

# Ryan Tal Fellers, Joseph Brent Greer, Bryan Patrick Early, AJ van Nispen

## **External User Guide for Performing High Throughput Searches on TDPortal**

Part 1: Transferring Project Files to TDPortal	Step 1
Part 2: TDPortal Page Orientation	Step 2
Part 3: Preparing Raw Files for the TDPortal Search	Step 3 - Step 14
Part 4: Running Searches on TDPortal	Step 15 - Step 18
Part 5: Monitoring Progress of TDPortal Searches	Step 19 - Step 26
Part 6: Generating and Viewing the Top Down Report	Step 27 - Step 28



#### **Part 1: Transferring Project Files to TDPortal**

1. Log in to TDPortal at <u>https://portal.nrtdp.northwestern.edu</u> with your Northwestern email and password.

#### Part 2: TDPortal Page Orientation

TDPortal	Analyze Data Monthlee Dated Inter- Vacaditation- Inter- Date-	III Uning 200.1 CR
Tools Panel Welcome to the National Resource The NRTDP TDPortal Global Known Issues The NRTDP TDPortal Global Known Issues The NRTDP Will address urgent in	Six subheadings Six subheadings Translational and Developmental Proteomics (NRTDP) TDPortal for performance and beyond state-of-the-art in mass spectrometry-based proteometry-based proteometry	Interning Top Down

- a. There are six subheadings that link to different pages.
- b. The History Panel is on the right. When data is uploaded or an analysis is performed, each output generates a dataset. These datasets are stored by TDPortal in the History Panel.
- c. All of the available tools are located in the blue Tools Panel on the left.
- d. The NRTDP designed tools are located in the NRTPD Tools menu in the Tools Panel.

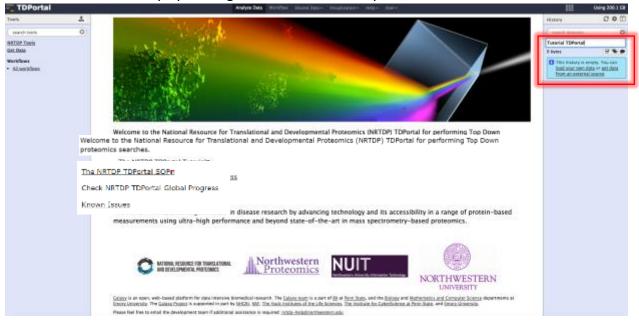


#### Part 3: Preparing Raw files for the TDPortal Search

3. Create a new history. Go to the history column and click on the gear in the top right A drop down menu will appear. Click on "Create New."



4. Rename the new history by clicking on "Unnamed History."





5. Go to Tools Panel and click on "NRTDP Tools." A menu containing all the available NRTDP tools will appear.

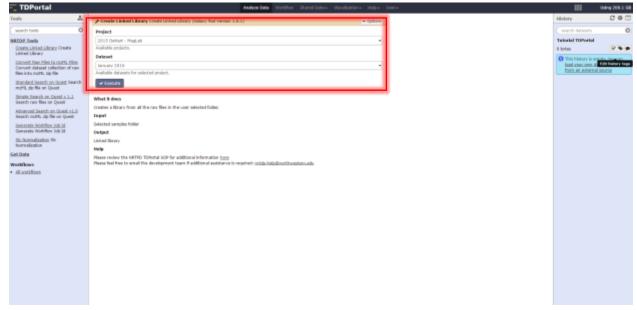


#### 6. Click on "Create Linked Library" to create a linked library.





7. Select the dataset that will be searched. Click "Execute."



8. Click on the "Shared Data" subheading and <u>"Data Libraries</u>" view available linked library.

🔤 Galaxy	Analyze Data Workflor 9	Narod Data- Ve dipation - Help - User-		Using 178.2 GB
Tesic 1			Hictory	00
( search tools 0)	1 job has been successfully added to the queue - resulting in the following and to Library		search datasets	0
NRTDP Tools Create Linked Library Create	You can check the status of guessed jobs and view the resulting data by refre		2015 Dec DeHart - Lisheum	
Unked Ubrary Convert Ray Files to milML Files		Published Vousilizations Published Pages	0 bytes	12 <b>9 9</b>
Convert dataset collection of raw files into math, sip file		r somer not i nay to	L: Linked Library	● / ×
Standard Search on Quest Search mzHL zip file on Quest				
Advanced Search on Quest Search mgHL zip file on Quest				
Generata Workflow Job Id Generata Workflow Job Id				
Get.Deta				
Workflows     Sil workflows				

9. The Data Libraries screen displays available data libraries. Click on the desired data set library

- Galaxy		Analyze Data	Workflow	Shared Data -	Visualization +	Help+	User +	 Using 0 bytes
Data Libraries								
search dataset name, info, message, dbkey Advanced Search	Q							
Data library name;					Dat	Brary	description	
2015 Detait - Hadab December 2015								



hing 178.2 GB

Using 179.2 GB

10. The next screen displays the dataset raw files. Select files to import to history by checking their checkboxes. At the bottom of the page there is a "For selected datasets" dropdown menu. Set the dropdown next to "Import to current history" and click "Go."

Name	Hennage	Data type	Date uploaded	file size
ELGA_CD_DLD1par_10par_f1_freehd/01/_D9Agradientx9_1000A_0um_PLRPS_alt_param.raw *		RAM	Thu Jan 7 18:10:38 2016 (UTC)	585.8 HB
LCA_CD_DLD1psr_10ptr_12_fresh5i3X_3PAprotient/9_1000A_8um_PLRP5_sN_parom.row •		RAM	Thu Jan 7 18:10:40 2016 (UTC)	\$77.3 MB
LCA_CD_DLDiper_10per_10_freehdi0X_IPAgradientx9_1000A_0um_PLRPS_aR_peram.rev *		RAN	Thu Jan 7 18:10:39 2016 (UTC)	657.0 MB
LCA_CD_DLD1par_10per_14_freshfillX_PMgradientv9_1000A_lhum_PLRP6_alt_param.raw *		RAM	Thu Jan 7 10:10:35 2016 (UTC)	753.7 MB
LCA_CD_DLD1pm_10pm_f5_fmahdi5X_IPAgradientv9_1000A_8um_PLRP5_a8_param.raw +		RAM	Thu Jan 7 18:10:39 2006 (UTC)	853.1 M8
LCA_CD_DLDtpar_10per_16_freehd/SX_3PAgradientx9_1000A_0um_PLRP6_aR_param.raw *		RAW	Thu Jan 7 18:10:37 2016 (UTC)	984.8 MB
LCA_CD_DLD1psr_11psr_17_freshdiSX_3PAgradientv9_1000A_8um_PLRP5_aN_param.row •		RAM	Thu Jan 7 18:10:40 2016 (UTC)	958.2 M
LCA_CD_DLDipar_10per_10_freehdISX_IPAgradientx9_1000A_0um_PLRPS_aR_param.raw *		RAN	Thu Jan 7 18:10:41 2016 (UTC)	1.1.08
LCA_CD_DLD1pw_Npw_f1_fresh6i311_0Mgradientv9_1000A_Num_PURPS_alt_param.raw *		RAM	Thu Jan 7 10:10:34 2016 (UTC)	760.4 M
LCA_CD_DLD1pw_Eper_f1_freehd111K_IPAgradientv9_1000A_8um_PLRPS_aR_param_025o_ray *		RAN	Thu Jan 7 18:10:33 2006 (UTC)	795.0 M
LCA_CD_DLDtpar_Sper_f1_freehditor_3PAgradientx9_1000A_Sum_PLRPS_aR_param_366HS1.raw *		RAW	Thu Jan 7 18:10:40 2016 (UTC)	730.3 M
LCA_CD_DLD1pm_Rpm_f2_freshdi3X_3PApredientx9_10084_dum_PURP5_skt_param.raw •		RAM	Thu Jan 7 18:10:33 2016 (UTC)	1.0 GB
LCA_CD_DLD1par_8par_12_freehd13K_3PAgradientx9_1008A_0um_PLRP5_elt_param_825q.raw =		RAN	Thu Jan 7 18:10:36 2016 (UTC)	1.1 08
LCA_CD_DLD1par_Rper_f2_freshill3t_3PAgradientv9_10004_0um_PUPP5_ait_param_386P80_raw +		RAM	Thu Jan 7 10:10:42 2016 (UTC)	1021.1
LCA_CD_DLD1pw_8pw_95_freehd13K_394gredientx9_10084_6um_PLRP5_elt_paramuraw •		RAN	Thu Jan 7 18:10:38 2006 (UTC)	1.0 08
LCA_CD_DLD1par_Epar_F1_freehd135_3Higradeetx4_10064_dum_F1RP5_s4t_param_0256q.raw *		RAW	Thu Jan 7 18:10:34 2016 (UTC)	1.1 68
LCA_CD_DLD1psr_8psr_61_freshdi3X_3Ngradientx6_1000A_8um_PLRP5_dt_param_355P51_attempt3.rax •		RAM	Thu Jan 7 18:10:42 2016 (UTC)	997.9 M
LCA_CD_DLDipar_Sper_M_freehd13X_3Mgradientx0_1006A_0um_FLRMS_sHt_paraminax *		RAW	Thu Jan 7 18:10:37 2016 (UTC)	1.1 08
LCA_CD_DLD1par_Rper_F4_fresh6i3x_3P4gradientv5_10004_0um_PUPP5_Ait_param_6250_raw *		RAII	Thu Jan 7 10:10:42 2016 (UTC)	1.1 GB
LCL CD Di Diem Rem M. Anniel CD. (Manufactur), 10004, damp <sup>1</sup> 1095_alt_param_305953.nex •		RAW	Thu Jan 7 18:10:36 2036 (UTC)	1.1 08

T3P: Several compression options are available for downloading multiple library datasets simultaneously:

- Galaxy

- 11. Once the files are successfully imported this confirmation will be displayed at the top of the page.

		,
Data Library "2015 DeHart - MagLab:December 2015"	Add datasets Add folder	Library Actions
20 30 datasets imported into 1 history; 2015 Dec DeHart - MagLab		

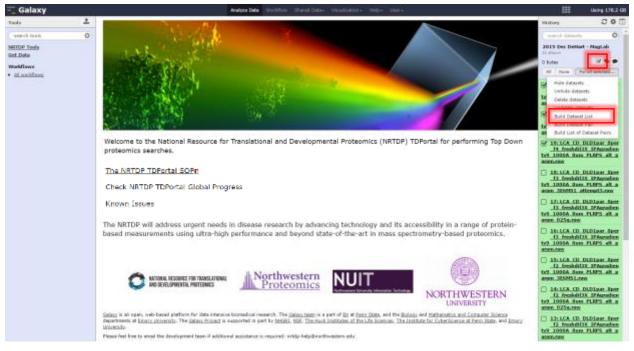


12. Click the Galaxy icon at the top of the page to return to the home page. The history has been populated with a green box for each imported raw file.





- 13. Next create the dataset list. To do this, navigate to the History Panel.
  - a. Click on the checkbox.
  - b. Select the raw files that will be searched.
  - c. Click For all selected
  - d. Select Build Dataset List





14. A new dataset list will appear in the History Panel.



### Step 4: Running searches on TDPortal

15. Navigate to Shared Data. Click on "Published Workflows" in the drop down menu. Import Published Standard Search Workflow 2.0 by clicking on the down arrow of the "Standard Search Workflow Version 2.0" tab.

– Galaxy	Analyze Data	Workflow	Shared Data -	Visualization +	Admin	Help∓	User≠		Using 210.3 GB
Published Workflows									
search name, annotation, owner, and tags	Q								
Advanced Search									
Name		Annot	ation <u>O</u>	wner	Commun	ity Rating	2 Comm	unity Tags	Last Updated
Standard Search Workflow Version 2.0			oi,	iseph-greer	***	kk -			~21 hours ago

16. When the workflow has successfully imported, this screen will appear.

Workflow "Standard Search Workflow Version 2.0" has been imported. You can <u>start using this workflow</u> or <u>return to the previous page</u> .
You can start using this workflow or return to the previous page.



17. Click the "Workflow" Subheading. Click "Run" in Standard Search Workflow Version 2.0 dropdown.

our workflows				Cres	ate new workflow	v 🕆 Upload o	or import workflow
lame						# of Steps	
imported: Standard Search Workflow 1	/ersion 2.0 -					2	
imported: Standard Search Workflow	Run					5	
imported: Standard Search Workflow	Share or Publish Download or Expo					5	
imported: Standard Search Workflow	Сору	`				5	
imported: Standard Search Workflow	Rename View					5	
imported: Old Standard Search Work	Delete					5	
imported: Old Standard Search Workfl	ow Version 1.0.1 •					5	

- 18. Input search parameters and click "Run Workflow."
  - a. Select Dataset list.
  - b. Input Parameters File (optional).
  - c. Select Organism to be searched.
  - d. If performing quant, select True to create SAS Input sheet.
  - e. Select Data file Resolution
  - f. Select Fragmentation Type.
  - g. Select Code Search v2.0.
  - h. Run Workflow.



Running workflow "imported: Standard Search Workflow Version 2.0"	Expand All	Collapse
Step 1: Generate Workflow Job Id (version 1.1.1)		
Step 2: Simple Search on Quest (version 1.1.1)		
Input Dataset List       4: New Dataset List \$		
Job Id		
Output dataset 'out_file1' from step 1 Search Parameters File		
7: Quant SAS Input 8: parameterss.json 10: Quant SAS Input Selection is Optional Organism Homo sapiens (Human) : (Taxon 9606 – July 2016) @ Create SAS Input Sheet for Quant False @		
Resolution High Resolution 🕼		
Fragmentation Type Auto 🕼		
Code Set Search v2.0		



#### **Step 5: Monitoring Progress of TDPortal Searches**

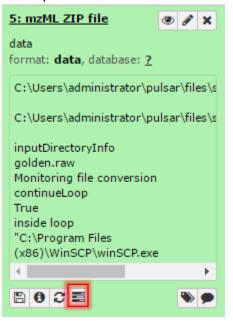
19. An email will be sent once the workflow has started containing history and job id information as well as a link to check progress outside of Galaxy.

G heply G	Reply All 🚱 Forward 😳 IM	
1	Tue 5/26/2015 8/85 PM	
	root <root@localdomain.localdomain></root@localdomain.localdomain>	
1 1	Galaxy Workflow or Job Finished	
to 📕 Suseph 1	test 0 eer	
from: ADM	M@Your.Galaxy.org	
from: ADM	NgWoor Galaxy.org	
from: ADM	NigYour Galaxy.org	
from: ADM	NUEPTour Golden and	
from: <u>ACM</u>	Nu@Proset.Gallery.org	
The follows	ng datasets meeded to be available before sending this notice:	
The follows		

20. The search will be added to History Panel.



21. Monitor Search progress by clicking on Check Progress button under mzML Zip file in the History Panel.





22. In progress, the date of the search, and user email will be visible.

Date	2016-04-25	
User	alexandra.vannispen@northwestern.edu	
Last Raw File Searched		

23. Four percentages are displayed in search progress. The page refreshes every 30 seconds.

Search F	Progress	
Step	Percent Complete	
Convert	a	100%
Prepare	b	100%
Search	c	22%
File	Percent Complete	Percentage of search completed per raw file.
Reporting	d	Queued

- a. The Convert percent is the percentage of raw files that have been converted to mzML.
- b. The Prepare percent is the percent of mzML files that have been prepared for search.
- c. The Search percent is the percent of searching all the mzML that has completed.
- d. The search percent for each mzML file is also displayed.
- e. The Reporting percent is the percentage of the creation of the tdReport that has completed.
- f. The Search Progress also displays errors during search with a descriptive message.



24. TDPortal will send an email immediately if a step in the workflow fails. The email will contain history and job id information.



25. An email will be sent when the search completes.

G MAY G	Reply All (	Forward	90 M					
	Tue 5/26/	2015 8-05 PM						
	root	<root@< th=""><th>localdor</th><th>main.loc</th><th>aldom</th><th>ain&gt;</th><th></th><th></th></root@<>	localdor	main.loc	aldom	ain>		
1 1	Galaxy	Workflow	or Job Fin	ished				
to Excepted	kent Greer							
from: ADM	IN@Your.0	inlawy.org						
from: <u>ADM</u>	IN@Your.G	Can Hansay, Joseph						
The follows	ng dataset:		i be availabk m Report	e before se	nding this r	votice:		

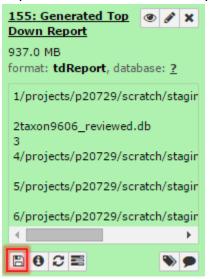
26. The History Panel items will turn green when the search fully completes.

	Realized Online Hard Name I Online - Annalization - A	dmin Melya Usera
High Throughput Galaxy Search Using	g Gwest	
Dete	315-12-88	
Uter	ryan fellen@nathweden.eda	
Last Raw File Searched		
Search Progress		
Seep		Percent Complete
Convert		1 1 94
Prepare		1 1 May 1 1
Seenh		1 1 Mar 1 1
File	Percent Complete	
111223-1112159/255mito-frac3atd	1 1 mm	
Reporting		10 MA



#### **Step 6: Generating and Viewing the Top Down Report**

27. Download search results by clicking "download" button under the Generated Top Down Report subsection in the History Panel.



28. Open results in TDViewer (topdownviewer.northwestern.edu)

•	Setup	Study Design		1% FDR +
		Use this viewer to setup your experiment before data processing begins. This is a crucial step, in order to property nor	e o study, it must be	
Я.	Proteins	Project Factors Use these factors to organize your data files into logical groups.		Search Result Sets Results from multiple search engines will be integrated and displayed together.
	Proteoforms		Add_ =	Add Toxit Tetra Ven Paranetes
8	Compare		Edit	III IS Active Name
1	Statistics		Penove	4646         R         Bio/Nerker           4647         R         Tight Absolute Mass           4648         R         Ind Unexpected Modifications
)	Reports			eole 👻 nos unepectes vicalitations
		Data Files and Categorization The files organized by factors below constitute the observations of your study Add Data Files.	_	
		golden/m2ML		