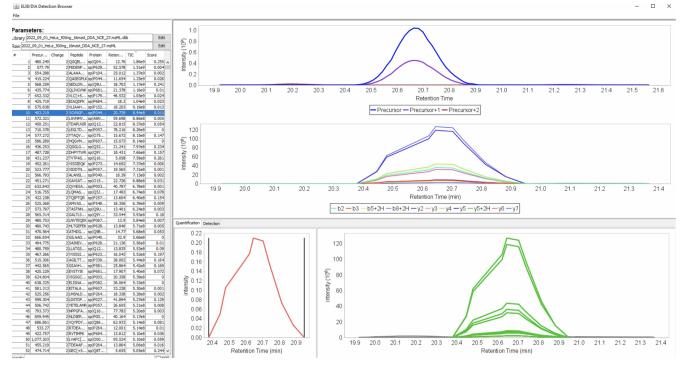


#### **ELIB Browser**

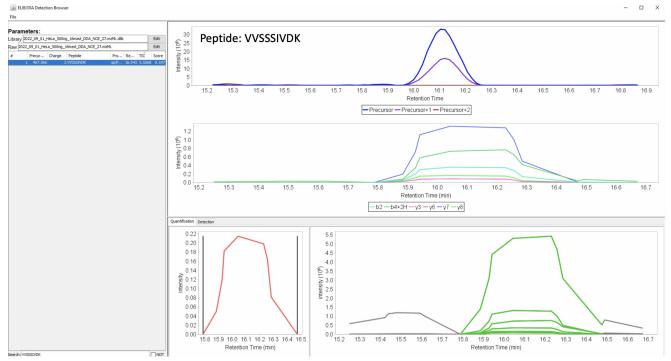
Click on "Launch ELIB Browser" to visualize a library relative to an individual sample. Click "edit" to select the library file you would like. Then do the same for the RAW file, in .mzML format. Note: EncyclopeDIA cannot read files in RAW form.

ELIB/DIA Detection Browser	-
File	
Parameters:	
ibrary Please select fie Edit	
Raw Please select fie Edt	
# Precur Charge Peptide Protein Retent TIC Score	
Precur Charge Peptoe Protein Retent IIC Score	
	1
	ect a peptide!
earch:	

Once the files are loaded, peptides will populate. The screen will pull up this view.



At the bottom, use the search bar to search for specific proteins or peptides. You can see in this example, the precursor spectra for the peptide "VVSSSIVDK" is symmetrical. For the MS2 spectra, the b2, b4+2H, y3, y6 y7 and y8 peptides were detected without interference. Visualization is presented as chromatograms, showing the relationship between ions in multiple spectral acquisitions of the same peptide.

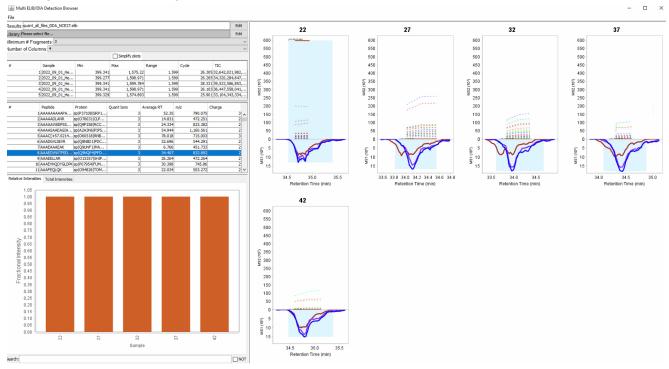


#### Multi-ELIB Browser

Click on "Launch Multi-ELIB Browser" to visualize data across all raw files in a quantitative result. Using this view, you can visualize peptides present in multiple samples.

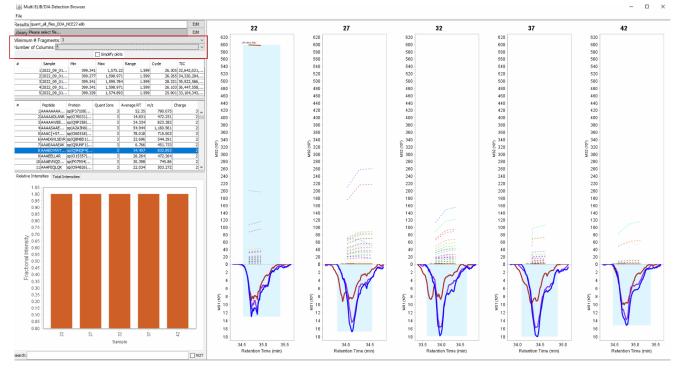
Results:Pleas	e select file					E			
brary:Please	e select file					Ec			
inimum # F	Fragments: 3 olumns: 2						~		
imper of C	olumns: 2		Simplify pl	ts			~		
	Sample	Min	Max	Range	Cycle	TIC			
,	Peptide	Protein	Quant Ions	Average RT	m/z	Charge	=		
	nsities Total Intens						=		
1.05		1023							
1.00									
0.95									
0.90									
0.85	5								
0.80	1								
0.75	5								
0.70	j.								
0.65							Select a peptide!		
							pelect a peptider		
usity of									
Inte									
0.60 Atjuensity 0.50 0.45									
0.40									
0.35									
0.30									
0.25	5								
0.20	1								
0.15	5								
0.10	1								
0.05	; -								
0.00									
			Sar	nple					

Click "edit" to upload the results file you are interested in viewing. You can insert the library file as well, although it is not necessary to visualize samples. Once uploaded, the view should look like this:

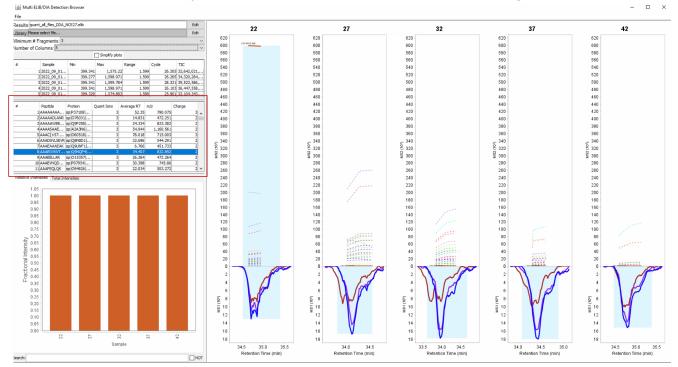


The minimum number of fragments per peptide can be set to limit the peptides that need to be examined. The number of columns can also be changed. The sample order will be in either numerical

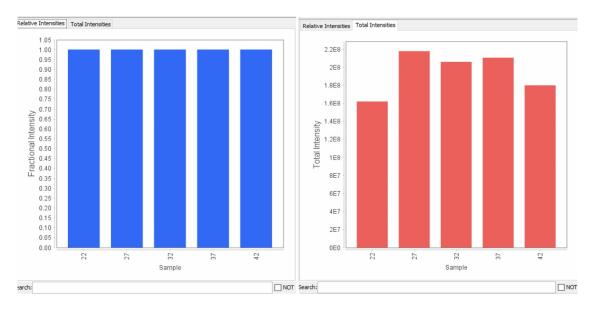
or alphabetical order. Visualization is presented as chromatograms that show the relationship between ions in multiple spectral acquisitions of the same peptide.



Similar to the ELIB browser, you can search for specific peptides, or click through the list of detections:

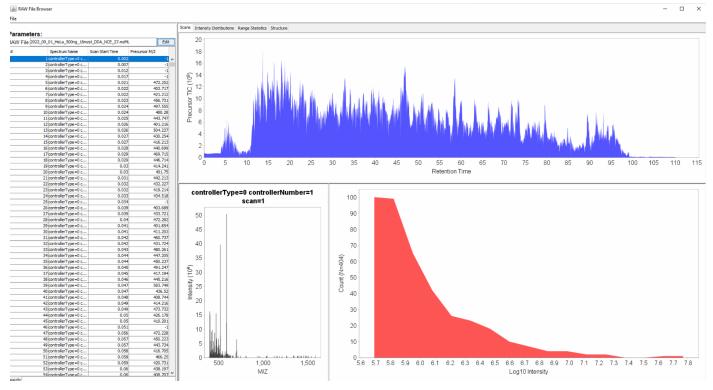


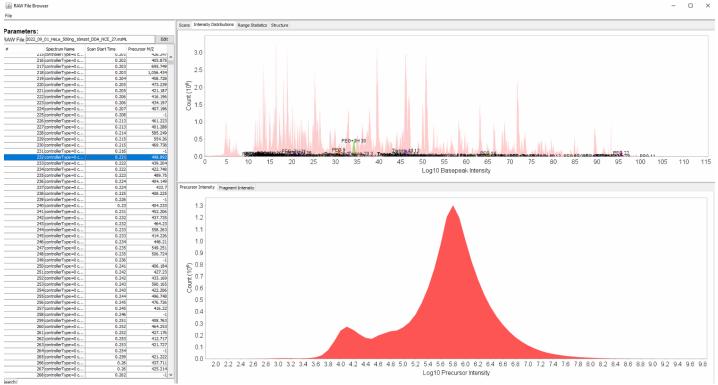
In the lower left panel, there are intensity bar graphs for "Relative Intensities" and "Total Intensities" across samples. Scribe uses precursors for quantitation, therefore the "Relative Intensities" bar graph will always consist of the intensity values for a single precursor across all samples. "Total Intensities" will allow you to compare quantitative intensities across samples. Below are both graphs for the peptide selected in the screenshot above.



#### Raw File Browser

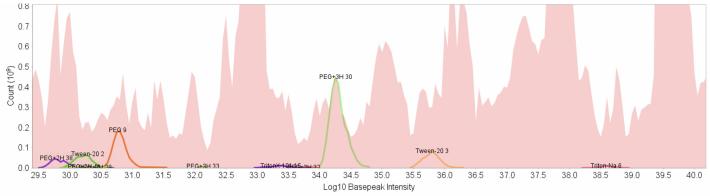
"Launch raw File Browser" will allow you to visualize samples, in mzML format. In this view, you can visualize unsearched data. The data is present in scans, allowing you to visualize the MS1 and MS2 spectra for specific scans throughout your injection





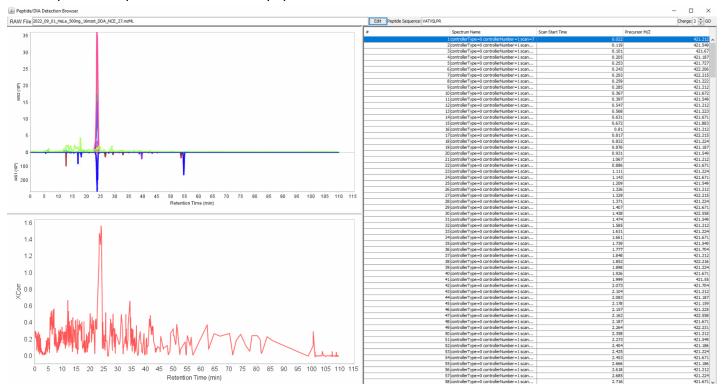
The next tab contains intensity distributions for precursor and fragment ions present in each scan in the bottom pane. In the top pain, you can see the base peak chromatogram trace.

Contaminant peaks from common molecules, such as PEG, triton, tween and polysiloxane are present in the base peak trace. If you look in the screenshot below, you will see PEG +3H 30 base peak present in the chromatogram around 34.5 minutes.

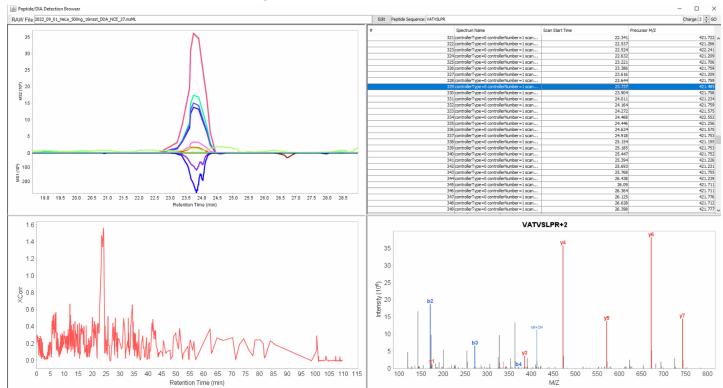


#### Peptide Browser

"Launch Peptide Browser" will allow you to search for a specific peptide in a sample file of interest. When you first open this browser, the peptide "VATVSLPR", a common tryptic peptide, is selected. The top left panel displays the chromatogram trace for this peptide across the entire gradient. The bottom left panel shows the XCorr score across the gradient. You should see the highest score where the most intense peak is for the peptide. On the right side of the screen are the scans throughout the gradient. Visualization is presented as chromatograms that show the relationship between ions in multiple spectral acquisitions of the same peptide.

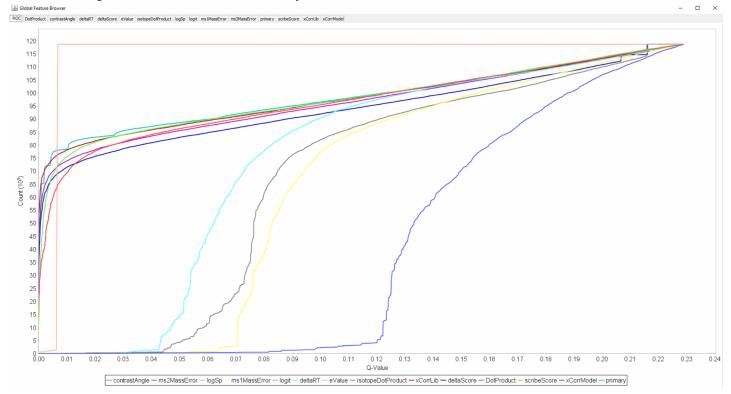


In this next view, the base peak has been zoomed in under the chromatogram panel in the top left. You can see this peak elutes around 23.8 min. A scan acquired at 23.737 min has been selected to pull up the MS2 spectra in the bottom-right corner.

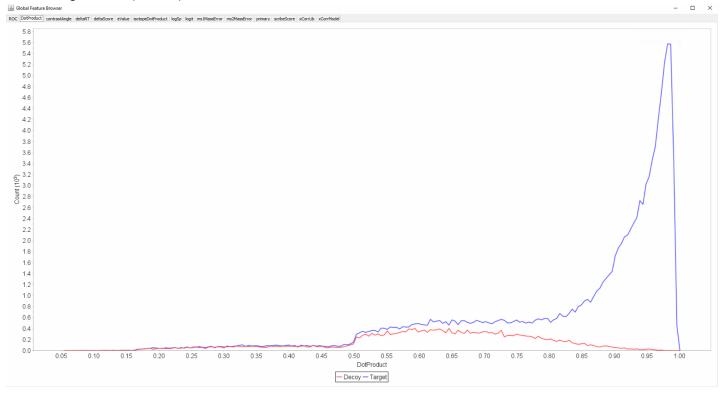


#### Feature Browser

"Launch feature browser" will allow you to visualize the features.txt files output from Scribe or EncyclopeDIA. The tabs of the features browser contain various metrics used in matching predicted spectra to peptides. Below is a screenshot of the ROC plot, which shows the performance of the 13 features available through this browser. On the x-axis is the Q-value, or the chances of a given feature resulting in a false positive rate (FPR). On the y-axis are counts, indicating the number of peptides detected at a given FPR. Features that are lower on the curve have worse performance than features higher on the curve. For example, "ms2MassError," denoted by the dark blue line, performs the worse of the given characteristics and is only useful as a correction factor in this case.



Each of the other tabs contains histograms of the individual scores for target and decoy peptides. For example, below is the dot product plot where a value closer to 1 indicates greater similarity between the actual and predicted spectra. In this plot, you should see the decoy trace (in red) reside lower than the target trace (in blue).



For the remaining features, you should also see the target trace (blue) relative to the decoy trace (red). This indicates the target spectra are performing better than the decoy spectra, which should occur when a peptide match is matched correctly.

#### C. Analyzing text files using Microsoft Excel

After saving a quant report, Scribe will perform a match-between-runs quantitation of all samples you obtained and then write a protein and peptide report. These are formatted as expression datasets of the proteins or peptides detected. This expression dataset allows you to use Open the text file in Excel by right-clicking on the file, open with excel. The following example shows how to process the protein reports for HeLa DDA injections at varying NCE processed against a library optimized against NCE = 27. Within this example, the file was "saved as" .XLS file to ensure that the equations are saved. Below is a screenshot of the file once it has been opened.

lome Insert	Draw P	age La	yout	Formulas	Data	Review	View	Automat	te Add	l-ins	♀ Tell	me							Comm	ents 🕑	🖻 Shar
r ۲ ۲ ۲	Calibri (Body)		v 12	✓ A <sup>^</sup> A	· =	= =	≫ •	ab v	General			•	<b>•</b>		🔠 In	nsert v	ΣčA Z	7. 0.			~
[ <u>-</u>															🎫 D	elete 🗸					
Paste 🗳	B I <u>U</u> √		~ <u>~</u>	• • <u>A</u> •	=	= =	€= →=	÷ •	\$ * %	• •	0. 0,→ ★ 00.		Conditional Forn Formatting as Ta		E Fe	ormat 🗸	✓ ✓ Fil	t & Find & ter Select	Analyze Data	Sensitivi	nty
Possible Data	Loss Some fea	tures r	night be	lost if you sa	ave this wo	orkbook in ti	he text (.tx	t) format. T	o preserve	these	features,	save	it in an Excel file	format.						Sa	ave As.
1 🌲 🛪	√ <i>f</i> x Pr	otein																			
A	B C		D	E	F	G	н	I	J		К	L	M	N	0	Р	Q	R	S	т	
Protein Nun	nPeptide PeptideS	equ 202	2_09_01_2	2022_09_01_2	022_09_01	2022_09_01	2022_09_01	_HeLa_500n	g_16mzst_[	DDA_NC	E_42.mzM	IL									
sp Q5JS54 F	1 YDSIPVS	ISLL 2	2.89E+07	3.29E+07	4.38E+07	4.09E+07	3.20E+07	'													
sp Q04637	50 AALSEEE		2.95E+10	2.94E+10	3.75E+10																
sp P08195 4	24 ADLLLST		1.78E+10	2.00E+10	2.17E+10																
sp Q9UHB9	18 ALLQQQ	PEDI 1	1.25E+09	1.45E+09	1.42E+09	1.51E+09	1.08E+09	)													
sp P07237 F	28 ALAPEYA		2.22E+10	2.36E+10	2.65E+10																
sp Q6NUM9	10 ATVQSVI		5.67E+08	7.21E+08	8.01E+08		5.43E+08														
sp Q13131 .	6 FEC[+57.		3.00E+08	3.75E+08	3.94E+08		3.00E+08														
sp Q09666	271 ADIDISG		1.02E+11	1.08E+11	1.21E+11		9.81E+10														
sp Q13151	6 EDIYSGG		1.70E+09	1.76E+09	1.97E+09		1.66E+09														
sp P31150 C	19 DWNVDL		4.19E+09	4.82E+09	5.60E+09		4.26E+09														
sp Q01813	26 AMEWIT		7.65E+09	6.37E+09	8.18E+09		6.67E+09														
sp Q9H3P7	10 DC[+57.0		3.55E+08	4.17E+08	5.34E+08		3.17E+08														
sp 015213	9 ALAEVD		7.21E+08	8.28E+08	9.90E+08		7.67E+08														
sp Q8N9T8	13 AQEEAD		1.59E+09	1.54E+09	1.60E+09		1.40E+09														
sp P20290 E	9 APLATGE		2.61E+09	3.90E+09	3.18E+09		2.38E+09														
sp Q9NWV8	7 AVGAQA		1.59E+09	1.75E+09	2.13E+09		1.72E+09														
sp P49419 /	27 DLPLAQO		5.78E+09	6.55E+09	6.92E+09		5.63E+09														
sp Q96E11	10 DTVSEDT		2.05E+09	2.51E+09	2.66E+09		1.99E+09														
sp Q13813	137 ADVVES		3.06E+10	3.15E+10	3.84E+10		3.17E+10														
sp P62873 0	9 AC[+57.0		1.44E+09	1.59E+09	1.52E+09		1.40E+09														
sp P42765 1	19 AANDAG		5.54E+09	6.90E+09	7.28E+09		6.69E+09														
sp Q92616	96 AASQSTO		1.83E+10	1.88E+10	2.11E+10																
sp Q99832	30 ALEIIPR;4		1.46E+10	1.54E+10	1.74E+10																
sp Q9BWM7	15 AGVVTP		2.89E+09	3.30E+09	3.65E+09		2.71E+09														
sp Q14980	82 AADALEE		1.30E+10	1.37E+10	1.72E+10		1.42E+10														
sp Q9Y263	28 AINC[+57		3.28E+09	3.25E+09	3.88E+09		3.18E+09														
sp Q8WY22	3 SSPSGPS		1.24E+09	8.03E+08	1.68E+09		1.39E+09														
sp P60891 F	4 ENISEWE		7.40E+08	7.61E+08	8.84E+08		7.82E+08														
sp 075489	10 AANWYE		1.30E+09	1.50E+09	1.85E+09		1.19E+09														
sp Q8NF37	6 AAPASSA		1.29E+09	1.47E+09	1.52E+09		1.33E+09														
sp P56192 5	31 AEVLIST		1.00E+10	1.26E+10	1.45E+10												_				
sp P30084 E	18 AFAAGA t all files DDA		8.32E+09	8.84E+09 +	9.35E+09	9.24E+09	8.07E+09	,													

Add column annotations to describe the sample names/conditions for each injection. In this example, NCE was added as an annotation row.

• • A	utoSave 💽 🏠		8 / * ·	~ . •			quant_a	I_IIIC3_00	A_NCE27.el		manzeu.	proteins		_						Q
	Calibri (Body)	× 12	~ A^ A	× =	= =	* ~	ab ce Wrap	Text 🗸	General		•	<b>•</b>		7.	Insert ∨ Insert ∨ Delete ∨	Σ •	Zv · C	) - [		
iste 🎸	B I <u>U</u> ∽	⊞ •   <	× <u>A</u> ×	≡	≡ ≡	<del>•</del> = <del>•</del> =	🚊 Merge	& Center 🗸	\$ ~ %		€0.00 .00 →0	Conditional Formatting		cell vies	Eormat ∨	-	Sort & Fine Filter Sel	d& A		Sensitiv
	× ✓ fx NCE											Tormatting	us fubic of	yies		•	Tinter Oct		butu	
A	B C	D	E	F	G	н	1	1	к	L	м	N	0	Р	Q	R	s	т	U	J
	umPeptide PeptideSequ						HeLa 500n	z 16mzst DD							4			· · ·		_
ICE		22	27	32	37	42														
p Q5JS54 F	1 YDSIPVSTSLL	2.89E+07	3.29E+07	4.38E+07	4.09E+07	3.20E+07														
p Q04637	50 AALSEEELEK	2.95E+10	2.94E+10	3.75E+10	3.31E+10	3.36E+10														
p P08195 4	24 ADLLLSTQPG	1.78E+10	2.00E+10	2.17E+10	2.10E+10	1.82E+10														
p Q9UHB9	18 ALLQQQPED	1.25E+09	1.45E+09	1.42E+09	1.51E+09	1.08E+09														
p P07237 F	28 ALAPEYAK;D	2.22E+10	2.36E+10	2.65E+10	2.47E+10	2.34E+10														
Q6NUM9	10 ATVQSVLLDS	5.67E+08	7.21E+08	8.01E+08	7.77E+08	5.43E+08														
Q13131	6 FEC[+57.021		3.75E+08	3.94E+08	4.05E+08	3.00E+08														
Q09666	271 ADIDISGPNV		1.08E+11	1.21E+11	1.17E+11	9.81E+10														
Q13151	6 EDIYSGGGG		1.76E+09	1.97E+09	1.95E+09	1.66E+09														
P31150 (	19 DWNVDLIPK		4.82E+09	5.60E+09	5.15E+09	4.26E+09														
Q01813	26 AMEWITAK;		6.37E+09	8.18E+09	9.46E+09	6.67E+09														
Q9H3P7	10 DC[+57.0214		4.17E+08	5.34E+08	4.37E+08	3.17E+08														
015213	9 ALAEVDVISL		8.28E+08	9.90E+08	8.96E+08	7.67E+08														
Q8N9T8	13 AQEEADYIEV		1.54E+09	1.60E+09	1.41E+09	1.40E+09														
P20290   E	9 APLATGEDD		3.90E+09	3.18E+09	2.51E+09	2.38E+09														
Q9NWV8	7 AVGAQASVO		1.75E+09	2.13E+09	1.74E+09	1.72E+09														
P49419 /	27 DLPLAQGIK;E		6.55E+09	6.92E+09	6.77E+09	5.63E+09														
Q96E11	10 DTVSEDTIR;E 137 ADVVESWIG		2.51E+09 3.15E+10	2.66E+09 3.84E+10	2.22E+09 3.56E+10	1.99E+09 3.17E+10														
Q13813	9 AC[+57.0214		1.59E+09	1.52E+09	1.52E+09	1.40E+09														
P62873 (	19 AANDAGYFN		6.90E+09	7.28E+09	7.06E+09	1.40E+09 6.69E+09														
Q92616	96 AASQSTQVP		1.88E+10	2.11E+10	2.04E+10	1.73E+10														
Q92616	30 ALEIIPR;ATIS		1.54E+10	1.74E+10	2.04E+10 1.61E+10	1.52E+10														
Q98WM	15 AGVVTPGITE		3.30E+09	3.65E+09	3.83E+09	2.71E+09														
Q14980	82 AADALEEQQ		1.37E+10	1.72E+10	1.49E+10	1.42E+10														
Q14360	28 AINC[+57.02		3.25E+09	3.88E+09	3.92E+09	3.18E+09												-		
Q8WY22	3 SSPSGPSNPS		8.03E+08	1.68E+09	1.14E+09	1.39E+09														
P60891 F	4 ENISEWR:FS		7.61E+08	8.84E+08	9.39E+08	7.82E+08														
0754891	10 AANWYER:D		1.50E+09	1.85E+09	1.51E+09	1.19E+09														
Q8NF37	6 AAPASSAGA		1.47E+09	1.52E+09	1.43E+09	1.33E+09														
P56192	31 AEVLISTVGP		1.26E+10	1.45E+10	1.31E+10	1.11E+10														
P30084   E	18 AFAAGADIK;		8.84E+09	9.35E+09	9.24E+09	8.07E+09														
Q9BZK7	9 EGGQDVPSN		5.97E+08	7.08E+08	6.70E+08	5.18E+08														
Q9H0A0	33 AGFVPVYLR;		4.12E+09	4.45E+09	4.20E+09	3.80E+09														
A2A3N6	7 AAAASAAEA	6.98E+08	7.31E+08	8.96E+08	8.97E+08	8.39E+08														
Q9NX24	4 ADPDGPEAQ	8.86E+08	9.51E+08	1.12E+09	1.10E+09	8.85E+08														
P55809 S	16 AGGAGVPAF	3.12E+09	3.87E+09	4.38E+09	4.34E+09	3.42E+09														
P25490 1	7 DIDHETVVEE		5.50E+08	4.02E+08	4.27E+08	3.26E+08														
P42166 l	28 ALEESESSQL	6.03E+09	6.95E+09	8.64E+09	7.72E+09	6.22E+09														
qu	ant all files DDA NO	E27 elib	+																	

In excel, sort values in increasing order. For example, in HeLa, a mammalian immortalized cancer cell line, the most abundant proteins are often metabolic enzymes (G3P) or filament proteins (ACTB).

ome Insert Draw	Page Layout	Formula	s Data	Review	View	Automat	e Add-	ins 🔉 1	Fell me						Comm	ents 🖻	Shar
Paste S I		• A^ • • <u>A</u> •	E		v _	neral ~ % 9	▼ 00. 0→ 0€ 00.	Fo	nditional Forr rmat as Table Il Styles  v	-	Insert Delete	✓	Sort 8	Find & Select	Analyze Data	Sensitivity	
1 🌲 🗙 🗸 j	x sp P04406 G3P	HUMAN															
А	B C	D	E	F	G	н	I.	J	к	L	М	N	0	Р	Q	R	
	nPeptide PeptideSequ 2		2022_09_01_2		022_09_01		HeLa_500ng_	16mzst_DD/	A_NCE_42.mz	ML							
NCE 💌	<b>v v</b>	22 💌	27 🚽	32 💌	37 💌	42 💌											
sp P60709 ACTB_HUM	22 C[+57.02146	3.03E+11	3.43E+11	3.92E+11	3.61E+11	2.91E+11											
sp P68104 EF1A1_HUN	18 EHALLAYTLG	1.93E+11	1.98E+11	2.36E+11	2.21E+11	2.03E+11											
sp P06733 ENOA_HUN	31 AAVPSGAST	1.68E+11	1.94E+11	2.35E+11	2.03E+11	1.83E+11											
sp P04908 H2A1B_HUN	5 AGLQFPVGR	1.88E+11	1.82E+11	2.73E+11	1.77E+11	2.15E+11											
sp P14618 KPYM_HUM	39 AEGSDVANA	1.75E+11	1.80E+11	2.13E+11	1.89E+11	1.75E+11											
sp P68431 H31_HUMA	6 EIAQDFK;QL/	1.69E+11	1.76E+11	2.09E+11	2.03E+11	1.84E+11											
sp P05787 K2C8_HUM/	42 AEAESMYQI	1.39E+11	1.54E+11	1.67E+11	1.54E+11	1.41E+11											
sp Q92621 NU205_HU	40 AQIEQVIANC	5.52E+10	1.47E+11	1.76E+11	1.73E+11	1.46E+11											
p P04406 G3P_HUMA	25 GALQNIIPAS	1.30E+11	1.43E+11	1.55E+11	1.44E+11	1.27E+11											
sp P07900 HS90A_HUN	46 ADLINNLGTI	1.20E+11	1.34E+11	1.43E+11	1.40E+11	1.19E+11											
sp P78527 PRKDC_HUN	198 AALSALESFLI	1.26E+11	1.30E+11	1.48E+11	1.40E+11	1.29E+11											
sp P00558 PGK1_HUM	36 AC[+57.0214	1.23E+11	1.30E+11	1.50E+11	1.43E+11	1.26E+11											
sp P55060 XPO2_HUM	46 AAC[+57.021	1.15E+11	1.23E+11	1.76E+11	1.66E+11	1.17E+11											
sp P62805 H4_HUMAN	11 DAVTYTEHAN	9.66E+10	1.16E+11	1.33E+11	1.24E+11	1.14E+11											
sp P10412 H14_HUMA	5 ALAAAGYDV	9.42E+10	1.12E+11	1.19E+11	1.11E+11	8.55E+10											
sp Q09666 AHNK_HUN	271 ADIDISGPNV	1.02E+11	1.08E+11	1.21E+11	1.17E+11	9.81E+10											
p   P00338   LDHA_HUM	20 DDVFLSVPC[	8.98E+10	1.03E+11	1.08E+11	1.08E+11	9.79E+10											
sp P21333 FLNA_HUM	122 AEAGVPAEF	8.78E+10	1.02E+11	1.15E+11	1.09E+11	9.46E+10											
sp P05783 K1C18_HUN	33 AQIFANTVD	1.00E+11	9.71E+10	1.05E+11	9.88E+10	9.95E+10											
sp 060814 H2B1K_HUI	8 AMGIMNSF\	9.36E+10	9.70E+10	1.21E+11	1.10E+11	1.08E+11											
sp P31327 CPSM_HUM	82 AADTIGYPVN	8.72E+10	9.39E+10	1.04E+11	9.76E+10	8.82E+10											
sp P13639 EF2_HUMAI	66 ALLELQLEPE	8.66E+10	9.15E+10	1.05E+11	9.90E+10	8.64E+10											
p Q06830 PRDX1_HUI	17 ADEGISFR;A	8.17E+10	8.94E+10	1.01E+11	9.51E+10	8.31E+10											
sp Q15149 PLEC_HUM	255 AAEEAEEAR;	8.09E+10	8.75E+10	1.02E+11	9.53E+10	8.45E+10											
sp P62937 PPIA_HUMA	10 EGMNIVEAN	7.94E+10	8.68E+10	8.98E+10	8.94E+10	7.60E+10											
sp P11142 HSP7C_HUN	31 ARFEELNADI	7.77E+10	8.67E+10	9.20E+10	9.11E+10	8.00E+10											
sp P08670 VIME_HUM	37 ARVEVER;DG	7.64E+10	8.41E+10	9.80E+10	8.87E+10	8.16E+10											
sp P10809 CH60_HUM/	44 AAVEEGIVLG	7.65E+10	8.26E+10	9.21E+10	8.57E+10	7.38E+10											
sp PODPH7 TBA3C_HUI	15 AFVHWYVGE	8.22E+10	8.20E+10	9.34E+10	9.04E+10	8.11E+10											
sp P08238 HS90B_HUN	37 ADHGEPIGR;	7.09E+10	8.06E+10	8.34E+10	8.00E+10	7.17E+10											
sp P08729 K2C7_HUM/	30 AEAEAWYQT	7.03E+10	7.90E+10	8.48E+10	8.24E+10	7.03E+10											
sp P04075 ALDOA_HUI	26 AAQEEYVK;A	7.21E+10	7.85E+10	8.88E+10	8.53E+10	7.59E+10											
splA5A3E01POTEF HUI	8 AGFAGDDAP	7.58E+10	7.76E+10	8.38E+10	8.11E+10	7.26E+10											
quant all files	DDA_NCE27.elib_	+															

### Calculate the sum of all intensities in each sample. Divide each intensity by this sum.

Home	Insert	Draw	Page Layout	Formula	as Data	Review	View	Automat	te Add	ine O	Tell me					ſ			≳ e <sup>g</sup>
Home	Insert	Draw	Page Layout	Formula	as Data	Review	view	Automa	te Add	-ins ⊉	rell me						Comme	ents 🖻	Share
Paste	X [≞ ~ ≪	Calibri (B	ody) ~ 1 ⊔ ~   ⊞ ~	2 • A^	A <sup>×</sup>   ≡ ×   ≡		Ge Ge	neral • % 9			conditional For format as Table cell Styles 🗸		Insert Delete Forma	~ <b>J</b>	Sort & Filter	Find & Select	Analyze Data	Sensitivity	
SUM	\$ ×	✓ fx	=SUM(D3:D49	36)															
	A	1	з с	D	E	F	G	н	1	J	K	L	м	N	0	Р	Q	R	
1 Protein	1	NumP	eptide PeptideSequ	2022 09 01	2022 09 01	2022 09 01	2022 09 01	2022 09 01	HeLa 500ne	Protein	NumPeptide	PeptideSequ	2022 09 01	2022 09 0	1 2022 09 0	1 2022 09 0	1 2022 09 0	1 HeLa 500r	ng 16m7
2	NCE		v v		27 +	32 💌	37 💌	42 💌	_neea_seena	NCE	rtain eptide	replaceda	22	27	32	37	42	<u></u>	-6_x0.112
3 spl P60	709 ACTB		22 C[+57.02146		3.43E+11	3.92E+11	3.61E+11	2.91E+11		Summed in	ntensity		=SUM(D3:D4						
	104   EF1A1		18 EHALLAYTLO		1.98E+11	2.36E+11	2.21E+11	2.03E+11											
	733 ENOA		31 AAVPSGAST	1.68E+11	1.94E+11	2.35E+11	2.03E+11	1.83E+11											
	908   H2A1B		5 AGLQFPVG		1.82E+11	2.73E+11	1.77E+11	2.15E+11											
	618   KPYM		39 AEGSDVAN		1.80E+11	2.13E+11	1.89E+11	1.75E+11											
	431 H31 H		6 EIAQDFK;QL		1.76E+11	2.09E+11	2.03E+11	1.84E+11											
9 spl P05	787   K2C8_H	ними	42 AEAESMYQ		1.54E+11	1.67E+11	1.54E+11	1.41E+11											
	2621 NU205		40 AQIEQVIAN		1.47E+11	1.76E+11	1.73E+11	1.46E+11											
	406 G3P_H		25 GALQNIIPAS		1.43E+11	1.55E+11	1.44E+11	1.27E+11											
	900 HS90A		46 ADLINNLGT	1.20E+11	1.34E+11	1.43E+11	1.40E+11	1.19E+11											
	527 PRKDC		198 AALSALESFI	1.26E+11	1.30E+11	1.48E+11	1.40E+11	1.29E+11											
	558 PGK1	-	36 AC[+57.021		1.30E+11	1.50E+11	1.43E+11	1.26E+11											
15 spl P55	060   XPO2	ним	46 AAC[+57.02	1 1.15E+11	1.23E+11	1.76E+11	1.66E+11	1.17E+11											
	805 H4_HU		11 DAVTYTEHA		1.16E+11	1.33E+11	1.24E+11	1.14E+11											
	412 H14_H		5 ALAAAGYD		1.12E+11	1.19E+11	1.11E+11	8.55E+10											
	666 AHNK		271 ADIDISGPN		1.08E+11	1.21E+11	1.17E+11	9.81E+10											
	338 LDHA		20 DDVFLSVPC		1.03E+11	1.08E+11	1.08E+11	9.79E+10											
	333   FLNA_		122 AEAGVPAER	8.78E+10	1.02E+11	1.15E+11	1.09E+11	9.46E+10											
	783   K1C18		33 AQIFANTVD		9.71E+10	1.05E+11	9.88E+10	9.95E+10											
	0814 H2B1K		8 AMGIMNSF	9.36E+10	9.70E+10		1.10E+11	1.08E+11											
	327 CPSM		82 AADTIGYPV		9.39E+10	1.04E+11	9.76E+10	8.82E+10											
	639 EF2_H		66 ALLELQLEPE		9.15E+10	1.05E+11	9.90E+10	8.64E+10											
	830   PRDX1		17 ADEGISFR;A	8.17E+10	8.94E+10	1.01E+11	9.51E+10	8.31E+10											-
	149 PLEC	-	255 AAEEAEEAR		8.75E+10	1.02E+11	9.53E+10	8.45E+10											
	937   PPIA H		10 EGMNIVEA		8.68E+10	8.98E+10	8.94E+10	7.60E+10											
	142   HSP7C		31 ARFEELNAD		8.67E+10	9.20E+10	9.11E+10	8.00E+10											
	670   VIME_		37 ARVEVER;D		8.41E+10	9.80E+10	8.87E+10	8.16E+10											
	809 CH60		44 AAVEEGIVL		8.26E+10	9.21E+10	8.57E+10	7.38E+10											
	PH7 TBA3C		15 AFVHWYVG		8.20E+10	9.34E+10	9.04E+10	8.11E+10											
	238   HS90B	-	37 ADHGEPIGR		8.06E+10	8.34E+10	8.00E+10	7.17E+10											
	729   K2C7_I		30 AEAEAWYQ		7.90E+10	8.48E+10	8.24E+10	7.03E+10											
	075 ALDOA		26 AAQEEYVK;		7.85E+10	8.88E+10	8.53E+10	7.59E+10											
	3E0 POTER		8 AGFAGDDA		7.76E+10	8.38E+10	8.11E+10	7.26E+10											
4			DA NCE27.elib	+															

Calculate the average of the summed intensity for all samples (using the values calculated in step 2). Multiply each value by the average summed intensity. Repeat for all values in the data matrices.

_				,						<u> </u>	•								
• •	Autos	Save 🔵 off	) A B C	<b>トッ</b>	*** 🄛 🤅	ə ··· 🖻	quant_all	_files_DDA_I	NCE27.	elib_uno	rmalized.	proteins ~							er 6
Home	Insert	Draw	Page Layout	Formula	s Data	Review	View	Automate	Add-i	ns 🄉	Tell me						🖵 Commei	nts 🖻 🖻 Sh	nare
Paste	× È≞ v ≪	Calibri (Bo B I L	ody)	<ul> <li>✓ A<sup>*</sup></li> <li>△ ✓ A ✓</li> </ul>			× .	neral ✔ % 🤊	▼ 00.00	F	onditional Fo ormat as Tab ell Styles  v		Insert Delete	~ .	Sort & Filter	Find & Select	Analyze Data	Sensitivity	
SUM	÷ ×	$\checkmark f_x$	=AVERAGE(M3:0	Q3)															*
	A	В	AVERAGE(num	ber1, [numbe	er2],)	F	G	н	1	J	к	L	М	N	0	Р	Q	R	S
1 Prote	in	NumPe	AVERAGE	01 2	022 09 01	2022 09 01 2	022 09 01	2022_09_01_HeL	a 500ng f	Protein	NumPeptio	de PeptideSegu	2022 09 01	2022 09 01	2022 09 01	2022 09 01	2022 09 01	HeLa_500ng	16mzs
2	NCE	-	Functions	•	27 ++	32 💌	37 💌	42 💌		NCE			22	27	32	37	42		
3 sp P6	50709 ACTB_H	IUM.	AVERAGE	11	3.43E+11	3.92E+11	3.61E+11	2.91E+11	9	Summed in	tensity			1.92E+13	2.21E+13	2.05E+13	1.83E+13	1	
	58104   EF1A1_		AVERAGEA	11	1.98E+11	2.36E+11	2.21E+11	2.03E+11	/	Average Su	mmed 👔	sit =AVERAGE(							
5 sp P0	06733   ENOA_	HUN	AVERAGEIF	11	1.94E+11	2.35E+11	2.03E+11	1.83E+11			_								
6 sp P0	04908 H2A1B	HUN		11	1.82E+11	2.73E+11	1.77E+11	2.15E+11											
7 sp P1	14618   KPYM_I	ним	AVERAGEIFS	11	1.80E+11	2.13E+11	1.89E+11	1.75E+11											
8 sp P6	58431 H31_HU	JMA	DAVERAGE	11	1.76E+11	2.09E+11	2.03E+11	1.84E+11											
9 sp P0	05787 K2C8_H	UM/	42 AEAESMYQI	1.39E+11	1.54E+11	1.67E+11	1.54E+11	1.41E+11											
10 sp Q	92621 NU205	_HU	40 AQIEQVIANC	5.52E+10	1.47E+11	1.76E+11	1.73E+11	1.46E+11											
11 sp P0	04406 G3P_H	JMA	25 GALQNIIPAS	1.30E+11	1.43E+11	1.55E+11	1.44E+11	1.27E+11											
	07900   HS90A_		46 ADLINNLGTI/	1.20E+11	1.34E+11	1.43E+11	1.40E+11	1.19E+11											
13 sp P7	78527   PRKDC_	HUN	198 AALSALESFLI	1.26E+11	1.30E+11	1.48E+11	1.40E+11	1.29E+11											
	00558 PGK1_H		36 AC[+57.0214	1.23E+11	1.30E+11	1.50E+11	1.43E+11	1.26E+11											
	55060 XPO2_H		46 AAC[+57.021	1.15E+11	1.23E+11	1.76E+11	1.66E+11	1.17E+11											
	52805   H4_HUI		11 DAVTYTEHAF	9.66E+10	1.16E+11	1.33E+11	1.24E+11	1.14E+11											
	10412 H14_HU		5 ALAAAGYDV	9.42E+10	1.12E+11	1.19E+11	1.11E+11	8.55E+10											
	09666 AHNK_		271 ADIDISGPNV	1.02E+11	1.08E+11	1.21E+11	1.17E+11	9.81E+10											
	00338 LDHA_H		20 DDVFLSVPC[-	8.98E+10	1.03E+11	1.08E+11	1.08E+11	9.79E+10											
	21333 FLNA_H		122 AEAGVPAEFS	8.78E+10	1.02E+11	1.15E+11	1.09E+11	9.46E+10											
	05783   K1C18_		33 AQIFANTVDI	1.00E+11	9.71E+10	1.05E+11	9.88E+10	9.95E+10											
	60814 H2B1K		8 AMGIMNSF\	9.36E+10	9.70E+10	1.21E+11	1.10E+11	1.08E+11											
	31327   CPSM_I		82 AADTIGYPVN	8.72E+10	9.39E+10	1.04E+11	9.76E+10	8.82E+10											
	13639 EF2_HL		66 ALLELQLEPE	8.66E+10	9.15E+10	1.05E+11	9.90E+10	8.64E+10											
	06830   PRDX1		17 ADEGISFR;A	8.17E+10	8.94E+10	1.01E+11	9.51E+10	8.31E+10											
	15149   PLEC_F		255 AAEEAEEAR;	8.09E+10	8.75E+10	1.02E+11	9.53E+10	8.45E+10										<u> </u>	
	52937   PPIA_H		10 EGMNIVEAN	7.94E+10	8.68E+10	8.98E+10	8.94E+10	7.60E+10											
	11142   HSP7C_		31 ARFEELNADL	7.77E+10	8.67E+10	9.20E+10	9.11E+10	8.00E+10											
	08670   VIME_H		37 ARVEVER;DG	7.64E+10	8.41E+10	9.80E+10	8.87E+10	8.16E+10 7.38E+10											
	10809   CH60_F DDPH7   TBA3C		44 AAVEEGIVLG 15 AFVHWYVGE	7.65E+10 8.22E+10	8.26E+10 8.20E+10	9.21E+10 9.34E+10	8.57E+10 9.04E+10	7.38E+10 8.11E+10											
	08238   HS90B		37 ADHGEPIGR;	7.09E+10	8.20E+10 8.06E+10	9.34E+10 8.34E+10	9.04E+10 8.00E+10	7.17E+10											
	08238 H590B_ 08729 K2C7_H		30 AEAEAWYQT	7.09E+10 7.03E+10	7.90E+10	8.34E+10 8.48E+10	8.00E+10 8.24E+10	7.03E+10											
	04075 ALDOA		26 AAQEEYVK;A	7.03E+10 7.21E+10	7.90E+10 7.85E+10	8.48E+10 8.88E+10	8.24E+10 8.53E+10	7.59E+10											
	5A3E01POTEF		8 AGFAGDDAP	7.58E+10	7.85E+10 7.76E+10	8.38E+10 8.38E+10	8.53E+10 8.11E+10	7.26E+10											
35 SUTA.			DA_NCE27.elib_	+	7.70L+10	0.001+10	0.111-10	7.202+10											
- P	quant	an_mes_D	DA_NOE21.enD_																

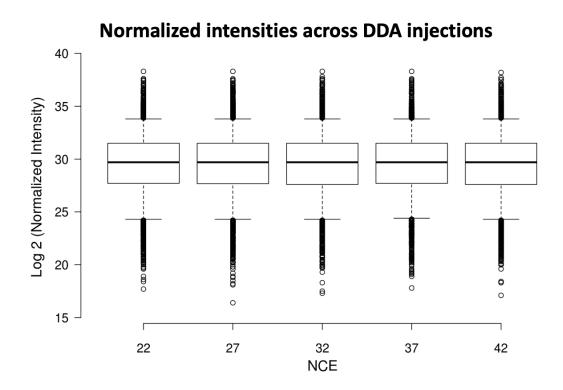
### Log<sub>2</sub> transform the data matrix.

Calib	T 12.m Protein Kr> NCE sp P60709 ACT3_HUM sp P60304 F2A1_HUM sp P64304 F2A1_HUM sp P64314 R3_HUM sp P6431 R3_HUM sp P6431 R3_HUM sp P63771 K2C8_HUM sp P64541 R3_HUM sp P65751 R3_HUM sp P65754 R3_HUM sp P657541 R3_HUM sp P65754 R3_HU	U V NumPeptide PeptideSequ 22 (1-57.02146 18 EHALAYTIG 31 AAVPSGASTI 5 AGLOPVGR 39 AEGSDVANA 6 EIAQPKGU 42 AEGSBVANA	W 2022_09_01 21 22	View	Automate	xt ~ « Center ~ Z	Scientifi \$ • 9 AA 022_09_01_	> Star Star Star Star Star Star Star Star	ditional Format Cell matting as Table Styles	E Delete V Format V	∑ • A ↓ Z Sort & Filter AG		Analyze Data	Sensitivi
te S 16mzst_DDA_NCE_42.1	I         U         -         -           fx         =LOG(W3, 2)         -         -           12.m         Protein         -         -           12.m         Protein         -         -           15.p         P607091 ACT8, HUM         -         -           15.p         P607091 ACT8, HUM         -         -           16.p         P07381 RAZALE, HUM         -         -           16.p         P076481 REALE, HUM         -         -           16.p         P076381 RAZALE, HUM         -         -           16.p         P0764981 REALE, HUM         -         -           16.p         P067381 RAZALE, HUM         -         -           16.p         P067381 RAZE, HUM         -         -	U V NumPeptide PeptideSequi 22 C(+57 02146 18 EHALAYTIG 31 AAVPSGASTI 5 AGLOPVGR 39 AEGSDVANA 6 EIAQPKGU 42 AEAESMOQH 42 AEAESMOQH	W 2022_09_01_22 3.335+11 2.12E+11 1.85E+11 2.07E+11	K 022_09_01 27 3.50E+11 2.02E+11	Y 2022_09_01_2 32	Z 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	\$ ~ % AA 022_09_01_	•         •	AD	E Delete V Format V	↓ ↓ ∠ U Sort & Filter	Find & Select	Analyze Data	
te → B ★ × ✓ S 16mzst_DDA_NCE_42.1		U V NumPeptide PeptideSequ 2 22 (1-57.02146) 18 EHALLAYTLG 31 AAVPSGAST 39 AEGSDVANA 6 EIAQDFXQL 42 AEASMQH 42 AEASMQH	W 2022_09_01_2/ 22 3.33E+11 2.12E+11 1.85E+11 2.07E+11	× 022_09_01_: 27 3.50E+11 2.02E+11	Y 2022_09_01_2 32	Z 022_09_01_2	AA 022_09_01_	•         •	AD	AE AF	Sort & Filter	Select	Analyze Data	
		U V NumPeptide PeptideSequ 2 22 (1-57.02146) 18 EHALLAYTLG 31 AAVPSGAST 39 AEGSDVANA 6 EIAQDFXQL 42 AEASMQH 42 AEASMQH	W 2022_09_01_2/ 22 3.33E+11 2.12E+11 1.85E+11 2.07E+11	× 022_09_01_: 27 3.50E+11 2.02E+11	Y 2022_09_01_2 32	Z 022_09_01_2	AA 022_09_01_	AB AC	matting as Table Styles AD	AE AF	AG	Select	Data	
S 16mzst_DDA_NCE_42.	T 12.m Protein IX-> NCE sp P60709 ACT8_HUM sp P60304 F2A1_HUM sp P60308 H2A1B_HUI sp P44518 KPVM_HUM sp P60371 H3_HUM sp P60377 K2C8_HUM sp P063761 K2C8_HUM sp P063761 K2C8_HUM	NumPeptide PeptideSequ 2 2 C[+57.02146 31 AAVPSGASTC 5 AGLQFPVGR 39 AEGSDVANA 6 ELAQPK;QL 42 AEAESMYQH	2022_09_01_2 22 3.33E+11 2.12E+11 1.85E+11 2.07E+11	022_09_01 27 3.50E+11 2.02E+11	2022_09_01_2 32	022_09_01_2	022_09_01_					AH	AI	AJ
16mzst_DDA_NCE_42.	I2.m         Protein           IX->         NCE           spl P60709 JACT8_HUM         spl P64090 JACT8_HUM           spl P64308 JENOA_HUN         spl P64308 JENOA_HUN           spl P146318 JENOA_HUN         spl P64331 JENOA_HUN           spl P146318 JENOA_HUN         spl P64331 JENOA_HUN           spl P64331 JENOA_HUN         spl P64331 JENOA_HUN           spl P64311 JENA_HUN         spl P64341 JENA_HUN           spl P6431 JENA_HUN         spl P643405 JENA_HUN           spl P64406 JG3P_HUN4         spl P64406 JG3P_HUN4	NumPeptide PeptideSequ 2 2 C[+57.02146 31 AAVPSGASTC 5 AGLQFPVGR 39 AEGSDVANA 6 ELAQPK;QL 42 AEAESMYQH	2022_09_01_2 22 3.33E+11 2.12E+11 1.85E+11 2.07E+11	022_09_01 27 3.50E+11 2.02E+11	2022_09_01_2 32	022_09_01_2	022_09_01_					AH	AI	AJ
	IX->         NCE           spl P660709   ACT8_HUM         spl P66104   EF1A1_HUM           spl P66104   EF1A1_HUM         spl P04508   HAA1B_HUI           spl P64308   HAA1B_HUI         spl P04518   KPVM_HUM           spl P65787   K2C8_HUM         spl P05787   K2C8_HUM           spl P04512   INJ25_HUM         spl P05787   K2C8_HUM	22 C[+57.02146 18 EHALLAYLG 31 AAVPSGAST( 5 AGLQFPVGR 39 AEGSDVANA 6 EIAQDFK;QL 42 AEAESMYQH	22 3.33E+11 2.12E+11 1.85E+11 2.07E+11	27 3.50E+11 2.02E+11	32									
IORMALIZED MATRIX-	sp  P60709   ACTB_HUM sp  P68104   EF1A1_HUP sp  P06733   ENOA_HUN sp  P04698   HA21B_HUU sp  P14618   KPYM_HUN sp  P68431   H31_HUMA sp  P05787   K2CB_HUM sp  P05787   K2CB_HUMA	18 EHALLAYTLG 31 AAVPSGAST 5 AGLQFPVGR 39 AEGSDVANA 6 EIAQDFK;QL/ 42 AEAESMYQIF	3.33E+11 2.12E+11 1.85E+11 2.07E+11	3.50E+11 2.02E+11		37				NumPeptide: PeptideSe				
	sp   P68104   EF1A1_HUN sp   P06733   ENOA_HUN sp   P04908   H2A1B_HUN sp   P4618   KPYM_HUN sp   P68431   H31_HUMA sp   P05787   K2C8_HUM sp   P05421   NU205_HU sp   P04406   G3P_HUMA	18 EHALLAYTLG 31 AAVPSGAST 5 AGLQFPVGR 39 AEGSDVANA 6 EIAQDFK;QL/ 42 AEAESMYQIF	2.12E+11 1.85E+11 2.07E+11	2.02E+11	3.47E+11		42	LOG2 TRANSFORM			22	27	32	37
	sp   P06733   ENOA_HUN sp   P04908   H2A1B_HUI sp   P14618   KPYM_HUN sp   P68431   H31_HUM sp   P05787   K2C8_HUM sp   Q2621   NU205_HU sp   P04406   G3P_HUMA	A 31 AAVPSGAST( 5 AGLQFPVGR 39 AEGSDVANA 6 EIAQDFK;QL/ 42 AEAESMYQIF	1.85E+11 2.07E+11			3.44E+11	3.12E+11		sp P60709 ACTB_HUM		46 =LOG(W3,2)	3.83E+01	3.83E+01	3.83E-
	sp P04908 H2A1B_HUI sp P14618 KPYM_HUN sp P68431 H31_HUMA sp P05787 K2C8_HUM sp Q92621 NU205_HU sp P04406 G3P_HUMA	AGLQFPVGR 39 AEGSDVANA 6 EIAQDFK;QL/ 4 42 AEAESMYQI	2.07E+11	1.98E+11	2.09E+11	2.10E+11	2.18E+11		sp P68104 EF1A1_HUN	18 EHALLATT		3.76E+01	3.76E+01	3.76E+
	sp P14618  KPYM_HUM sp P68431 H31_HUMA sp P05787 K2C8_HUM sp Q92621 NU205_HU sp P04406 G3P_HUMA	4 39 AEGSDVANA 6 EIAQDFK;QL/ 4 42 AEAESMYQI			2.08E+11	1.94E+11	1.96E+11		sp P06733 ENOA_HUM	31 AAVPSGAS		3.75E+01	3.76E+01	3.75E-
	sp P68431 H31_HUMA sp P05787 K2C8_HUM sp Q92621 NU205_HU sp P04406 G3P_HUMA	6 EIAQDFK;QU 42 AEAESMYQI	1.92E+11	1.86E+11	2.42E+11	1.68E+11	2.31E+11		sp P04908 H2A1B_HUM	5 AGLQFPVC		3.74E+01	3.78E+01	3.73E-
	sp P05787 K2C8_HUM sp Q92621 NU205_HU sp P04406 G3P_HUMA	42 AEAESMYQI	4.005.44	1.84E+11	1.89E+11	1.80E+11	1.88E+11		sp P14618 KPYM_HUM	39 AEGSDVA		3.74E+01	3.75E+01	3.74E
	sp Q92621 NU205_HU sp P04406 G3P_HUMA		1.86E+11 1.53E+11	1.80E+11 1.57E+11	1.85E+11 1.48E+11	1.94E+11 1.47E+11	1.97E+11 1.51E+11		sp P68431 H31_HUMA	6 EIAQDFK;C 42 AEAESMYC		3.74E+01 3.72E+01	3.74E+01 3.71E+01	3.75E
	sp P04406 G3P_HUMA		6.07E+10	1.5/E+11 1.50E+11	1.48E+11 1.56E+11	1.4/E+11 1.65E+11	1.51E+11 1.56E+11		sp P05787 K2C8_HUM/ sp Q92621 NU205_HU	42 AEAESIVIT		3.71E+01	3.72E+01 3.72E+01	3.73E
			1.43E+11	1.46E+11	1.37E+11	1.37E+11	1.36E+11		sp P04406 G3P_HUMA	25 GALQNIIP		3.71E+01	3.72E+01 3.70E+01	3.70E
	sp P07900 HS90A_HU		1.43E+11 1.32E+11	1.40E+11 1.37E+11	1.37E+11 1.26E+11	1.37E+11	1.38E+11		sp P07900 HS90A_HUN	46 ADLINNLG		3.70E+01	3.69E+01	3.70E
	sp P78527 PRKDC_HU		1.39E+11	1.33E+11	1.31E+11	1.34E+11	1.38E+11		sp1P785271PRKDC_HUN	198 AALSALES		3.69E+01	3.69E+01	3.70E
	sp P00558 PGK1_HUM		1.36E+11	1.32E+11	1.33E+11	1.36E+11	1.35E+11		sp P00558 PGK1_HUM	36 AC[+57.02		3.69E+01	3.70E+01	3.70E
	sp P55060 XPO2_HUM		1.26E+11	1.26E+11	1.56E+11	1.58E+11	1.25E+11		sp P55060 XPO2_HUM	46 AAC[+57.0		3.69E+01	3.72E+01	3.72E
	sp P62805 H4_HUMAN		1.06E+11	1.18E+11	1.18E+11	1.18E+11	1.22E+11		sp P62805 H4_HUMAN	11 DAVTYTEH		3.68E+01	3.68E+01	3.68E
	sp P10412 H14_HUMA		1.04E+11	1.14E+11	1.06E+11	1.06E+11	9.16E+10		sp P10412 H14_HUMA	5 ALAAAGYI		3.67E+01	3.66E+01	3.66E
	sp Q09666 AHNK_HUN		1.12E+11	1.10E+11	1.07E+11	1.11E+11	1.05E+11		sp Q09666 AHNK_HUN	271 ADIDISGP		3.67E+01	3.66E+01	3.67E
	sp P00338 LDHA_HUM		9.87E+10	1.05E+11	9.59E+10	1.03E+11	1.05E+11		sp P00338 LDHA_HUM	20 DDVFLSVP		3.66E+01	3.65E+01	3.66E
	sp P21333 FLNA_HUM		9.65E+10	1.04E+11	1.02E+11	1.04E+11	1.01E+11		sp P21333 FLNA_HUM	122 AEAGVPA		3.66E+01	3.66E+01	3.66E
	sp  P05783   K1C18_HUN		1.10E+11	9.91E+10	9.28E+10	9.42E+10	1.07E+11		sp  P05783  K1C18_HUN	33 AQIFANTV		3.65E+01	3.64E+01	3.65E
	sp 060814 H2B1K_HU		1.03E+11	9.90E+10	1.07E+11	1.05E+11	1.16E+11		sp 060814 H2B1K_HUI	8 AMGIMNS		3.65E+01	3.66E+01	3.66E
	sp P31327 CPSM_HUN	82 AADTIGYPVN	9.59E+10	9.59E+10	9.19E+10	9.30E+10	9.45E+10		sp P31327 CPSM_HUM	82 AADTIGYP	VN 3.65E+01	3.65E+01	3.64E+01	3.64E
	sp P13639 EF2_HUMA		9.51E+10	9.34E+10	9.30E+10	9.44E+10	9.26E+10		sp P13639 EF2_HUMAI	66 ALLELQLEF	PEE 3.65E+01	3.64E+01	3.64E+01	3.65E
	sp Q06830 PRDX1_HU		8.98E+10	9.13E+10	8.97E+10	9.06E+10	8.91E+10		sp Q06830 PRDX1_HUI	17 ADEGISFR	;A 3.64E+01	3.64E+01	3.64E+01	3.64E
	sp Q15149 PLEC_HUM	255 AAEEAEEAR;	8.89E+10	8.93E+10	9.01E+10	9.09E+10	9.05E+10		sp Q15149 PLEC_HUM	255 AAEEAEEA	R; 3.64E+01	3.64E+01	3.64E+01	3.64E
	sp   P62937   PPIA_HUM/	10 EGMNIVEAN	8.72E+10	8.87E+10	7.96E+10	8.52E+10	8.15E+10		sp P62937 PPIA_HUMA	10 EGMNIVE	AN 3.63E+01	3.64E+01	3.62E+01	3.63E
	sp P11142 HSP7C_HUM	31 ARFEELNADI	8.54E+10	8.86E+10	8.15E+10	8.68E+10	8.58E+10		sp P11142 HSP7C_HUN	31 ARFEELNA	DL 3.63E+01	3.64E+01	3.62E+01	3.63E
	sp P08670 VIME_HUM	I. 37 ARVEVER;DG	8.40E+10	8.59E+10	8.69E+10	8.45E+10	8.74E+10		sp P08670 VIME_HUM	37 ARVEVER;	DG 3.63E+01	3.63E+01	3.63E+01	3.63E-
	sp P10809 CH60_HUM		8.41E+10	8.43E+10	8.16E+10	8.17E+10	7.91E+10		sp P10809 CH60_HUM	44 AAVEEGIV		3.63E+01	3.62E+01	3.62E-
	sp PODPH7 TBA3C_HU	15 AFVHWYVGE	9.03E+10	8.37E+10	8.28E+10	8.61E+10	8.69E+10		sp PODPH7 TBA3C_HUI	15 AFVHWYV	GE 3.64E+01	3.63E+01	3.63E+01	3.63E
	sp P08238 HS90B_HU		7.79E+10	8.23E+10	7.39E+10	7.63E+10	7.69E+10		sp P08238 HS90B_HUN	37 ADHGEPIG		3.63E+01	3.61E+01	3.62E-
	sp P08729 K2C7_HUM		7.73E+10	8.07E+10	7.51E+10	7.85E+10	7.53E+10		sp P08729 K2C7_HUM/	30 AEAEAWY		3.62E+01	3.61E+01	3.62E-
	sp P04075 ALDOA_HU		7.93E+10	8.01E+10	7.87E+10	8.13E+10	8.13E+10		sp P04075 ALDOA_HU	26 AAQEEYV		3.62E+01	3.62E+01	3.62E-
	sp A5A3E0 POTEF_HU		8.33E+10	7.92E+10	7.43E+10	7.73E+10	7.78E+10		sp A5A3E0 POTEF_HUI	8 AGFAGDD		3.62E+01	3.61E+01	3.62E-
	sp A6NMY6 AXA2L_HL		7.64E+10	7.89E+10	7.79E+10	7.92E+10	7.94E+10		sp A6NMY6 AXA2L_HU	23 ALLYLC[+5		3.62E+01	3.62E+01	3.62E+
			7.79E+10	7.67E+10	8.75E+10	7.88E+10	7.99E+10		sp 075369 FLNB_HUM	140 AAGSGELO		3.62E+01	3.63E+01	3.62E-
	sp 075369 FLNB_HUN		7.87E+10	7.39E+10	6.39E+10	7.66E+10	5.52E+10		sp P62736 ACTA_HUM	5 GYSFVTTA		3.61E+01	3.59E+01	3.62E-
	sp P62736 ACTA_HUM		6.88E+10	6.99E+10	6.92E+10	6.55E+10	6.73E+10		sp P0DMV8 HS71A_HU	35 AAAIGIDLO		3.60E+01	3.60E+01	3.59E+
	sp P62736 ACTA_HUM sp P0DMV8 HS71A_HU		6.90E+10	6.64E+10	6.58E+10	6.48E+10	6.57E+10		sp P11021 BIP_HUMAN	33 DAGTIAGL		3.60E+01	3.59E+01	3.59E+
	sp P62736 ACTA_HUM	1 20 DYFEEYGK;EI	6.09E+10	6.21E+10	6.30E+10	6.29E+10	6.27E+10		sp P22626 ROA2_HUM	20 DYFEEYGK	;EI 3.58E+01	3.59E+01	3.59E+01	3.59E-

To visualize relative intensities across injections, you can copy and paste your matrix into a downstream tool, such as BoxplotR: <u>http://shiny.chemgrid.org/boxplotr/</u>

nosave		<b>A B</b> C		• • c				<u></u>	quant_all_	mes_DDA_	_NCE27.em	b_unormai	nzeu.pro	interns -							Q
ne	Insert	Draw P	age Layou	t For	mulas	Data	Review	View	Automate	Add-ins	♀ Tell ı	me							<b>P</b>	Comments	🖻 S
te	• ·	Calibri (Body)					_	≫~ ~	ab Co Wrap Te:		General			Conditional Fo	ormat Cell		nsert ∨ Ielete ∨	Σ×Aς J Z	√ × ∕ v ort & Find &	Analyze	Sensiti
<	2	BIU	· · · ·	<u>~</u> ~	<u>A</u> v	EE		•= •=	👥 Merge &	Center V	\$ • %	9 .00		Formatting as		F F	ormat 🗸		iter Select	Data	ochait
		√ fx 2																			
	Y	Z	AA	AB		C		AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ.
01 2	022_09_01 32	2022_09_01 37	2022_09_01_ 42		ng_16mzst_ LOG2 TRAN			NCE	NumPeptide	PeptideSequ	2022_09_01	2022_09_01	2022_09_0 32	37	2022_09_01_H 42	eLa_500ng_	16mzst_Dl	DA_NCE_42.m	zML		
+11	3.47E+11		42 3.12E+11		LUGZ TRAN	SFORIVIED-		NCE 9 ACTB_HUN	22	C[+57.02146		3.83E+01	3.83E+0								
+11	2.09E+11		2.18E+11					4 EF1A1_HUI		EHALLAYTLG	3.76E+01	3.76E+01	3.76E+0								
E+11	2.09E+11		1.96E+11					3 ENOA_HUN		AAVPSGAST	3.74E+01	3.75E+01	3.76E+0		3.75E+01						
+11	2.42E+11		2.31E+11					8 H2A1B_HU		AGLQFPVGR	3.74E+01	3.74E+01	3.78E+0		3.77E+01						
+11	1.89E+11		1.88E+11					8 KPYM_HUN		AGEGSDVANA	3.75E+01	3.74E+01 3.74E+01	3.75E+0								
+11	1.85E+11		1.97E+11					1 H31_HUM4		EIAQDFK;QL	3.74E+01	3.74E+01	3.74E+0		3.75E+01						
+11	1.48E+11		1.51E+11					7   K2C8_HUM		AEAESMYQI	3.72E+01	3.72E+01	3.71E+0		3.71E+01						
+11	1.56E+11		1.56E+11					21 NU205_HU		AQIEQVIANO	3.58E+01	3.71E+01	3.72E+0		3.72E+01						
+11	1.37E+11		1.36E+11					6 G3P_HUM		GALQNIIPAS	3.71E+01	3.71E+01	3.70E+0		3.70E+01						
+11	1.26E+11		1.28E+11					0 HS90A HU		ADLINNLGTI	3.69E+01	3.70E+01	3.69E+0		3.69E+01						
+11	1.31E+11		1.38E+11					7 PRKDC_HU		AALSALESFL	3.70E+01	3.69E+01	3.69E+0								
+11	1.33E+11		1.35E+11					8 PGK1_HUN		AC[+57.0214	3.70E+01	3.69E+01	3.70E+0		3.70E+01						
+11	1.56E+11		1.25E+11					0 XPO2_HUN		AAC[+57.021	3.69E+01	3.69E+01	3.72E+0		3.69E+01						
+11	1.18E+11		1.22E+11					5 H4_HUMAN		DAVTYTEHA	3.66E+01	3.68E+01	3.68E+0		3.68E+01						
+11	1.06E+11	1.06E+11	9.16E+10					2 H14_HUMA		ALAAAGYDV	3.66E+01	3.67E+01	3.66E+0	3.66E+01	3.64E+01						
+11	1.07E+11		1.05E+11					66 AHNK_HU		ADIDISGPNV	3.67E+01	3.67E+01	3.66E+0		3.66E+01						
+11	9.59E+10		1.05E+11					8 LDHA_HUN		DDVFLSVPC[	3.65E+01	3.66E+01	3.65E+0	3.66E+01	3.66E+01						
+11	1.02E+11	1.04E+11	1.01E+11					3 FLNA_HUN		AEAGVPAEF	3.65E+01	3.66E+01	3.66E+0	3.66E+01	3.66E+01						
+10	9.28E+10	9.42E+10	1.07E+11				sp  P0578	3 K1C18_HU	33	AQIFANTVD	3.67E+01	3.65E+01	3.64E+0	1 3.65E+01	3.66E+01						
+10	1.07E+11	1.05E+11	1.16E+11				sp 0608	14 H2B1K_HU	1 8	AMGIMNSF	3.66E+01	3.65E+01	3.66E+0	3.66E+01	3.68E+01						
+10	9.19E+10	9.30E+10	9.45E+10					7 CPSM_HUN		AADTIGYPVN	3.65E+01	3.65E+01	3.64E+0	3.64E+01	3.65E+01						
+10	9.30E+10	9.44E+10	9.26E+10					9 EF2_HUMA		ALLELQLEPE	3.65E+01	3.64E+01	3.64E+0	1 3.65E+01	3.64E+01						
+10	8.97E+10	9.06E+10	8.91E+10				sp Q068	30 PRDX1_HU	17	ADEGISFR;A	3.64E+01	3.64E+01	3.64E+0	3.64E+01	3.64E+01						
+10	9.01E+10	9.09E+10	9.05E+10				sp Q151	49 PLEC_HUM	255	AAEEAEEAR	3.64E+01	3.64E+01	3.64E+0	3.64E+01	3.64E+01						
+10	7.96E+10	8.52E+10	8.15E+10				sp P6293	7 PPIA_HUM	4 10	EGMNIVEAN	3.63E+01	3.64E+01	3.62E+0	3.63E+01	3.62E+01						
+10	8.15E+10	8.68E+10	8.58E+10				sp P1114	2 HSP7C_HU	31	ARFEELNAD	3.63E+01	3.64E+01	3.62E+0	3.63E+01	3.63E+01						
+10	8.69E+10	8.45E+10	8.74E+10				sp  P0867	O VIME_HUN	1. 37	ARVEVER;D0	3.63E+01	3.63E+01	3.63E+0	3.63E+01	3.63E+01						
+10	8.16E+10	8.17E+10	7.91E+10				sp P1080	9 CH60_HUN	44	AAVEEGIVLG	3.63E+01	3.63E+01	3.62E+0	3.62E+01	3.62E+01						
+10	8.28E+10	8.61E+10	8.69E+10				sp PODP	H7 TBA3C_HU	15	AFVHWYVG	3.64E+01	3.63E+01	3.63E+0	3.63E+01	3.63E+01						
+10	7.39E+10	7.63E+10	7.69E+10				sp P0823	8 HS90B_HU	N 37	ADHGEPIGR;	3.62E+01	3.63E+01	3.61E+0	3.62E+01	3.62E+01						
+10	7.51E+10		7.53E+10				sp P0872	9   K2C7_HUM		AEAEAWYQ	3.62E+01	3.62E+01	3.61E+0		3.61E+01						
+10	7.87E+10		8.13E+10				sp P0407	5 ALDOA_HU		AAQEEYVK;A	3.62E+01	3.62E+01	3.62E+0								
+10	7.43E+10		7.78E+10					EO   POTEF_HU		AGFAGDDAF	3.63E+01	3.62E+01	3.61E+0		3.62E+01						
+10	7.79E+10		7.94E+10					1Y6 AXA2L_HL		ALLYLC[+57.0	3.62E+01	3.62E+01	3.62E+0		3.62E+01						
+10	8.75E+10		7.99E+10				sp 0753	59 FLNB_HUN		AAGSGELGV	3.62E+01	3.62E+01	3.63E+0		3.62E+01						
+10	6.39E+10		5.52E+10					6 ACTA_HUN		GYSFVTTAER	3.62E+01	3.61E+01	3.59E+0		3.57E+01						
E+10	6.92E+10		6.73E+10					V8 HS71A_HU		AAAIGIDLGT	3.60E+01	3.60E+01	3.60E+0		3.60E+01						
E+10	6.58E+10		6.57E+10					1 BIP_HUMA		DAGTIAGLN	3.60E+01	3.60E+01	3.59E+0		3.59E+01						
E+10	6.30E+10	6.29E+10	6.27E+10				sp P2262	6 ROA2_HUN	20	DYFEEYGK;E	3.58E+01	3.59E+01	3.59E+0	01 3.59E+01	3.59E+01						

Select destination and press ENTER or choose Paste



### 7. WALKTHROUGH WITH DEMO DATA

Tutorial for how to process data using Scribe and a Prosit generated library.

#### A. You can download data used in the Scribe manuscript from:

https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=99b781c9c0b84ac9bfc3f93734e9ddab

To start, we recommend downloading the following files for this walkthrough:

- 2022\_09\_01\_HeLa\_500ng\_16mzst\_DDA\_NCE\_27.mzML
- Uniprot\_human-reference\_reviewed\_2022mar02.fasta
- uniprot\_human-reference\_reviewed\_2022mar02.prosit\_input.trypsin\_nce29\_hcd2020.dlib

B. Specify [uniprot\_human-reference\_reviewed\_2022mar02.prosit\_input.trypsin\_nce29\_hcd2020.dlib] as the library and [uniprot\_human-reference\_reviewed\_2022mar02.fasta] as the background FASTA. The DLIB library was generated using the method explained in part 3 of the manual. We will use a "Normal Target/Decoy" approach, and select "Trypsin" as the enzyme. The fragmentation is set to CID/HCD (B/Y). Precursor mass tolerance, fragment mass tolerance, and library mass tolerance are all set at 10 PPM. The latest Percolator version should be selected, which at this time is v3-01. The settings should match the screenshot below.

Parameters:	
Library:uniprot_human-reference_reviewed_2022mar02.prosit_input.trypsin_nce29_hcd2020.dlib	Edit
Background:uniprot_human-reference_reviewed_2022mar02.fasta	Edit
Target/Decoy Approach: Normal Target/Decoy	~
Enzyme: Trypsin	~
Fragmentation: CID/HCD (B/Y)	~
Precursor Mass Tolerance: 10.0 PPM	~
Fragment Mass Tolerance: 10.0 PPM	~
Library Mass Tolerance: 10.0 PPM	~
Percolator Version: v3-01	$\sim$
Number of Cores:	10 韋
Additonal Command Line Options:	

C. Then queue up the [2022\_09\_01\_HeLa\_500ng\_16mzst\_DDA\_NCE\_27.mzML] mzML file by clicking the "Add MZML" button."

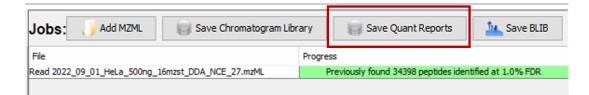
Jobs: 🤳 Add MZML	Save Chromatogram Library	Save Quant Reports	La Save BLIB
File	Progre	55	

Processing should begin immediately. Detailed analysis information will be printed in the console portion of the screen, and a blue bar will move across as the job is processed:

File	Progress
Read 2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27.mzML	#29.2 to 431.2 m/z (1272 MS/MS, 25905 total entries). Jobs in v

Scribe will complete several steps in order: conversion of the mzML to DIA format, library search scoring, Percolator filtering, retention time alignment, Refined Percolator filtering, and then quantification. The resulting search should detect and quantify around 34,398 unique peptides, matching the manuscript result from Figure 3. The specific number may change by a small amount depending on your operating system because of sort ordering in Percolator.

D. Once the file for this library has concluded running, the file bar will turn green. Although there is only one file, and match-between-runs quantitation will not be performed, we can click "Save Quant Reports."



Once the quantitation has concluded, the screen will look the same as the screenshot below. At this point, you should exit from Scribe, and re-open a new window before running another search.

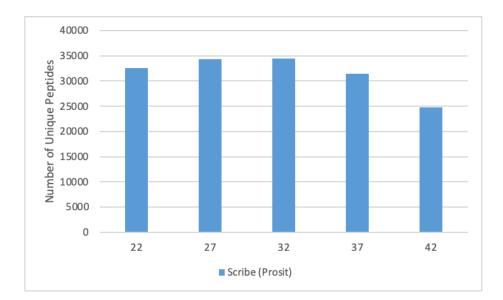
File	Progress
Read 2022_09_01_HeLa_500ng_16mzst_DDA_NCE_27.mzML	Previously found 34398 peptides identified at 1.0% FDR
Write Library quantreport_DDA_NCE27_trypsin.elib	34398 peptides identified at 1.0% FDR

E. To process the other files found in the repository, consult the table below. DLIB Libraries are paired to HeLa mzMLs based on experimental NCE and Prosit NCE, where NCEs are tuned to the specific instrument. Here are the pairings for this experiment in Figure 3 and Figure 4:

Experimental NCE	Experimental File (mzML)	Library File (DLIB)
22	2022_09_01_HeLa_500ng_16mzst_ DDA_NCE_22.mzML	uniprot_human-reference_reviewed_2022mar02.pro sit_input.trypsin_nce16_hcd2020.dlib
27	2022_09_01_HeLa_500ng_16mzst_ DDA_NCE_27.mzML	uniprot_human-reference_reviewed_2022mar02.pro sit_input.trypsin_nce29_hcd2020.dlib
32	2022_09_01_HeLa_500ng_16mzst_ DDA_NCE_32.mzML	uniprot_human-reference_reviewed_2022mar02.pro sit_input.trypsin_nce34_hcd2020.dlib
37	2022_09_01_HeLa_500ng_16mzst_ DDA_NCE_37.mzML	uniprot_human-reference_reviewed_2022mar02.pro sit_input.trypsin_nce41_hcd2020.dlib
42	2022_09_01_HeLa_500ng_16mzst_ DDA_NCE_42.mzML	uniprot_human-reference_reviewed_2022mar02.pro sit_input.trypsin_nce50_hcd2020.dlib

If you run into error messages or Scribe shuts down, please check your version of Java. Ideally, you will be using Java version 1.8. We know of errors that were introduced with Java versions 17 and 18 and we recommend using Java 16 or earlier to ensure stability.

F. Plot the number of peptides identified in each run in any downstream software (e.g., R or Excel). The resulting graph should be similar to the figure below.

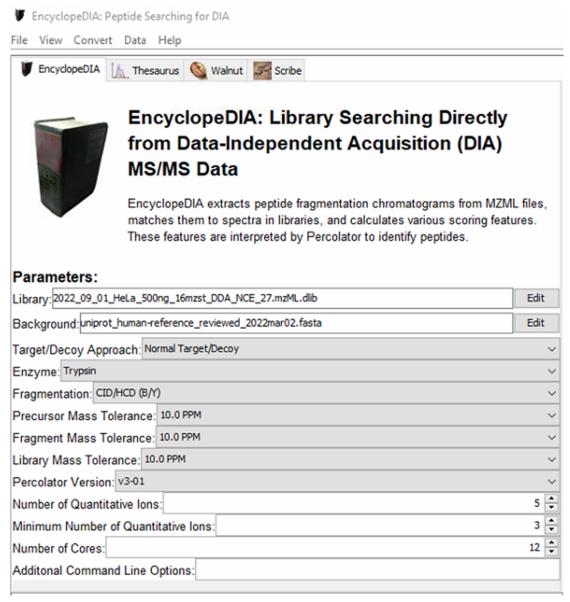


### 8. USING SCRIBE TO BUILD LIBRARIES FOR DIA SEARCHES

How to use spectral libraries made from DDA to search against DIA experiments

A spectral library, .DLIB, file is output from Scribe. With the latest version of EncyclopeDIA, you can use the acquired spectral library to search DIA injections. In this example, we are using the collision energy optimized library acquired from a HeLa DDA injection with NCE = 27. Open EncyclopeDIA, and upload "2022\_09\_01\_HeLa\_500ng\_16mzst\_DDA\_NCE\_27.DLIB." Specify

"uniprot\_human-reference\_reviewed\_2022mar02.fasta" as the background file. The settings should match the screenshot below.



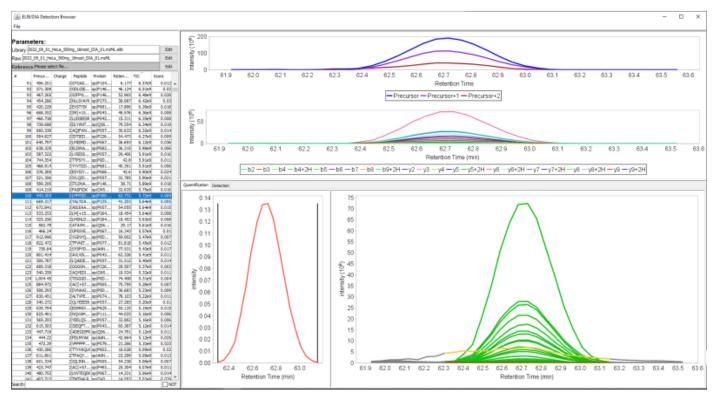
#### Click "Add mzML," then find the DIA injection named "2022\_09\_01\_HeLa\_500ng\_16mzst\_DIA\_01.mzML" to begin the search.

EncyclopeDIA: Pept	tide Searching for DIA									- 🗆 ×		
File View Convert	Data Help					-						
EncyclopeDIA	_ Thesaurus    Wali	nut 🌌 Scribe			Jobs	Add MZML	Save Chromatogram	Library	Save Quant Reports	kave BLIB		
			•	i	File		•	Progress				
			-	hing Directly								
	from Data-I	ndepend	lent Acqu	isition (DIA)								
	MS/MS Dat	V Select a RAV	W file									×
	EncyclopeDIA extr matches them to s	Look in			· 🗿 🛛	• 🗈 😢						
	These features are	<b>_</b>	Name	~		Date	Туре	Size	Tags			
			2022_09_01	_HeLa_500ng_16mzst_DIA_01.rav	v	9/2/2022 11:55 PM	RAW File	2,811,59	КВ			
Parameters:		Quick access		_HeLa_500ng_16mzst_DIA_01.mz	:ML	12/18/2022 6:57 PM	MZML File	1,723,300	KB			
Library: 2022_09_01_He	eLa_500ng_16mzst_DD/		Scribe_mar			12/11/2022 4:40 PM	File folder					
Background:uniprot_h	uman-reference_review	Desktop	p_falcariun			12/13/2022 8:01 PM	File folder File folder					
Target/Decoy Approa			Lysis_burre	r_scribe_encyclopeDIA_comparis	ion	12/11/2022 11:14 PM	File folder					
Enzyme: Trypsin	CIT. Normal Target/Dec											
	co (0.60	Libraries										
Fragmentation: CID/H												
Precursor Mass Tole												
Fragment Mass Toler		This PC										
Library Mass Toleran		<b>1</b>										
Percolator Version: v		Network										
Number of Quantitativ												
Minimum Number of	Quantitative lons:											
Number of Cores:												
Additonal Command	Line Options:											
Console:												
EncyclopeDIA Graphical	Interface (version 2.12.0											
			Object name:	2022_09_01_HeLa_500ng_16m	zst_DIA_0	1.mzML					~	Open
			Objects of type:	All Files (*.*)							~	Cancel
	L											_
				34 of 16312 MB us	ed							
•												

#### Once the search is complete, the detected number of peptides should be "31,349."

				_		×
Jobs:	Add MZML	Save Chromatogram Library	Save Quant Reports		ta Save E	BLIB
File		Progress				
Read 2022_09_01_HeLa_500ng_16mzst_DIA_01.mzML	Wrote 31439 peptides identified at 1.0% FDR					

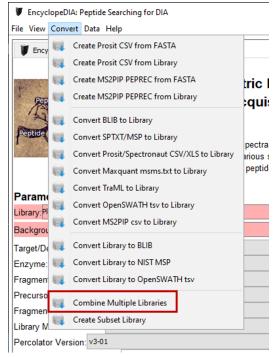
You can then visualize detected peptides as chromatograms using the "ELIB Browser." The peptide below is an example of a peptide that was detected with little interference and multiple fragment ions.



### 9. FAQ

A. Can multiple organisms be searched at once?

Yes, you can by generating a Prosit library for each organism using the method outlined in part 2, then using Encyclopedia/Scribe's "Combine Multiple Libraries" feature to merge both DLIBs. With this method, both organisms can be searched at once. The specified FASTA file contains both organisms. The "Combine Multiple Libraries" feature is outlined below:



Click on "Add Files or Drag Into Box" to add files to the queue. Then select "Library" at the bottom of the window to name the merged library. Click okay to start the process.

V Combine Libraries	× 🗸 Combine Libraries		×	
Add Files or Drag Into Box	Parameters:			
-		Add Files or Drag Into Box		
Library File (.dlib or .elb)	Lbrary File (.dlb or .elb)			
Library: Please select file	Edit Library: Please select file		Edit	
✓ RT align samples		RT align samples		
Remove duplicates		Remove duplicates		
Higher scores are better		Higher scores are better		
OK Cancel		OK Cancel		

#### B. When are additional decoys necessary?

Scribe can search for additional decoys as entrapment peptides by adjusting the "Target/Decoy Approach". When doing this, Scribe adds additional entrapment peptides that are generated by shuffling the sequence of target peptides. Decoy entrapment peptides (reverse shuffled) are also added. If the percentage of detected entrapment peptides is higher than the filtered FDR percentage, then this can help indicate potential errors that escape target/decoy analysis performed by Percolator.

For more information on entrapment as a tool for estimating confidence in proteomics datasets, we recommend the following paper:

Granholm, V., Fernández Navarro, J., Noble, W.S., Käll, L.: Determining the calibration of confidence estimation procedures for unique peptides in shotgun proteomics. J. Proteom. 80, 123–131 (2013)