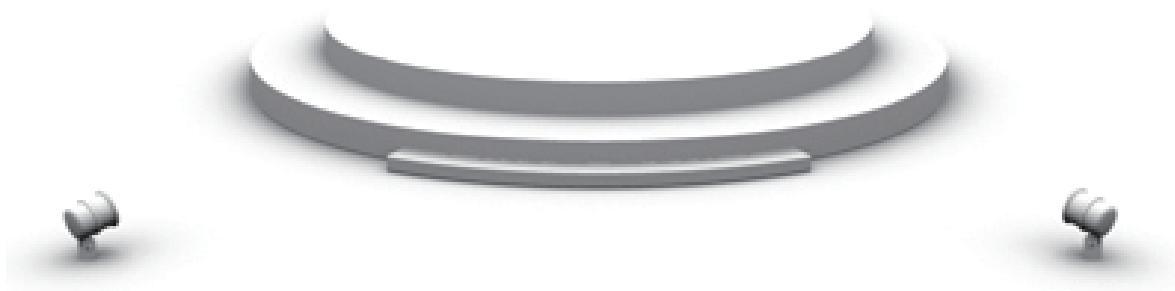


Professional MS Training Class

Proteome Discoverer



ThermoFisher
SCIENTIFIC

目录

软件主界面布局&Study 编辑	01
蛋白质数据库添加	08
可变翻译后修饰添加	09
酶切试剂添加	11
定量方法添加	12
定性分析流程	13
定量分析流程	18
定性分析流程	25
后处理模块	33
定量分析流程	35
结果展示	37
结果导出	48

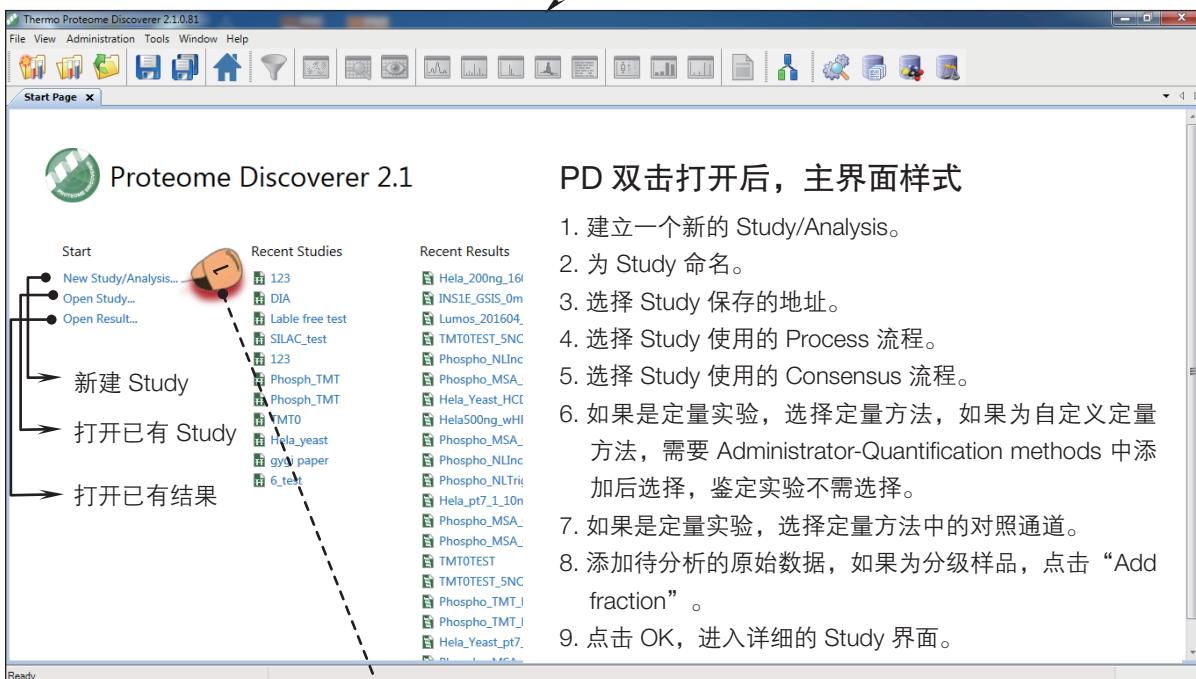
Study&
Administrator ►
编辑

Processing
分析流程 ►

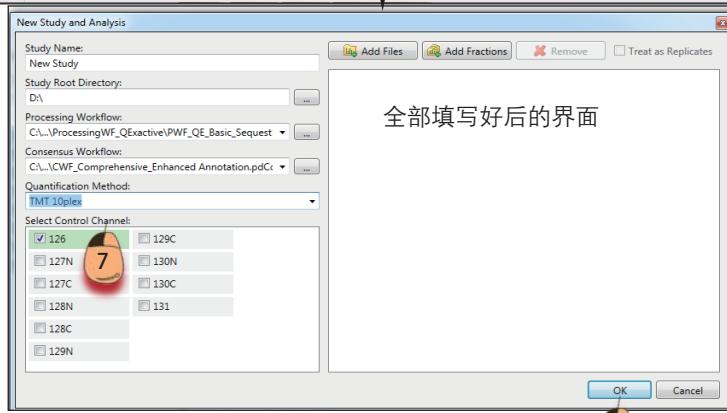
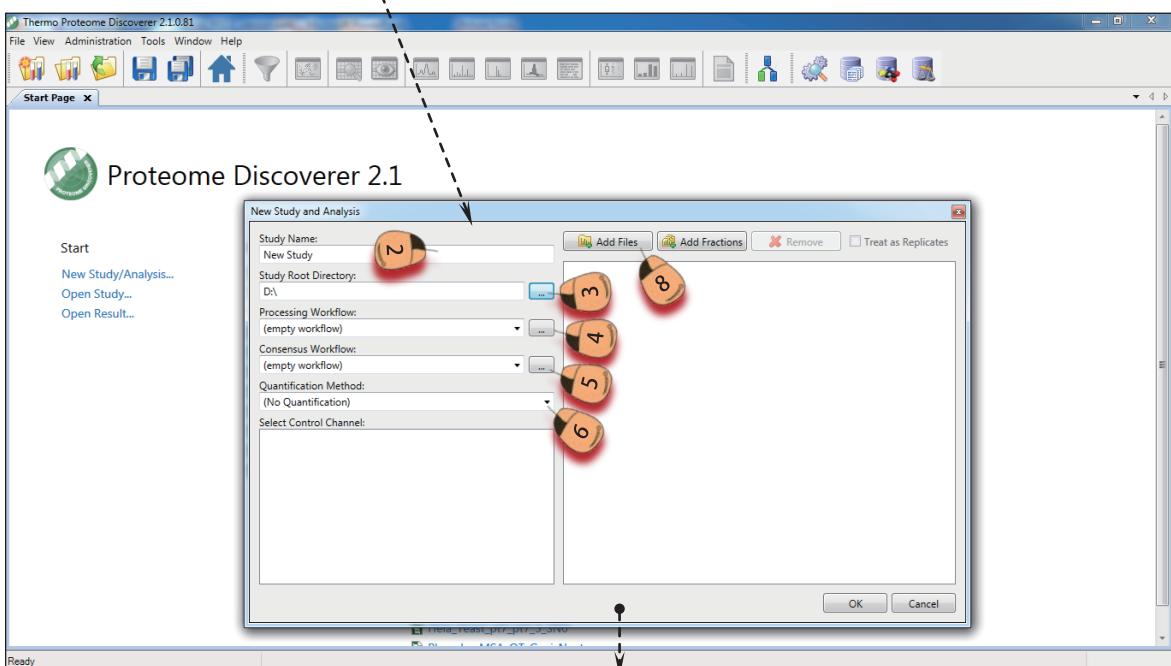
Consensus
分析流程 ►

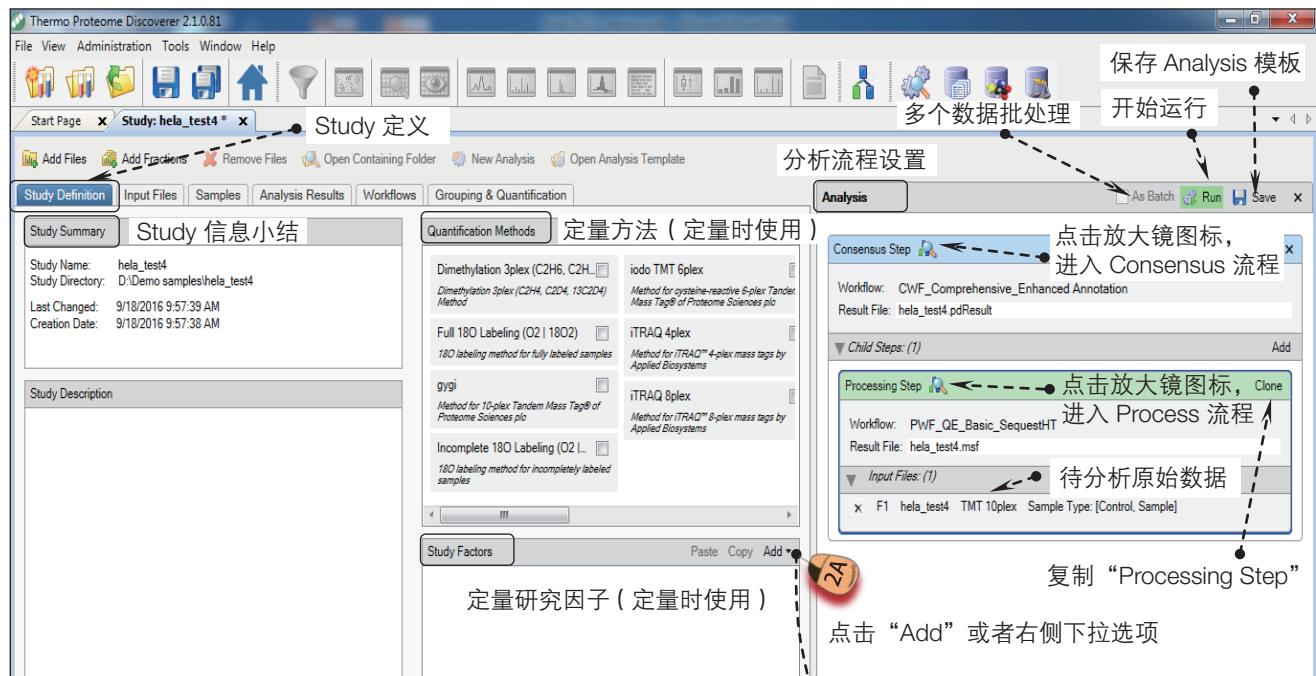
结果展示 ►

双击桌面图标打开软件

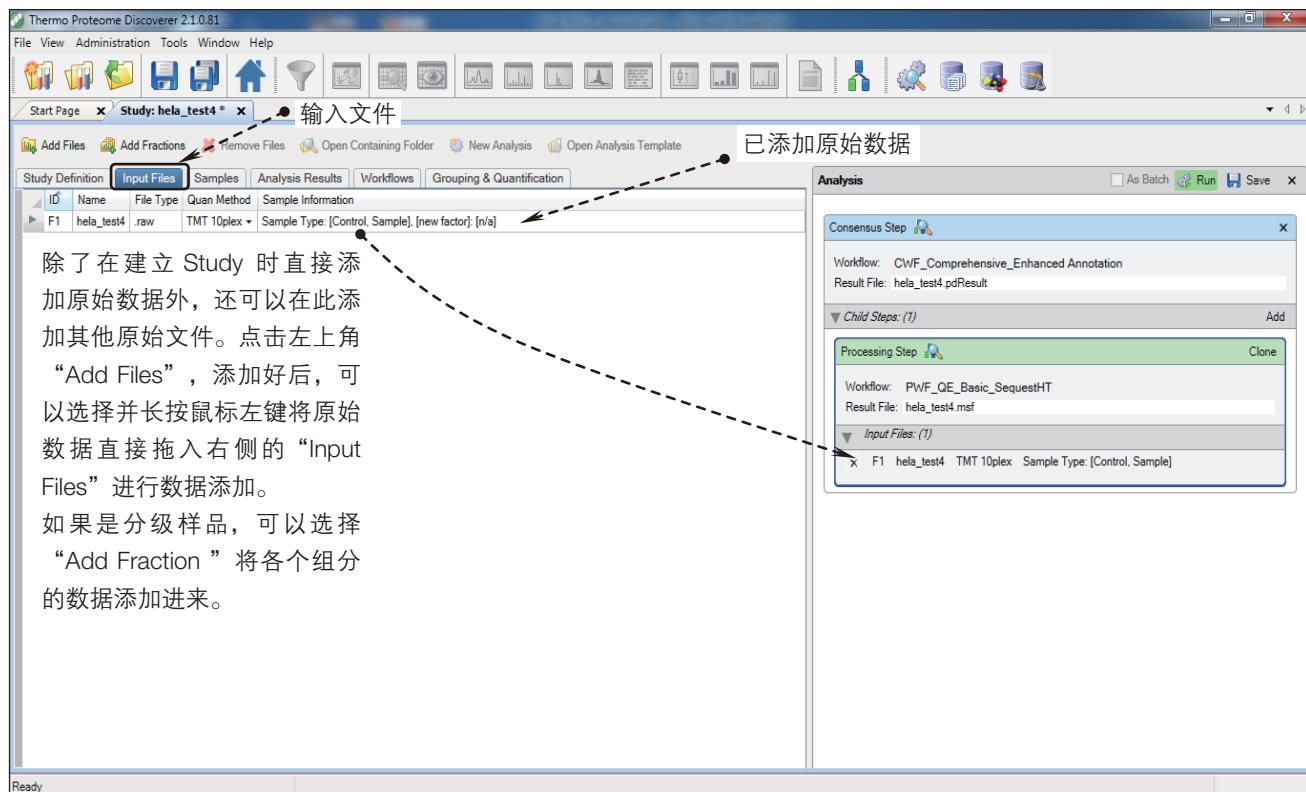


01



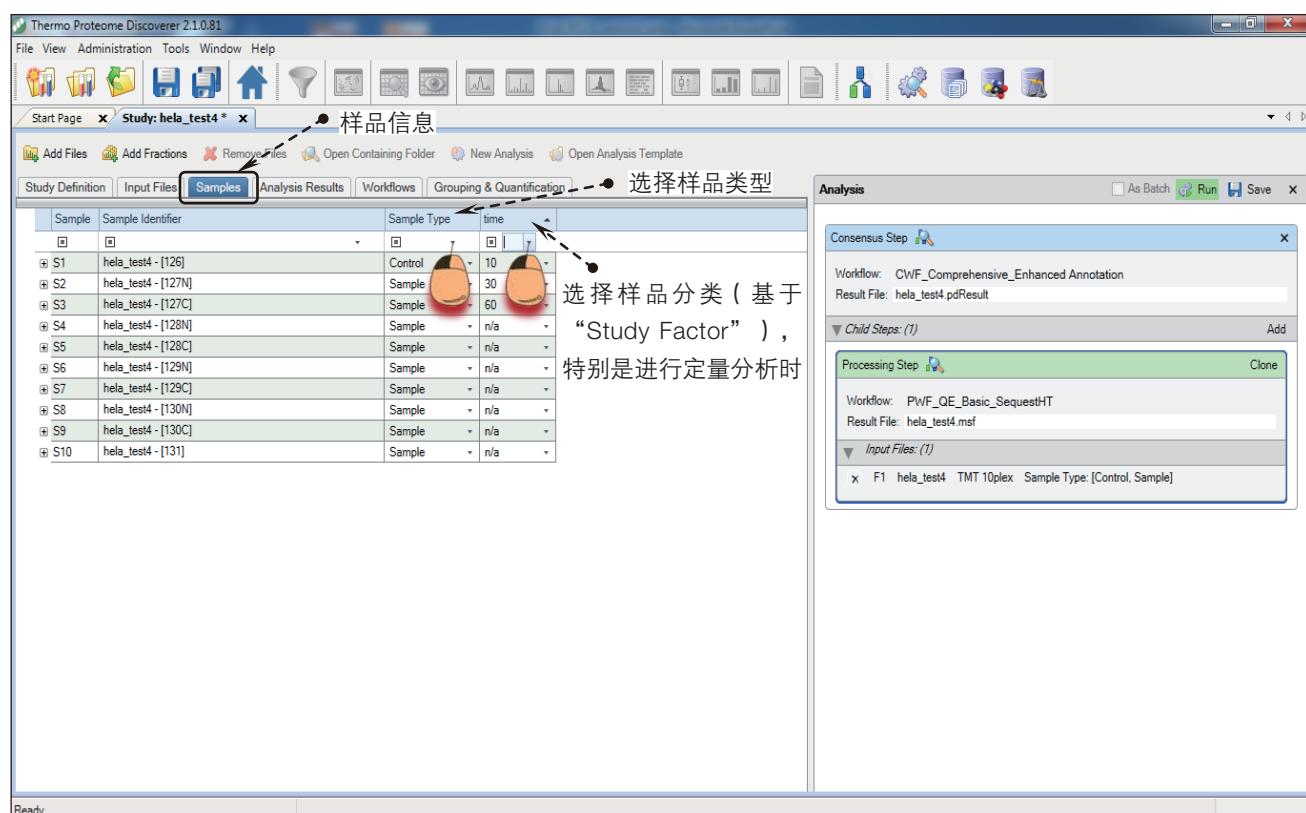


1. 确认 Study 信息和使用的定量方法。
2. 如果为定量实验，添加 Study Factor，具体见上述鼠标指示。
3. 在 Analysis 部分，选择 Processing 和 Consensus 流程。
4. 确认待分析的原始数据，在同时导入多个原始数据的情况下，如需采用相同的 Processing 和 Consensus 流程可以勾选“As Batch”进行批处理，可以得到与原始数据一一对应的多个结果报告，如果不勾选则生成一个结果报告，其中包括了来源于所有原始数据的鉴定和定量信息。
5. 参数确定无误后，可以点击右上角绿色背景的“Run”，开始运行数据处理。
6. 点击“Analysis”模块中右上角的“Save”，保存 Analysis 模板（包含 Processing 和 Consensus 流程），今后分析可直接调用。
7. 点击“Analysis”模块中右上角的“×”，可以关闭“Analysis”页面。
8. 点击“Consensus Step”模块中右上角的“×”，可以删除“Consensus Step”。
9. 点击“Processing Step”模块中右上角的“Clone”，可以进行“Process Step”模块的复制。

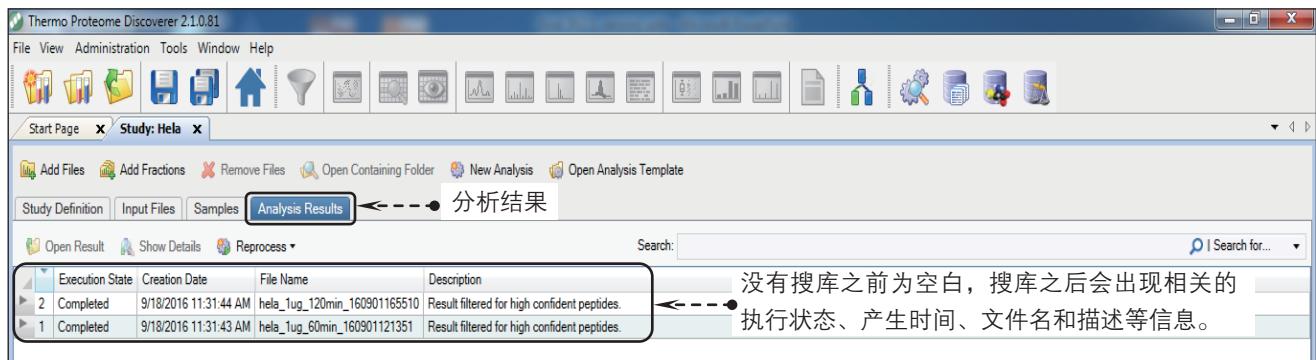


除了在建立 Study 时直接添加原始数据外，还可以在此添加其他原始文件。点击左上角“Add Files”，添加好后，可以选择并长按鼠标左键将原始数据直接拖入右侧的“Input Files”进行数据添加。

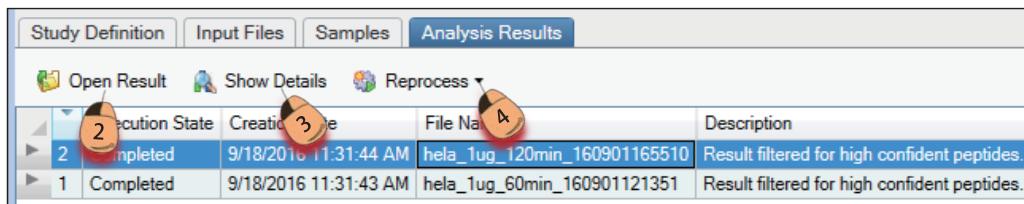
如果是分级样品，可以选择“Add Fraction”将各个组分的数据添加进来。



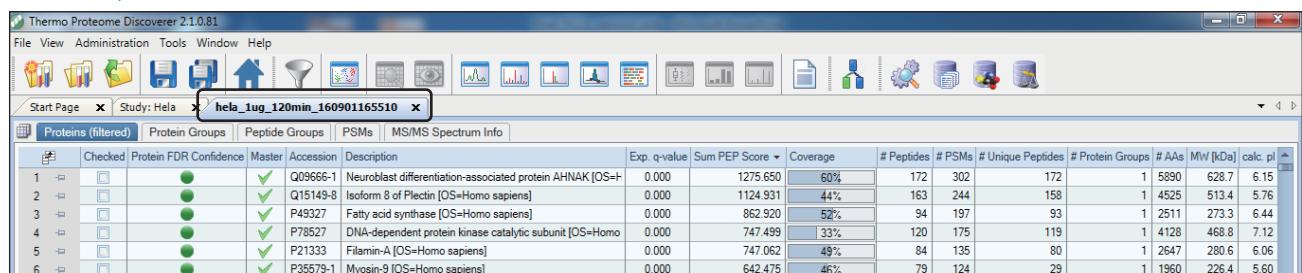
选择样品类型
选择样品分类（基于
“Study Factor”），
特别是进行定量分析时



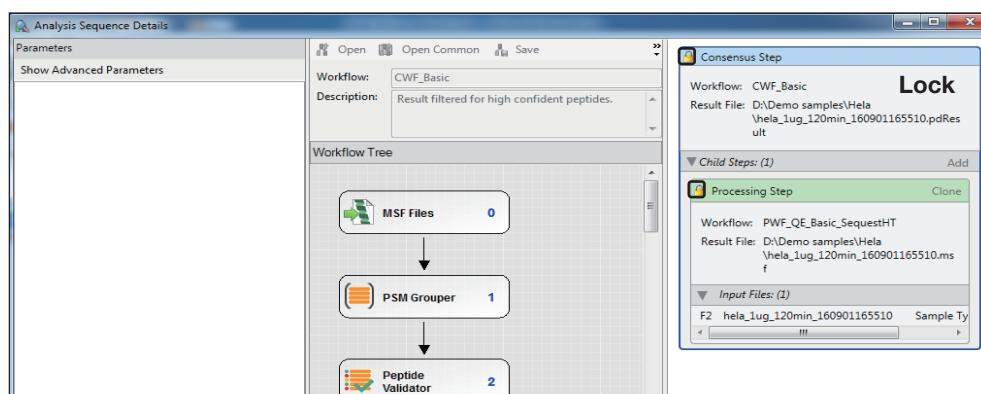
1. 点击待查看的搜库结果。



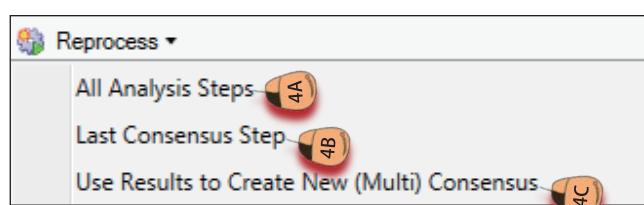
2. 点击“Open Result”按钮,结果将呈现在新的标签页中,有关结果介绍请见“结果展示”部分。

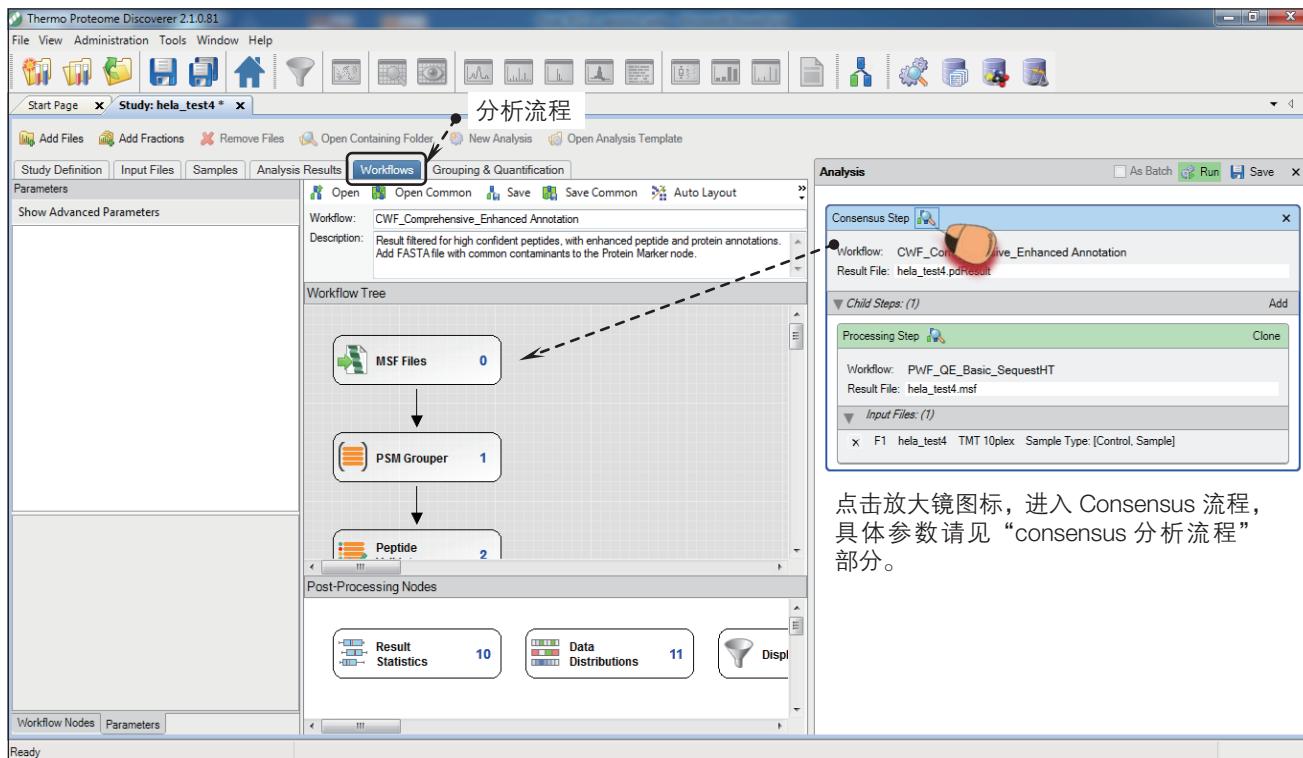


3. 点击“Show Details”按钮,搜库使用的Consensus和Processing流程将显示,此时两个流程均为Lock状态,只可查看不可编辑。

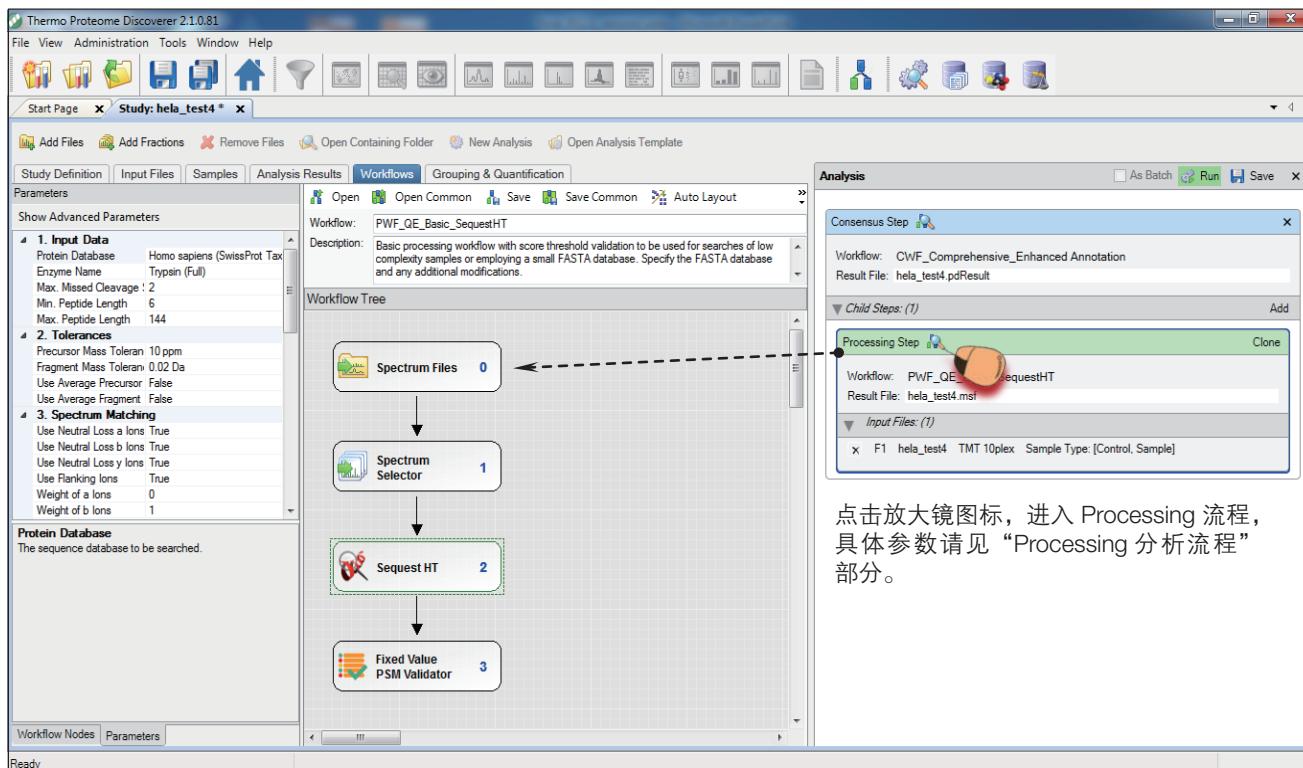


4. 点击“Reprocess”按钮,可以针对原始数据分别进行重新搜索: (4A)分别对Consensus和Processing重新编辑参数并检索; (4B)不进行Processing处理,只进行Consensus重新检索; (4C)可以选择一个或者多个processing后的msf结果文件,只进行Consensus处理,适合将多数据整合到一个结果。(备注: PD 2.1只能打开Consensus处理过的结果文件,如果想把几个结果合并打开,请使用该功能)。

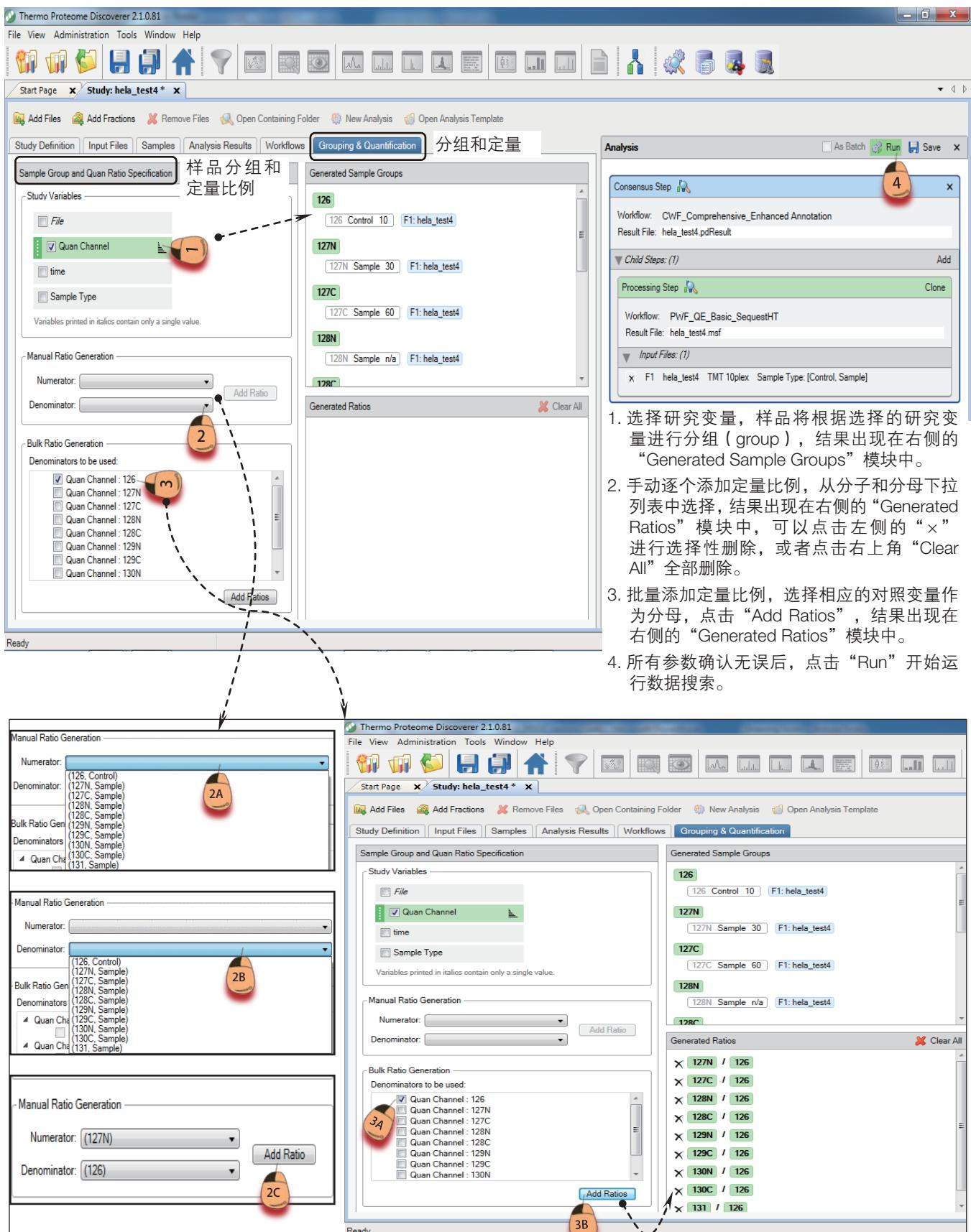


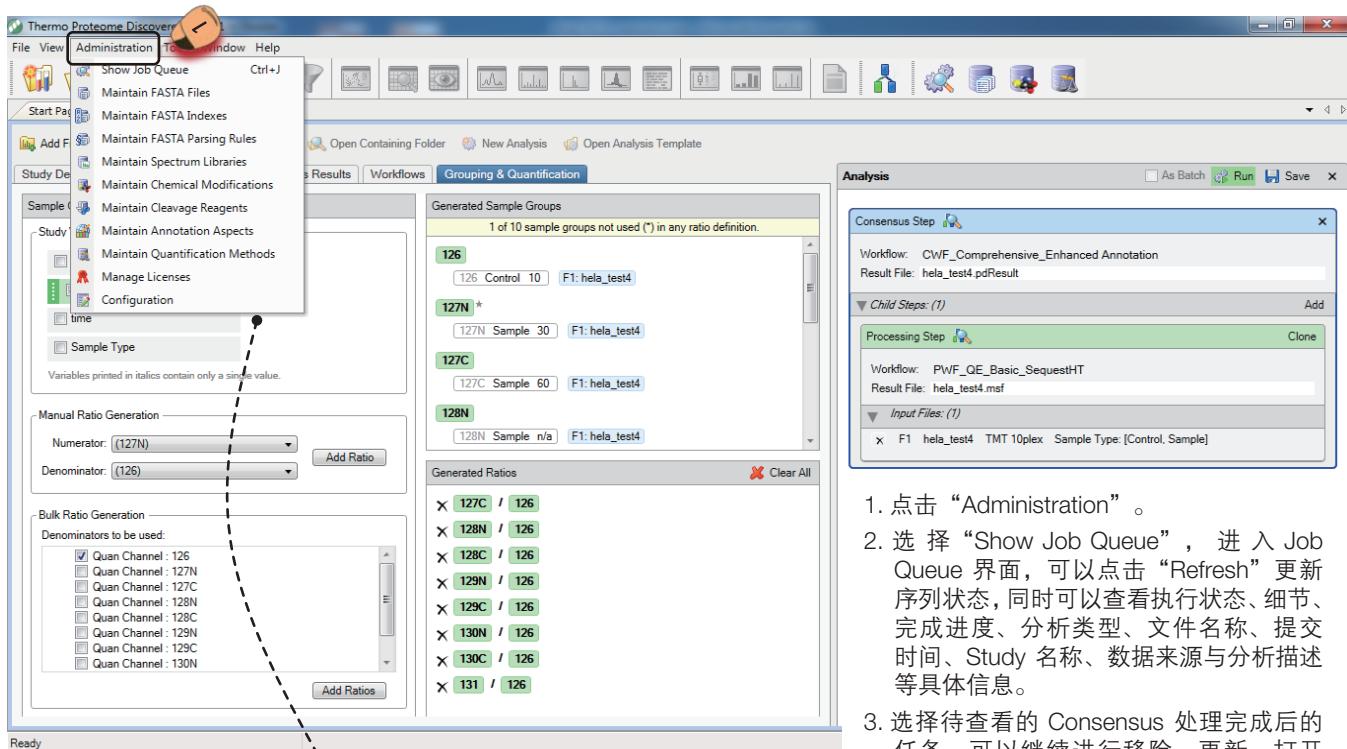


点击放大镜图标，进入 Consensus 流程，
具体参数请见“consensus 分析流程”
部分。

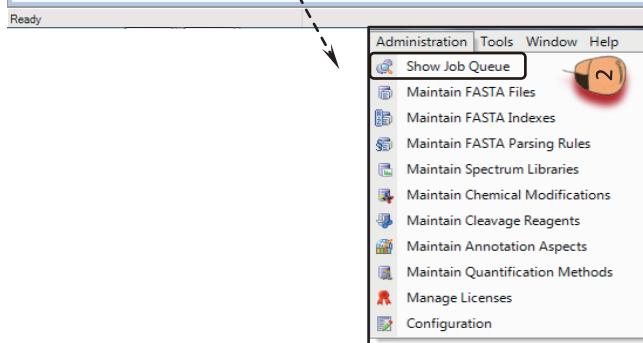


点击放大镜图标，进入 Processing 流程，
具体参数请见“Processing 分析流程”
部分。





1. 点击“Administration”。
2. 选择“Show Job Queue”，进入 Job Queue 界面，可以点击“Refresh”更新序列状态，同时可以查看执行状态、细节、完成进度、分析类型、文件名称、提交时间、Study 名称、数据来源与分析描述等具体信息。
3. 选择待查看的 Consensus 处理完成后的任务，可以继续进行移除、更新、打开结果与打开 Study 等后续操作。



Start Page x Study: Hela x helia_1ug_120min_160901165510 x Administration x

Pause Resume Abort Remove Refresh Open Results Open Study Display Verbose Messages

Job Queue:

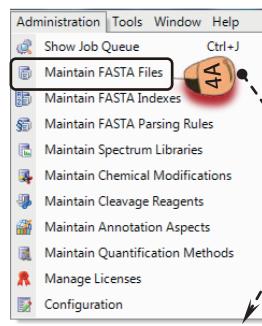
Execution State	Details	Progress	Type	Name	Submitted at	Study	Data Source	Description
Completed	OK	100 %	Consensus	helia_1ug_120min_160901165510	9/18/2016 11:31 AM	Hela	D:\Demo samples\Hela\helia_1ug_12...	Result filtered fo
Completed	OK	100 %	Processing	helia_1ug_120min_160901165510	9/18/2016 11:31 AM	Hela	D:\Demo samples\2016\Training\ika...	Basic processing
Completed	OK	100 %	Consensus	helia_1ug_60min_160901121351	9/18/2016 11:31 AM	Hela	D:\Demo samples\Hela\helia_1ug_60...	Result filtered fo
Completed	OK	100 %	Processing	helia_1ug_60min_160901121351	9/18/2016 11:31 AM	Hela	D:\Demo samples\2016\Training\ika...	Basic processing

Job Queue:

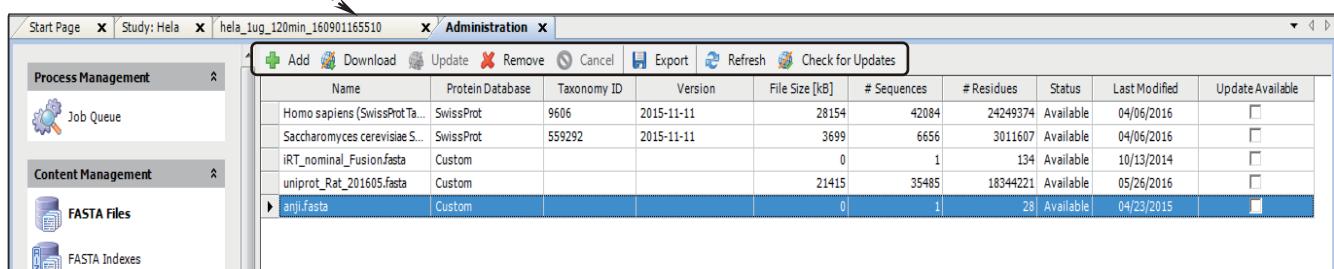
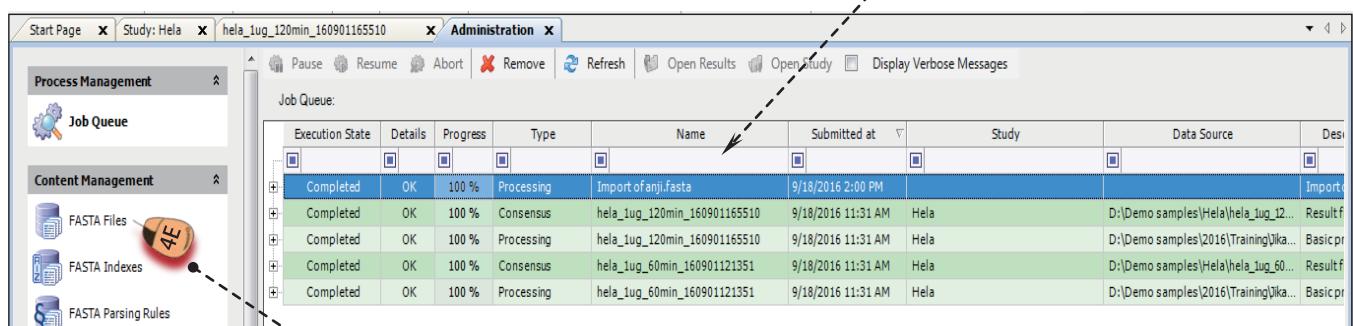
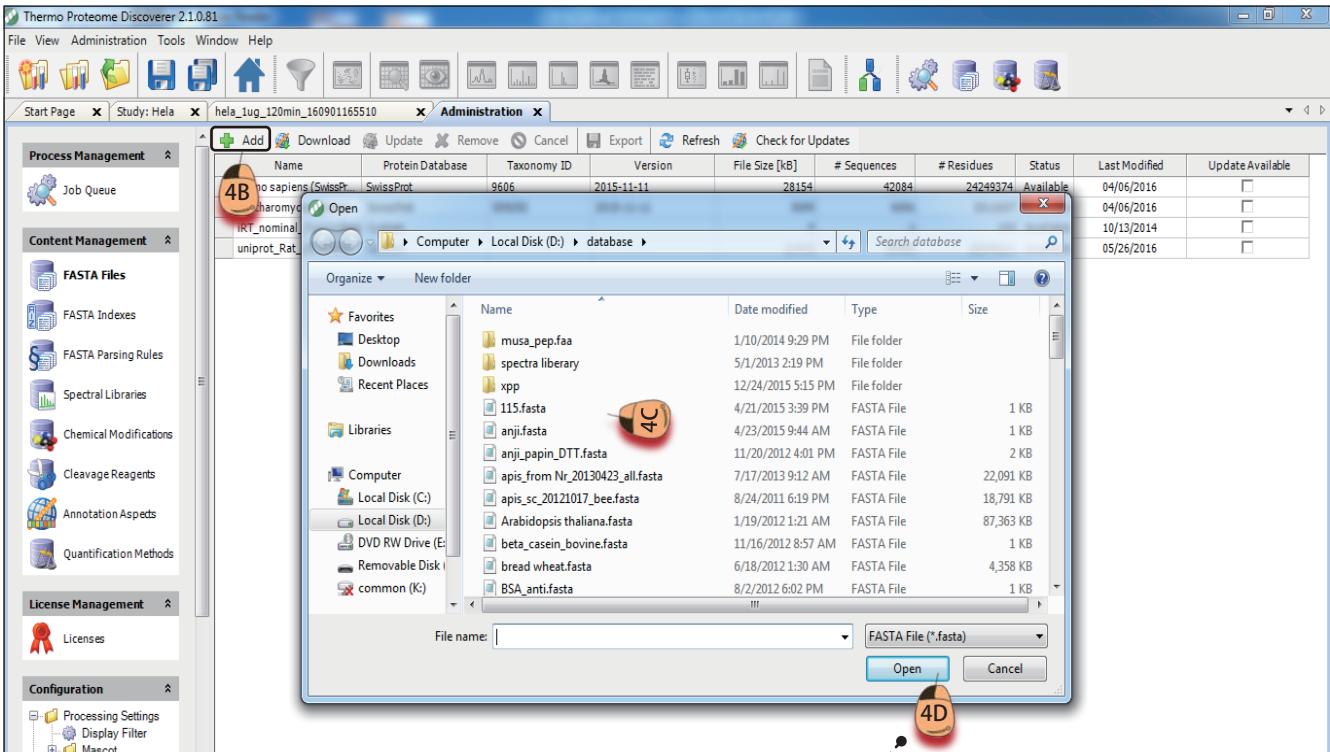
Execution State	Details	Progress	Type	Name	Submitted at	Study	Data Source	Description
Completed	OK	100 %	Consensus	helia_1ug_120min_160901165510	9/18/2016 11:31 AM	Hela	D:\Demo samples\Hela\helia_1ug_12...	Result filtered fo
Completed	OK	100 %	Processing	helia_1ug_120min_160901165510	9/18/2016 11:31 AM	Hela	D:\Demo samples\2016\Training\ika...	Basic processing
Completed	OK	100 %	Consensus	helia_1ug_60min_160901121351	9/18/2016 11:31 AM	Hela	D:\Demo samples\Hela\helia_1ug_60...	Result filtered fo
Completed	OK	100 %	Processing	helia_1ug_60min_160901121351	9/18/2016 11:31 AM	Hela	D:\Demo samples\2016\Training\ika...	Basic processing

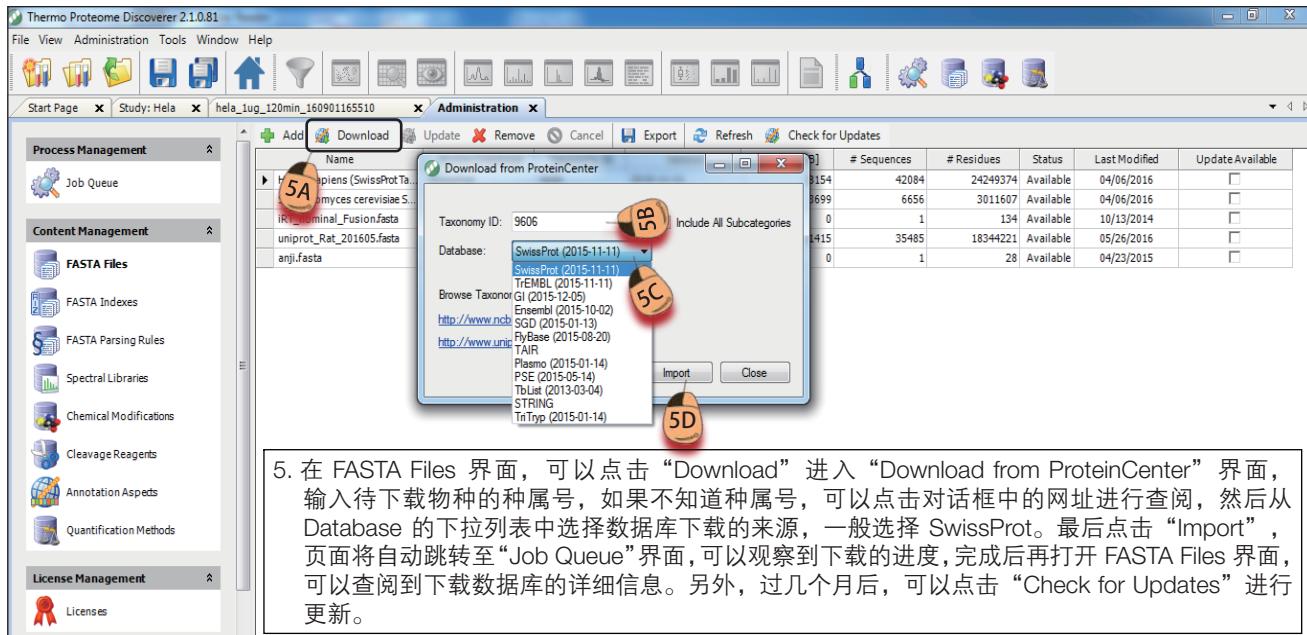


Administration 设置-手工添加已下载 FASTA 数据库



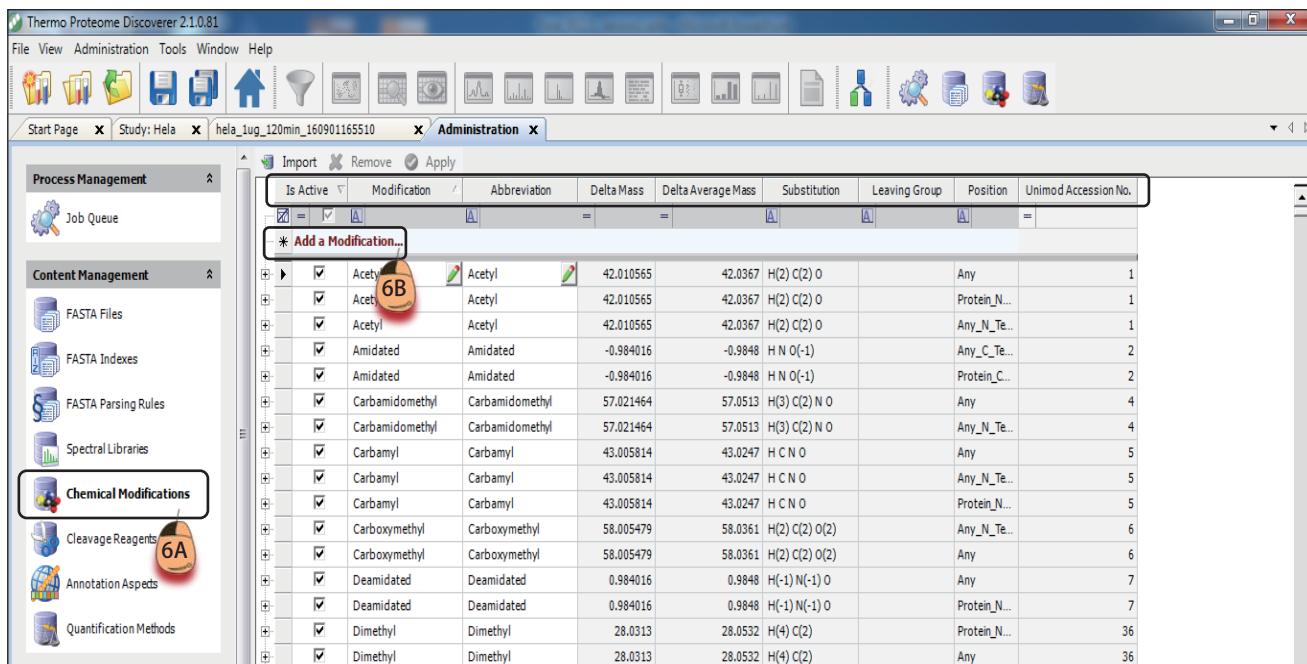
4. 选择“Maintain FASTA Files”，进入 FASTA Files 界面，可以点击“Add”进入“Open”界面，再选择已经下载好的待添加的.fasta 文件，点击“Open”，自动跳转至“Job Queue”界面显示数据库添加状态。当显示数据库添加完成后，点击左侧的“FASTA Files”返回 FASTA Files 界面，一般按照时间顺序排列，可以进一步查看添加数据库的信息，比如数据库的名字、来源、种属号、版本号、大小、包含的氨基酸序列数目、氨基酸残基数目、状态、最后一次修改的时间和是否需要自动更新等信息。另外也可以使用移除、导出等按钮对文件进行操作。





5. 在 FASTA Files 界面，可以点击“Download”进入“Download from ProteinCenter”界面，输入待下载物种的种属号，如果不知道种属号，可以点击对话框中的网址进行查阅，然后从 Database 的下拉列表中选择数据库下载的来源，一般选择 SwissProt。最后点击“Import”，页面将自动跳转至“Job Queue”界面，可以观察到下载的进度，完成后再打开 FASTA Files 界面，可以查阅到下载数据库的详细信息。另外，过几个月后，可以点击“Check for Updates”进行更新。

6. 在 Administration 界面，点击“Chemical Modifications”进入化学修饰自定义界面。当添加新的修饰时，建议点击“Delta Mass”进行升序或者降序排列，然后在待增加的质量数附近进行查找，如果找到相同质量的修饰条目，可以先确认“Is Active”是否已勾选，如果没有可以进行勾选，然后再点击“Apply”，即可在后续的数据分析中进行该修饰的选择；如果找不到相同质量的修饰条目，则需要点击“Add a Modification”，在出现铅笔标志的空白位置进行相应的修饰名称、简写、修饰质量差、修饰平均质量差、化学元素组成、丢失基团（一般不需要填写）、发生位置、Unimod 编号等信息输入；在“Position”区域，可以根据修饰发生的位置进行相应的选择，当选择了“Any”后，鼠标移动至其它区域，则在新加入的修饰之前会出现“+”，点击“+”，再点击出现的“Add an Amino Acid site”，则可以从下拉列表中进行相应氨基酸的选择，同时还可以在“Classification”选项中进行修饰种类的选择。如需删除，选中待删除的条目，再点击“Remove”，确认无误后，点击“Apply”则可以在后续的数据分析中进行选择，另外也可以从 Unimod 网页或者本地文件里直接导入化学修饰。



Import Remove Apply

Is Active	Modification	Abbreviation	Delta Mass	Delta Average Mass	Substitution	Leaving Group	Position	Unimod Accession No.
<input checked="" type="checkbox"/>	Arg-loss	Arg-loss	-156.101111	-156.1857	H(-12) C(-6) N(-...)			1287
<input checked="" type="checkbox"/>	Met-loss	Met-loss	-131.040485	-131.1961	H(-9) C(-5) N(-1...)			765
<input type="checkbox"/>	Trp->Gly	Trp->Gly	-129.057849	-129.1586	H(-7) C(-9) N(-1)			676
<input type="checkbox"/>	Lys-loss	Lys-loss	-128.094963	-128.1723	H(-12) C(-6) N(-...)			313

Any
Any_C_Terminus
Any_N_Terminus
Protein_C_Terminus
Protein_N_Terminus

6C

Import Remove Apply

Is Active	Modification	Abbreviation	Delta Mass	Delta Average Mass	Substitution	Leaving Group	Position	Unimod Accession No.																								
* Add a Modification...																																
<input checked="" type="checkbox"/>	new	new	123.123	123.135				0																								
<table border="1"> <thead> <tr> <th>Amino Acid Name</th> <th>One Letter Code</th> <th>Classification</th> </tr> </thead> <tbody> <tr> <td colspan="3">* Add an Amino Acid Site...</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Arg-loss</td> <td>Arg-loss</td> <td>-156.101111</td> <td>-156.1857</td> <td>H(-12) C(-6) N(-...)</td> <td></td> <td></td> <td>1287</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Met-loss</td> <td>Met-loss</td> <td>-131.040485</td> <td>-131.1961</td> <td>H(-9) C(-5) N(-1...)</td> <td></td> <td></td> <td>765</td> </tr> </tbody> </table>									Amino Acid Name	One Letter Code	Classification	* Add an Amino Acid Site...			<input checked="" type="checkbox"/>	Arg-loss	Arg-loss	-156.101111	-156.1857	H(-12) C(-6) N(-...)			1287	<input checked="" type="checkbox"/>	Met-loss	Met-loss	-131.040485	-131.1961	H(-9) C(-5) N(-1...)			765
Amino Acid Name	One Letter Code	Classification																														
* Add an Amino Acid Site...																																
<input checked="" type="checkbox"/>	Arg-loss	Arg-loss	-156.101111	-156.1857	H(-12) C(-6) N(-...)			1287																								
<input checked="" type="checkbox"/>	Met-loss	Met-loss	-131.040485	-131.1961	H(-9) C(-5) N(-1...)			765																								

6D

Import Remove Apply

Is Active	Modification	Abbreviation	Delta Mass	Delta Average Mass	Substitution	Leaving Group	Position	Unimod Accession No.																																																																														
* Add a Modification...																																																																																						
<input checked="" type="checkbox"/>	new	new	123.123	123.135				0																																																																														
<table border="1"> <thead> <tr> <th>Amino Acid Name</th> <th>One Letter Code</th> <th>Classification</th> </tr> </thead> <tbody> <tr> <td colspan="3">* Add an Amino Acid Site...</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Alanine</td> <td>A</td> <td>-156.101111</td> <td>-156.1857</td> <td>H(-12) C(-6) N(-...)</td> <td></td> <td></td> <td>1287</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Arginine</td> <td>R</td> <td>-131.040485</td> <td>-131.1961</td> <td>H(-9) C(-5) N(-1...)</td> <td></td> <td></td> <td>765</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Asparagine</td> <td>N</td> <td>-129.057849</td> <td>-129.1586</td> <td>H(-7) C(-9) N(-1)</td> <td></td> <td></td> <td>676</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Aspartic Acid</td> <td>D</td> <td>-128.094963</td> <td>-128.1723</td> <td>H(-12) C(-6) N(-...)</td> <td></td> <td></td> <td>313</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Average Asp/Asn</td> <td>B</td> <td>-115.042199</td> <td>-115.132</td> <td>H(-5) C(-8) N(-1)</td> <td></td> <td></td> <td>1224</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Average Glu/Gln</td> <td>Z</td> <td>-106.041865</td> <td>-106.1219</td> <td>H(-6) C(-7) O(-1)</td> <td></td> <td></td> <td>1239</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>C-Terminus</td> <td></td> <td>-99.079647</td> <td>-99.1344</td> <td>H(-9) C(-4) N(-3)</td> <td></td> <td></td> <td>646</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Arg->Gly</td> <td>Arg->Gly</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>									Amino Acid Name	One Letter Code	Classification	* Add an Amino Acid Site...			<input checked="" type="checkbox"/>	Alanine	A	-156.101111	-156.1857	H(-12) C(-6) N(-...)			1287	<input checked="" type="checkbox"/>	Arginine	R	-131.040485	-131.1961	H(-9) C(-5) N(-1...)			765	<input checked="" type="checkbox"/>	Asparagine	N	-129.057849	-129.1586	H(-7) C(-9) N(-1)			676	<input checked="" type="checkbox"/>	Aspartic Acid	D	-128.094963	-128.1723	H(-12) C(-6) N(-...)			313	<input checked="" type="checkbox"/>	Average Asp/Asn	B	-115.042199	-115.132	H(-5) C(-8) N(-1)			1224	<input checked="" type="checkbox"/>	Average Glu/Gln	Z	-106.041865	-106.1219	H(-6) C(-7) O(-1)			1239	<input checked="" type="checkbox"/>	C-Terminus		-99.079647	-99.1344	H(-9) C(-4) N(-3)			646	<input checked="" type="checkbox"/>	Arg->Gly	Arg->Gly						
Amino Acid Name	One Letter Code	Classification																																																																																				
* Add an Amino Acid Site...																																																																																						
<input checked="" type="checkbox"/>	Alanine	A	-156.101111	-156.1857	H(-12) C(-6) N(-...)			1287																																																																														
<input checked="" type="checkbox"/>	Arginine	R	-131.040485	-131.1961	H(-9) C(-5) N(-1...)			765																																																																														
<input checked="" type="checkbox"/>	Asparagine	N	-129.057849	-129.1586	H(-7) C(-9) N(-1)			676																																																																														
<input checked="" type="checkbox"/>	Aspartic Acid	D	-128.094963	-128.1723	H(-12) C(-6) N(-...)			313																																																																														
<input checked="" type="checkbox"/>	Average Asp/Asn	B	-115.042199	-115.132	H(-5) C(-8) N(-1)			1224																																																																														
<input checked="" type="checkbox"/>	Average Glu/Gln	Z	-106.041865	-106.1219	H(-6) C(-7) O(-1)			1239																																																																														
<input checked="" type="checkbox"/>	C-Terminus		-99.079647	-99.1344	H(-9) C(-4) N(-3)			646																																																																														
<input checked="" type="checkbox"/>	Arg->Gly	Arg->Gly																																																																																				

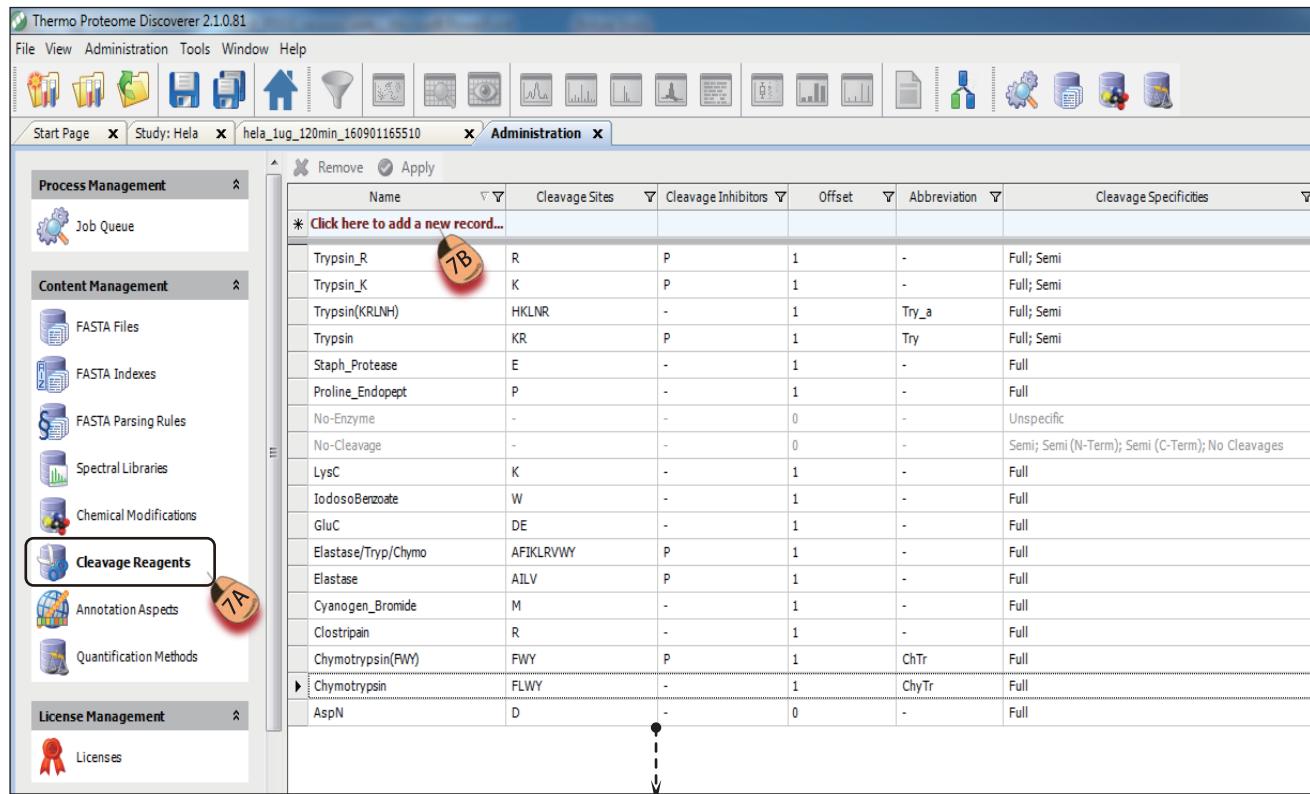
6E

Import Remove Apply

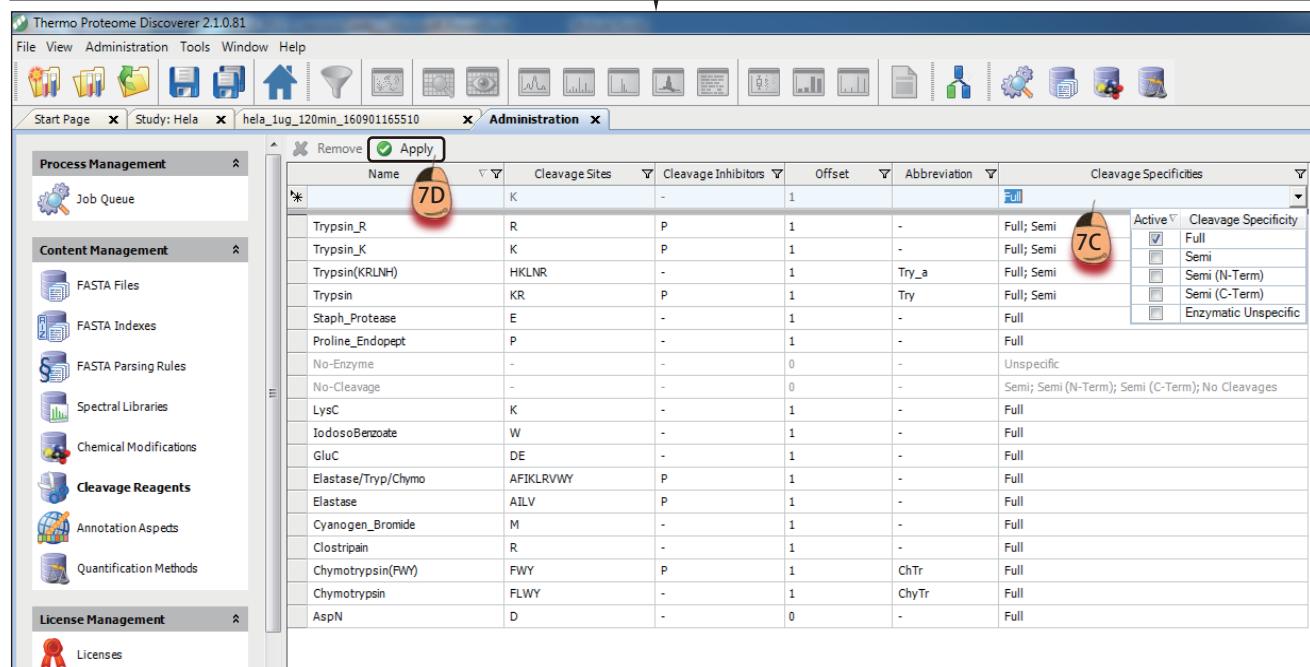
Is Active	Modification	Abbreviation	Delta Mass	Delta Average Mass	Substitution	Leaving Group	Position	Unimod Accession No.																																																																																													
* Add a Modification...																																																																																																					
<input checked="" type="checkbox"/>	new	new	123.123	123.135				0																																																																																													
<table border="1"> <thead> <tr> <th>Amino Acid Name</th> <th>One Letter Code</th> <th>Classification</th> </tr> </thead> <tbody> <tr> <td><input checked="" type="checkbox"/></td> <td>Arginine</td> <td>R</td> <td>Chemical derivative</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td colspan="3">* Add an Amino Acid Site...</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Met-loss</td> <td>Met-loss</td> <td>Chemical derivative</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Trp->Gly</td> <td>Trp->Gly</td> <td>Co-translational</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>Lys-loss</td> <td>Lys-loss</td> <td>Isotopic label</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td></td> <td></td> <td>Multiple</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td></td> <td></td> <td>N-linked glycosylation</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td></td> <td></td> <td>Non-standard residue</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td></td> <td></td> <td>O-linked glycosylation</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td></td> <td></td> <td>Other</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>									Amino Acid Name	One Letter Code	Classification	<input checked="" type="checkbox"/>	Arginine	R	Chemical derivative						* Add an Amino Acid Site...									<input checked="" type="checkbox"/>	Met-loss	Met-loss	Chemical derivative						<input checked="" type="checkbox"/>	Trp->Gly	Trp->Gly	Co-translational						<input checked="" type="checkbox"/>	Lys-loss	Lys-loss	Isotopic label						<input checked="" type="checkbox"/>			Multiple						<input checked="" type="checkbox"/>			N-linked glycosylation						<input checked="" type="checkbox"/>			Non-standard residue						<input checked="" type="checkbox"/>			O-linked glycosylation						<input checked="" type="checkbox"/>			Other					
Amino Acid Name	One Letter Code	Classification																																																																																																			
<input checked="" type="checkbox"/>	Arginine	R	Chemical derivative																																																																																																		
* Add an Amino Acid Site...																																																																																																					
<input checked="" type="checkbox"/>	Met-loss	Met-loss	Chemical derivative																																																																																																		
<input checked="" type="checkbox"/>	Trp->Gly	Trp->Gly	Co-translational																																																																																																		
<input checked="" type="checkbox"/>	Lys-loss	Lys-loss	Isotopic label																																																																																																		
<input checked="" type="checkbox"/>			Multiple																																																																																																		
<input checked="" type="checkbox"/>			N-linked glycosylation																																																																																																		
<input checked="" type="checkbox"/>			Non-standard residue																																																																																																		
<input checked="" type="checkbox"/>			O-linked glycosylation																																																																																																		
<input checked="" type="checkbox"/>			Other																																																																																																		

6F

7. 在 Administration 界面, 点击“Cleavage Reagents”进入酶切试剂自定义界面, 常用的酶切试剂基本都在该表格中, 如需添加, 点击“Click here to add a new record”, 然后在名字、酶切位点、酶切抑制位点、酶切位置 (Offset: C 端酶切为 1; N 端、非酶切和非特异性酶切为 0)、简写、酶切特异性 (全酶切: Full, 半酶切: Semi, N 端半酶切: SemiN, C 端半酶切: SemiC, 非特异性酶切: Enzymatic Unspecific) 等区域分别进行输入和选择, 然后点击“Apply”即可在后续的数据分析中直接使用, 如需删除, 选择相应条目后, 点击“Remove”即可。

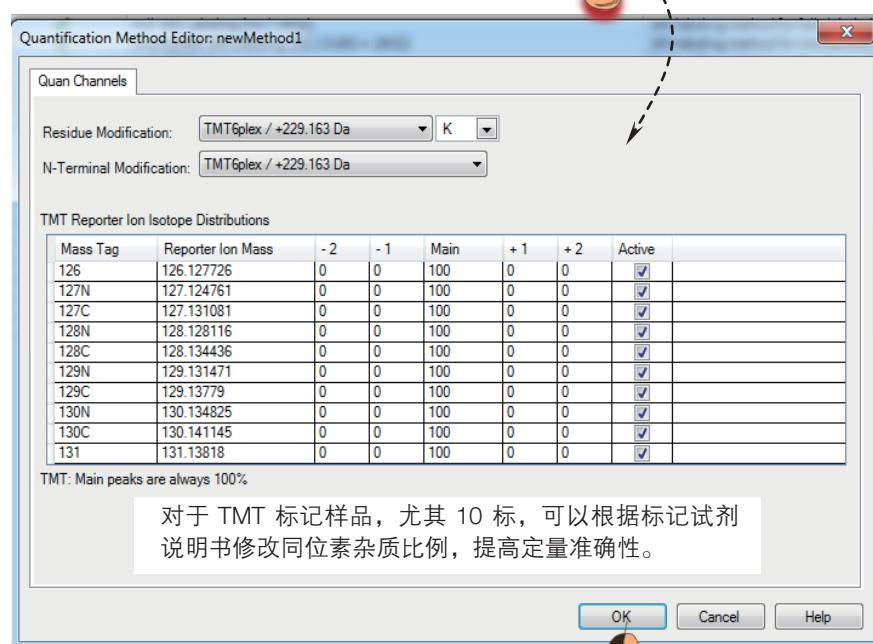
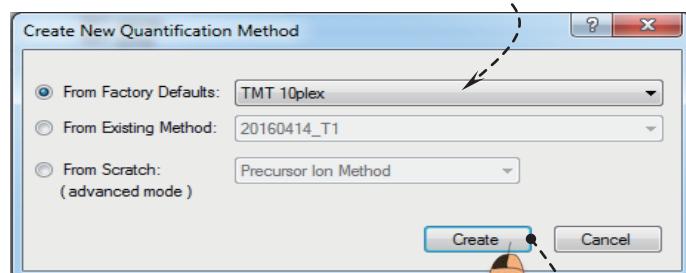
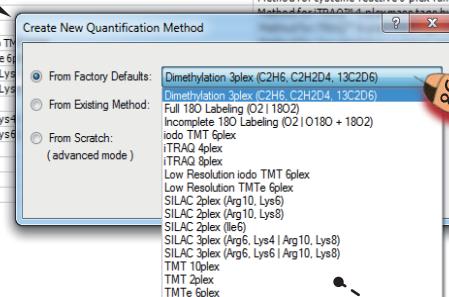
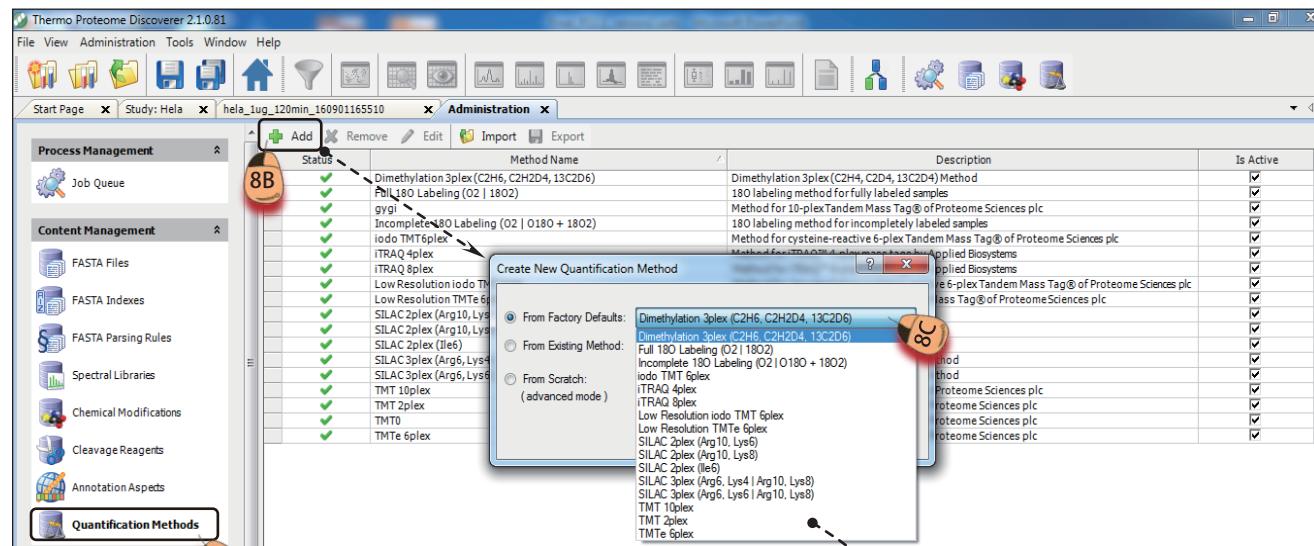


Name	Cleavage Sites	Cleavage Inhibitors	Offset	Abbreviation	Cleavage Specificities
Trypsin_R	R	P	1	-	Full; Semi
Trypsin_K	K	P	1	-	Full; Semi
Trypsin(KRLNH)	HKLRN	-	1	Try_a	Full; Semi
Trypsin	KR	P	1	Try	Full; Semi
Staph_Protease	E	-	1	-	Full
Proline_Endopept	P	-	1	-	Full
No-Enzyme	-	-	0	-	Unspecific
No-Cleavage	-	-	0	-	Semi; Semi (N-Term); Semi (C-Term); No Cleavages
LysC	K	-	1	-	Full
IodosoBenzote	W	-	1	-	Full
GluC	DE	-	1	-	Full
Elastase/Tryp/Chymo	AFIKLRLVWY	P	1	-	Full
Elastase	AILV	P	1	-	Full
Cyanogen_Bromide	M	-	1	-	Full
Clostripain	R	-	1	-	Full
Chymotrypsin(FWY)	FWY	P	1	ChTr	Full
Chymotrypsin	FLWY	-	1	ChyTr	Full
AspN	D	-	0	-	Full

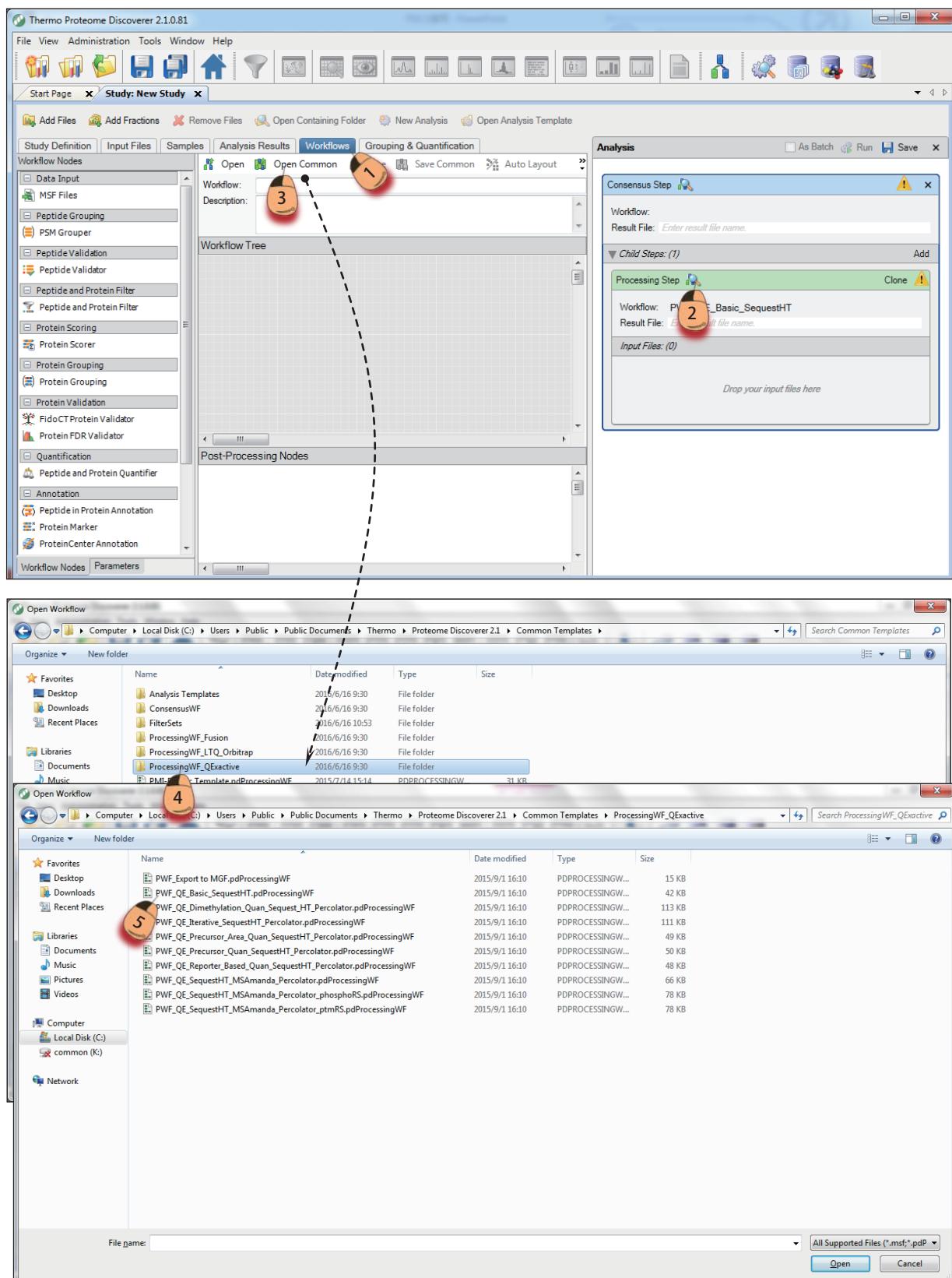


Name	Cleavage Sites	Cleavage Inhibitors	Offset	Abbreviation	Cleavage Specificities
7D	K	-	1	-	Full
Trypsin_R	R	P	1	-	Full; Semi
Trypsin_K	K	P	1	-	Full; Semi
Trypsin(KRLNH)	HKLRN	-	1	Try_a	Full; Semi
Trypsin	KR	P	1	Try	Full; Semi
Staph_Protease	E	-	1	-	Full
Proline_Endopept	P	-	1	-	Full
No-Enzyme	-	-	0	-	Unspecific
No-Cleavage	-	-	0	-	Semi; Semi (N-Term); Semi (C-Term); No Cleavages
LysC	K	-	1	-	Full
IodosoBenzote	W	-	1	-	Full
GluC	DE	-	1	-	Full
Elastase/Tryp/Chymo	AFIKLRLVWY	P	1	-	Full
Elastase	AILV	P	1	-	Full
Cyanogen_Bromide	M	-	1	-	Full
Clostripain	R	-	1	-	Full
Chymotrypsin(FWY)	FWY	P	1	ChTr	Full
Chymotrypsin	FLWY	-	1	ChyTr	Full
AspN	D	-	0	-	Full

8. 在 Administration 界面, 点击“Quantification Methods”进入定量方法自定义界面, 常用的各种定量方法基本都在该表格中, 例如 SILAC、TMT、iTRAQ、O¹⁸ 标记和二甲基化标记等。如需添加, 点击“Add”, 一般可以从工厂默认的模板中进行选择, 以 TMT 10 plex 为例, 点击“Create”之后, 对于定量方法引入的质量增加和修饰位点分别进行选择和修改, 针对需要高分辨设置的 TMT 10 plex 试剂, 还需要查阅定量试剂盒的说明书, 补充报告离子的同位素丰度信息, 然后点击“OK”, 再进行定量方法的命名和保存。另外还可以对定量方法进行删除、编辑、导入和导出操作, 导入和导出的定量方法的格式为 .method。

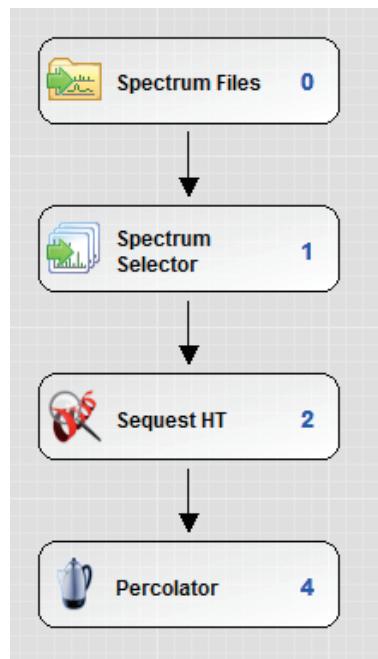


对于 TMT 标记样品, 尤其 10 标, 可以根据标记试剂说明书修改同位素杂质比例, 提高定量准确性。



Processing workflow 可从 PD 默认的模板中选择，根据需要再进行修改，打开步骤按照“鼠标”提示进行；第 1 步：在打开的 study 中点击 workflow；第 2 步：点击右侧 Processing step 中的放大镜；第 3 步：点击 open common 即可打开 PD 中已存储的默认模板；第 4 步：点击 Processing WF 所在文件夹；第 5 步：选取所需的 Processing 模板，QE_Basic_SequestHT 模板最常用，同时可在此模板基础上进行修改。

QE_Basic_SequestHT 模板中各模块意义说明



各模块作用如下：

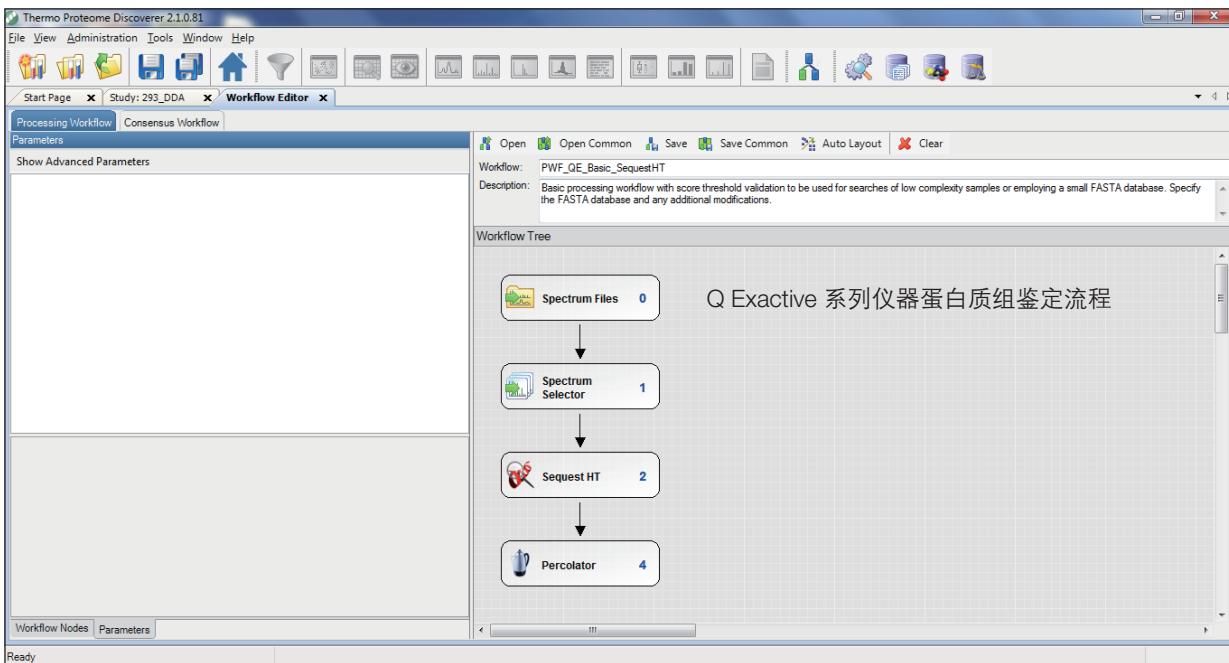
Spectrum files: 数据输入，无需设置任何参数。

Spectrum Selector: 从原始数据中抽提谱图送入搜索引擎模块进行搜索。

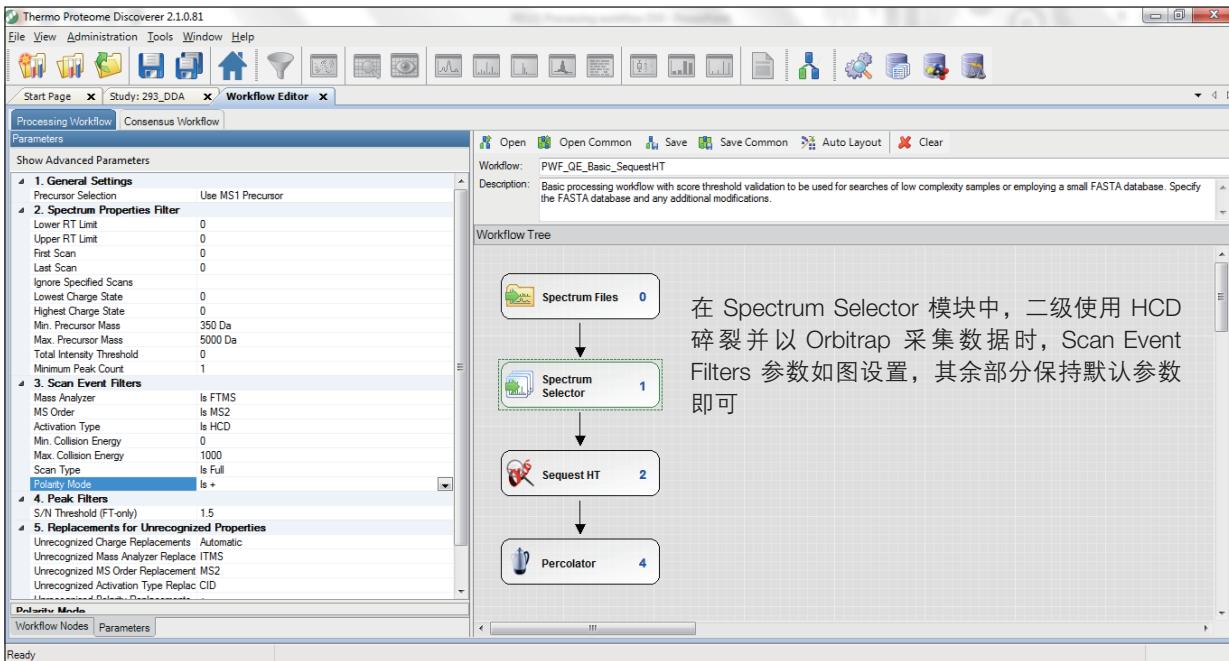
Sequest HT: 搜索引擎模块。

Percolator: 按照 delta Cn、q-value 等对 PSM (peptide-spectrum match) 进行卡值。

每个模块详细参数设置见下页



上图展示了 Q Exactive 系列仪器进行常规蛋白质组鉴定实验所用的搜库流程，LTQ-Orbitrap 和 Fusion 系列仪器二级使用 HCD 碎裂并以 Orbitrap 采集数据时，同样也可以使用此流程。



参数设置

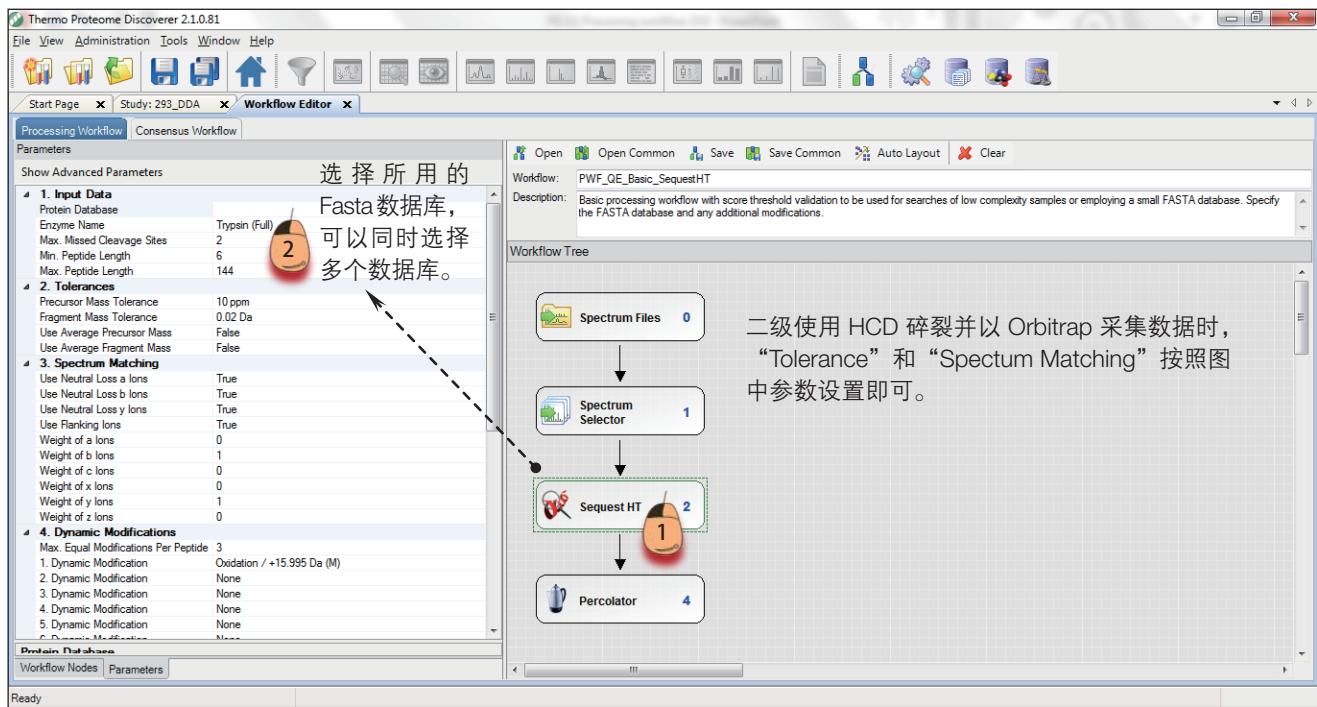
General settings: 前体离子为 MS 1 还是 MS (n-1)，通常设置为 MS 1，采集三级以上质谱时选择 MS (n-1)。

Spectrum Properties Filter: 根据谱图的保留时间、谱图编号、前体离子电荷价态、母离子质量范围、总强度阈值、最小谱峰数目等筛选特定谱图搜库，通常不设限制。

Scan Event Filters: 根据二级/三级谱图碎裂模式、检测器、碰撞能量、扫描类型和采集极性模式设定相应参数，筛选谱图进行搜库。

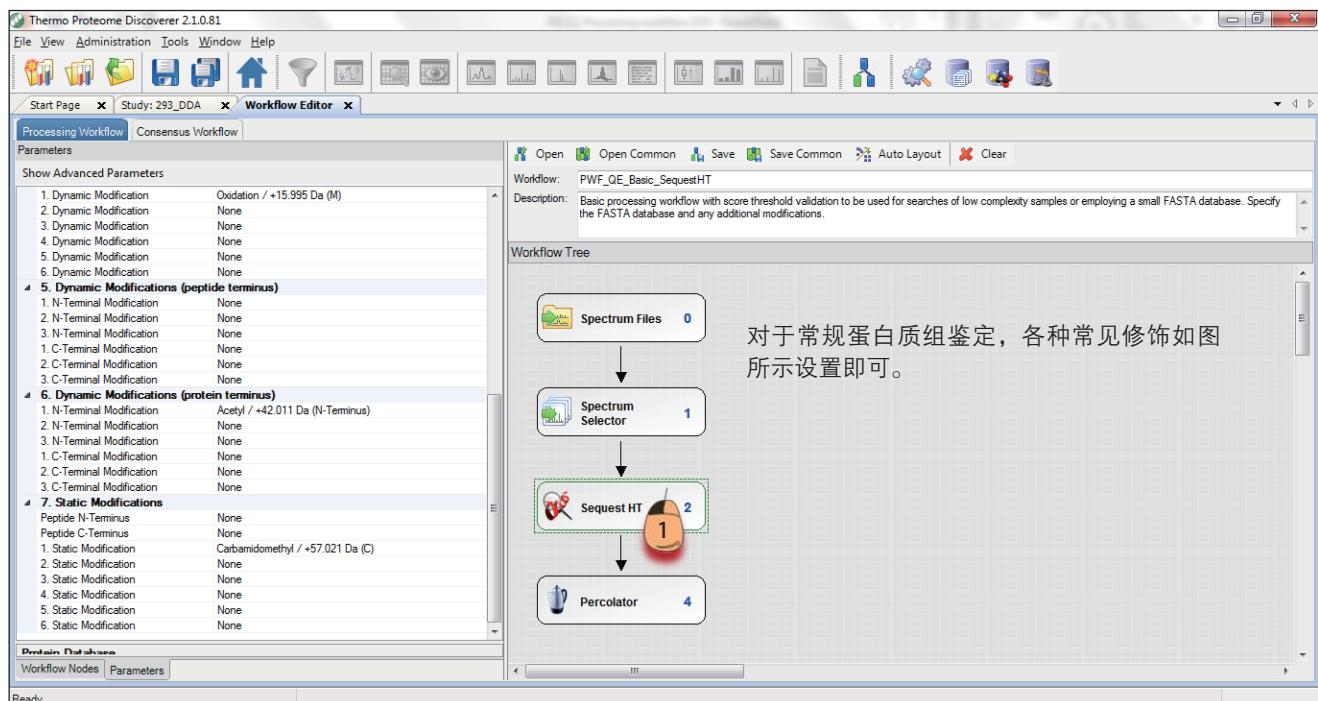
Peak Filters: 对谱图的信噪比设定阈值（仅适用于 Orbitrap）。

Replacements for Unrecognized Properties: 对于无法识别的电荷价态、质量检测器、碎裂方式等参数，使用此处设定的参数代替；通常使用默认参数。

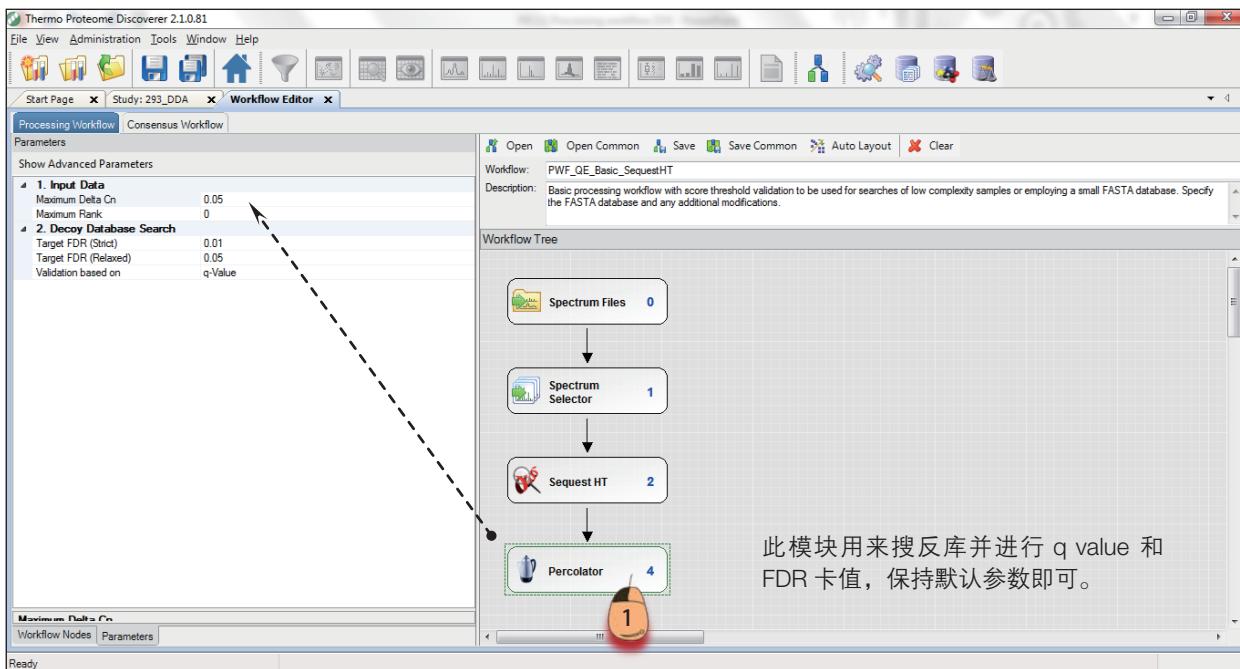


参数设置

1. Input Data: 设定搜索数据库、酶、最大漏切位点个数和最长/最短肽段长度信息。
2. Tolerances: 设定前体/碎片离子搜索时质量偏差容忍范围, 及是否使用前体/碎片离子的平均分子量。
3. Spectrum Matching: 设定在谱图匹配时是否使用中性丢失离子, 以及 a/b/c/x/y/z 离子所占的权重。



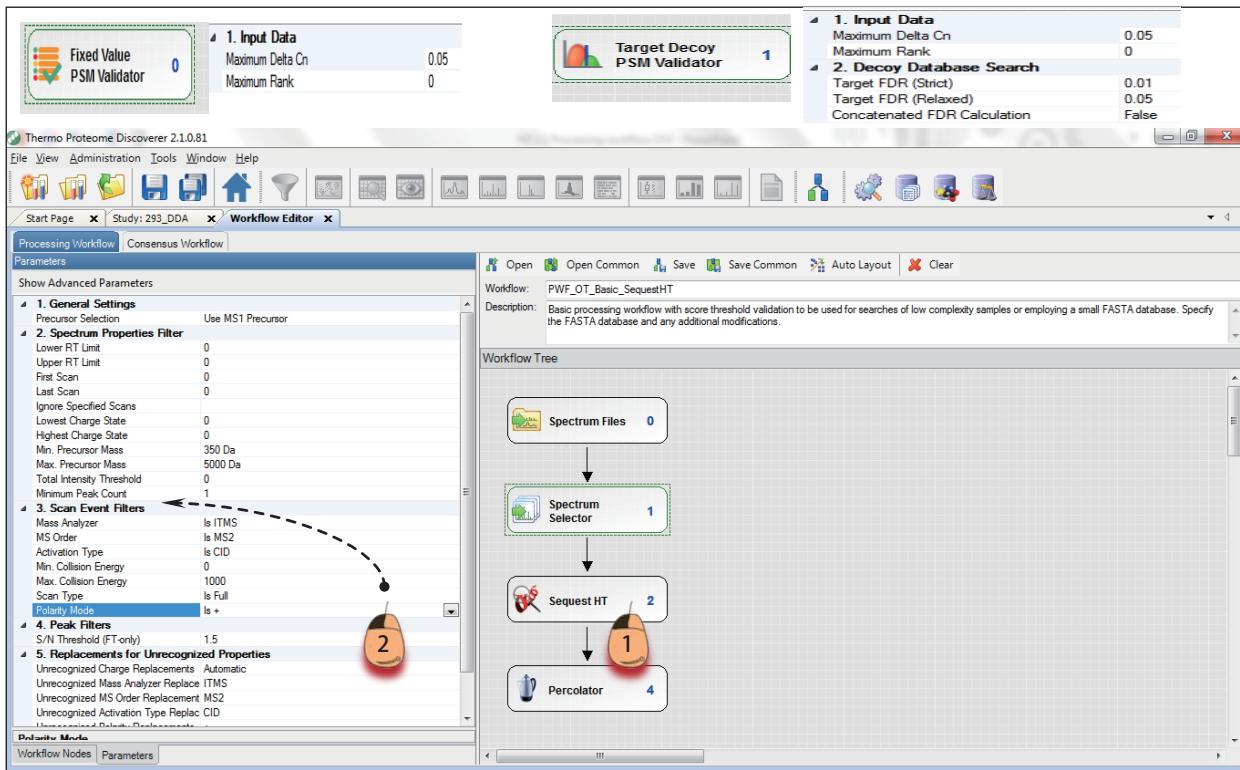
4~7项均为与修饰相关的设置: 其中 4“Dynamic Modifications”指的是在肽段侧链上可能发生的化学修饰(比如磷酸化、乙酰化、甲基化等) ; 5 “Dynamic Modifications (peptide terminus) ” 指的是在肽段末端上可能发生的化学修饰; 6 “Dynamic Modifications (protein terminus) ” 指的是在蛋白末端上可能发生的化学修饰; 7 “Static Modifications” 指的是在某个位点上一定会发生的修饰 (比如样品还原烷基化过程中引入的 Carbomidomethyl 化学修饰) 。



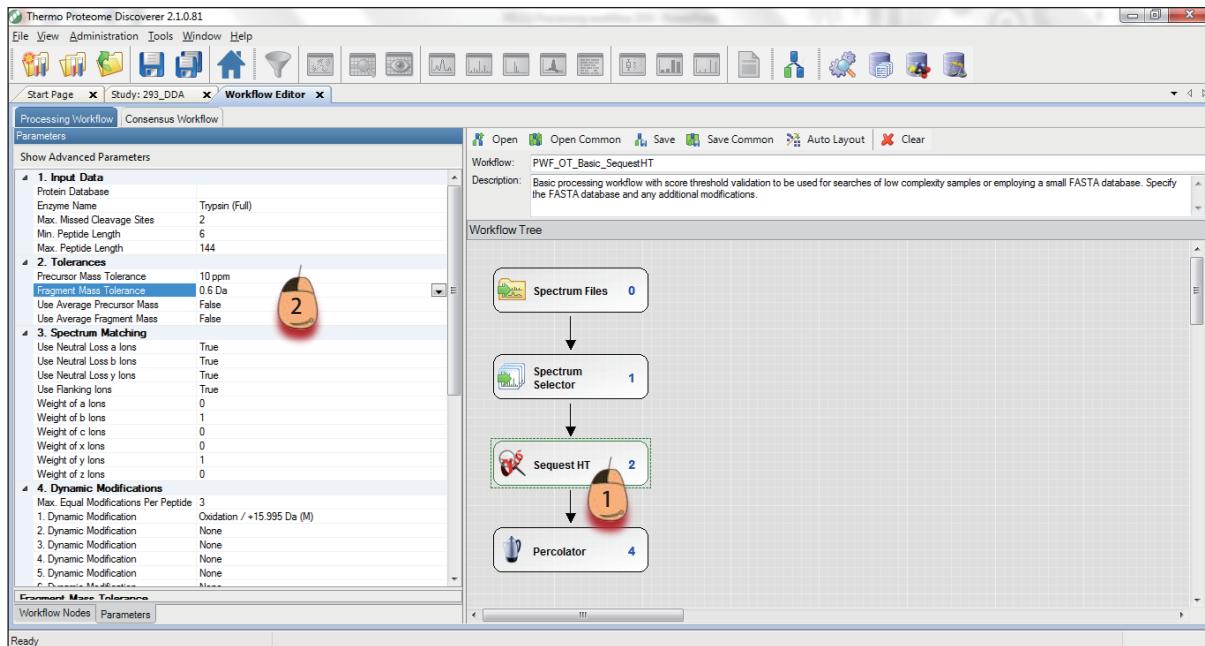
参数设置

- 1、Input Data: 对 PSM 的 Maximum Delta Cn 和 Maximum Rank 进行卡值。
- 2、Decoy Database Search: 搜索反库, 以 q-value 为基准进行 FDR 卡值, 通常设置 Target FDR (Strict) =0.01, Target FDR (Relaxed) =0.05。

另外两个 PSM Validation 的模块, Fixed Value PSM Validator 只对 PSM 的 Maximum Delta Cn 和 Maximum Rank 进行卡值, 适用于谱图库检索。Target Decoy PSM Validator 对在 PSM 的 Maximum Delta Cn 和 Maximum Rank 进行卡值的同时还会搜反库并对 FDR 进行卡值, 适用于数据库蛋白个数少时的数据库检索情况。



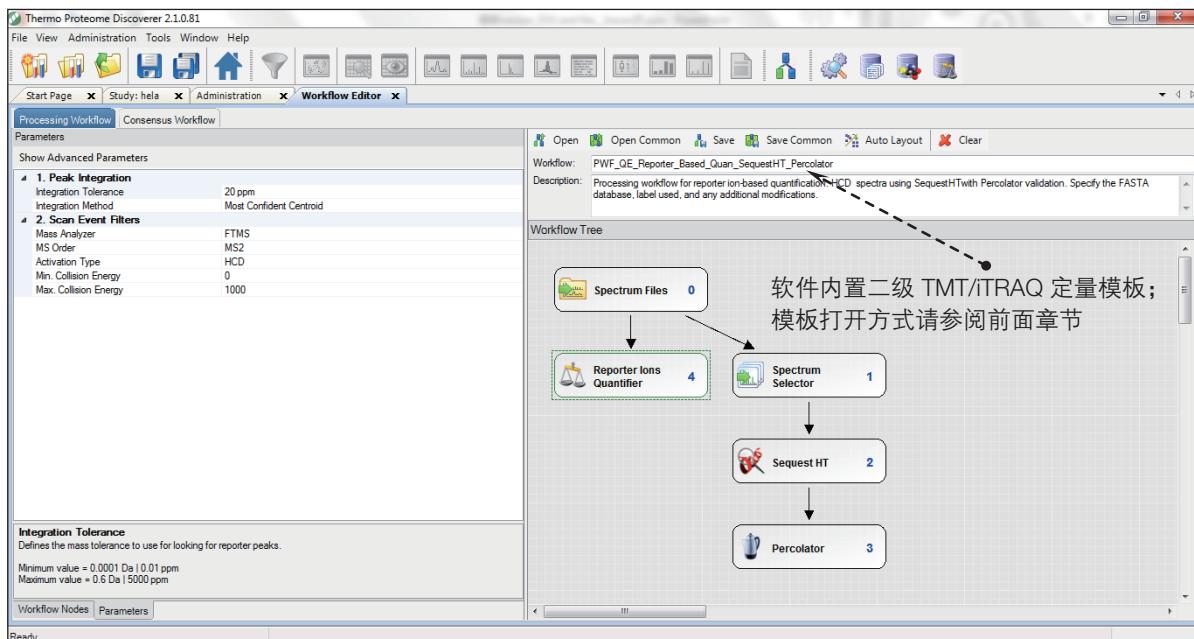
当二级使用 CID 碎裂, 离子阱采集数据时, 由于碎裂方式和检测器的变化, 要注意 Spectrum Selector-Scan Event Filters 的参数设置有所不同 (如上图所示); 各参数的详细含义请参阅前面章节。



二级使用 CID/HCD 碎裂并以离子阱采集数据时，“Tolerance”和“Spectrum Matching”按照图中参数设置即可。由于离子阱为低分辨率质量检测器，通常建议二级离子的质量精度在 0.6~0.8 Da 之间为宜。

备注：Proteome Discoverer 软件还可以对多种不同碎裂方式、不同实验采集方法的数据进行搜库，具体参数可参考相应的模板以及软件说明书。

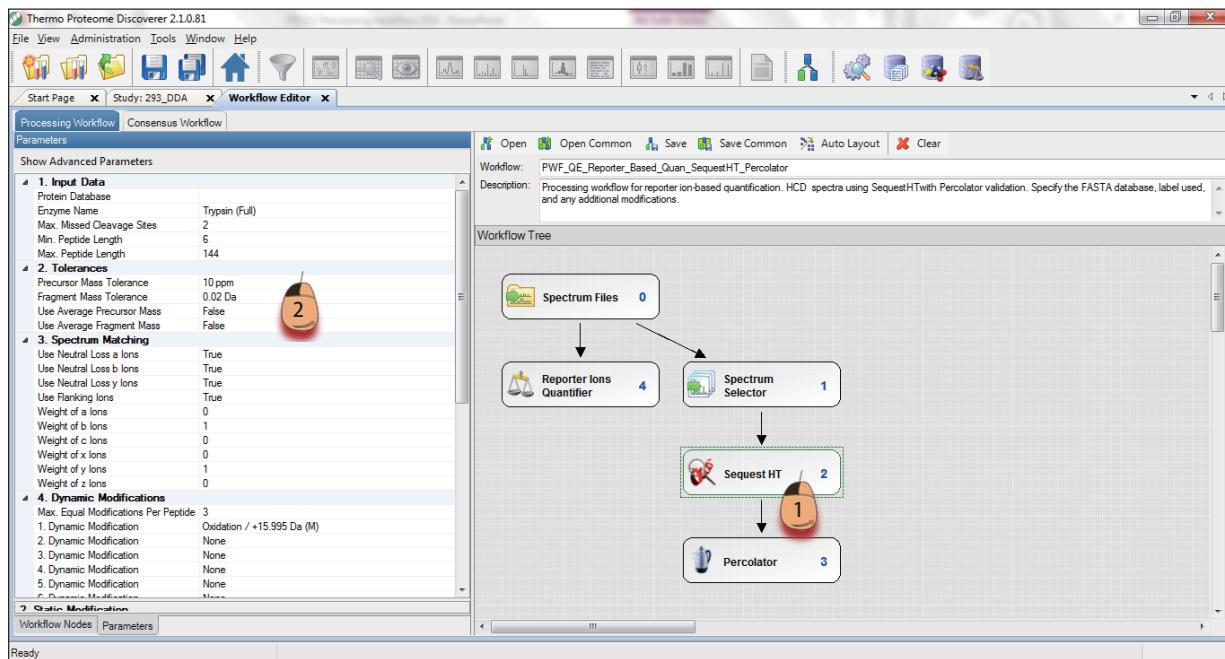
二级 TMT/iTRAQ 定量



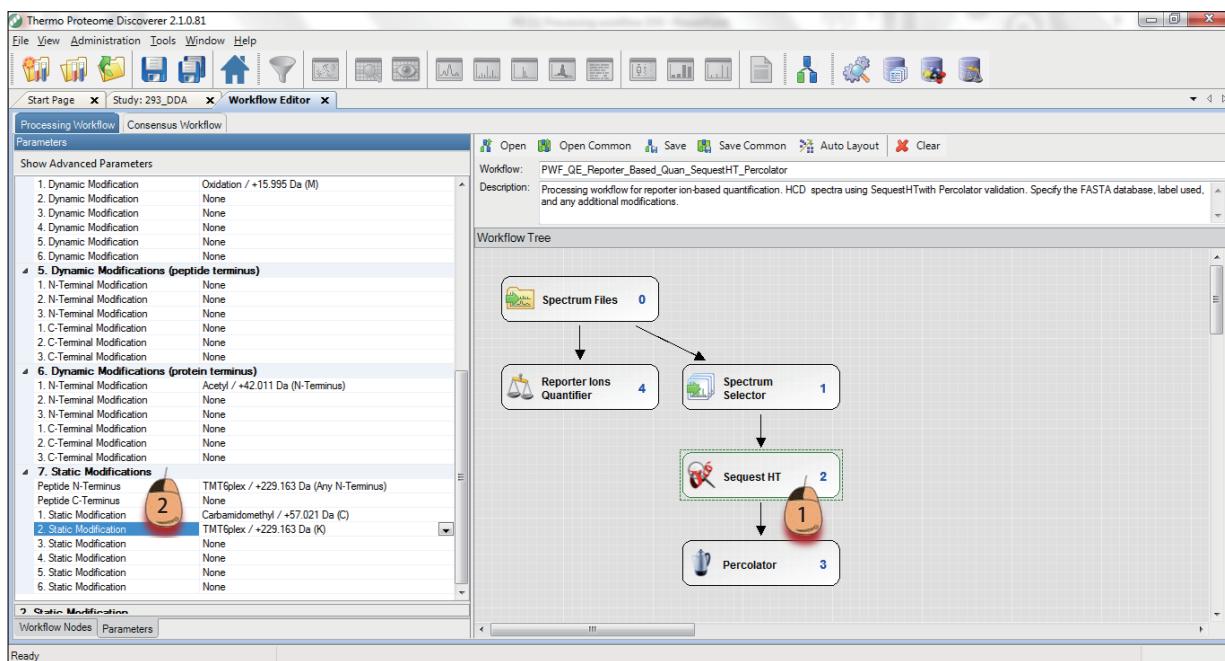
对于 Q Exactive/LTQ Orbitrap/Fusion 系列仪器，当使用二级 TMT/iTRAQ 定量时，请选择上图所示的内置模板。

Reporter Ions Quantifier 模块

1. Peak Integration：确定从原始谱图中提取报告离子谱峰时的允许质量偏差和提峰方法，通常使用默认的 20 ppm 和“Most Confident Centroid”即可。
2. Scan Event Filters：指定采集定量谱图时所用的质量检测器、谱图级数和碎裂方法。对于二级 TMT/iTRAQ 定量，设置为 Orbitrap 质量检测器、MS₂ 谱图、HCD 碎裂和碎裂能量从 0~1000 即可。

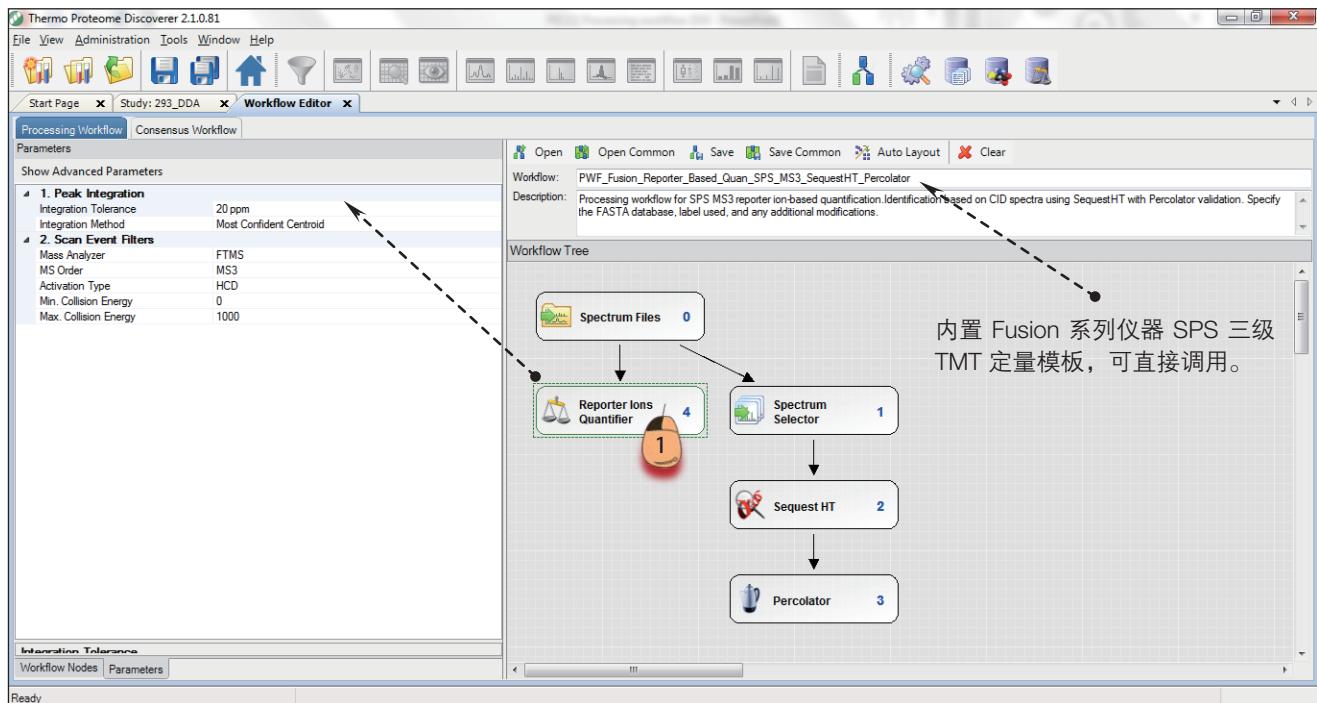


基于二级报告离子进行 TMT/iTRAQ 定量时，都会使用 HCD 碎裂，Orbitrap 采集二级谱图，所以需要按照上图所示进行 SequestHT-Tolerance 的参数设置。



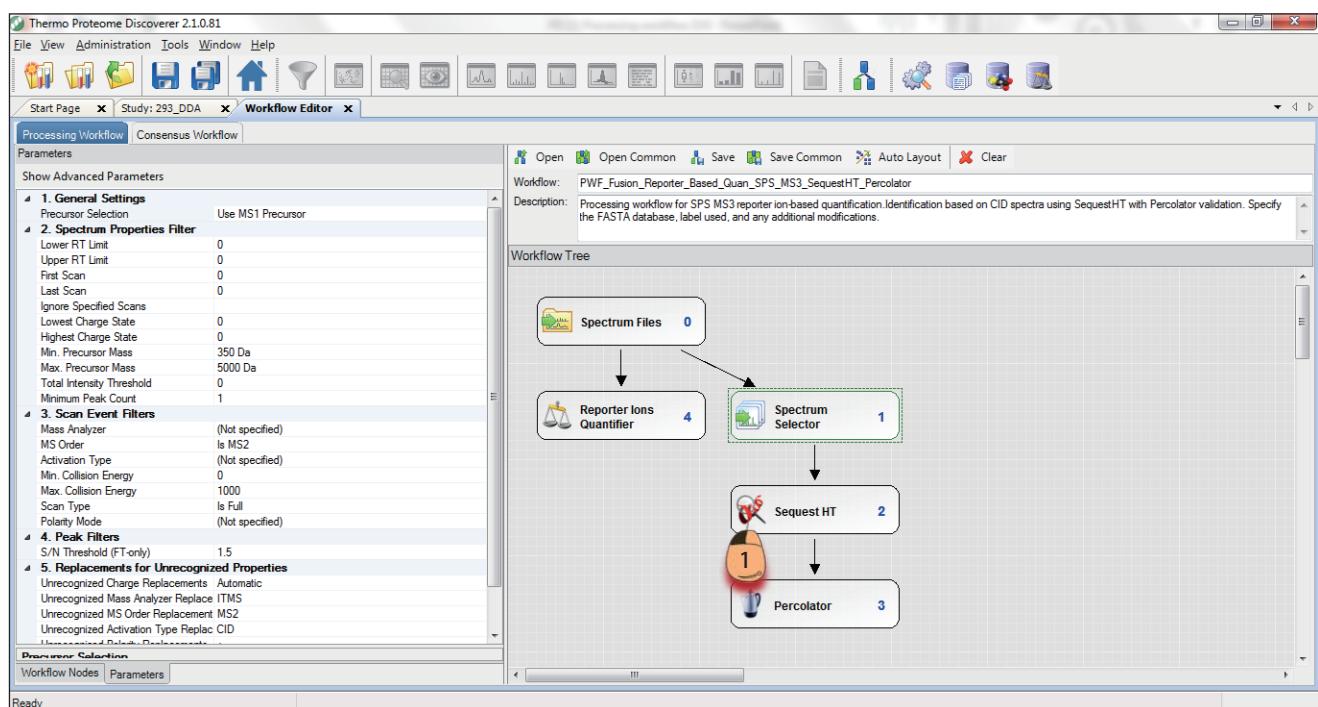
当对肽段进行 TMT/iTRAQ 标记后，理论上所有肽段的氨基（包括 N 末端和 K 侧链上的氨基）均会被标记，所以可以将 TMT/iTRAQ 标签设置为固定修饰。上图中以 TMT 6 plex 为例，展示了如何将 TMT/iTRAQ 标签设置为肽段 N 末端和 K 侧链固定修饰的步骤。

*：由于 iTRAQ 标记过程中会有副反应发生，建议在进行 iTRAQ 标记的搜库时，将其设置为 K, Y 侧链上的固定修饰。

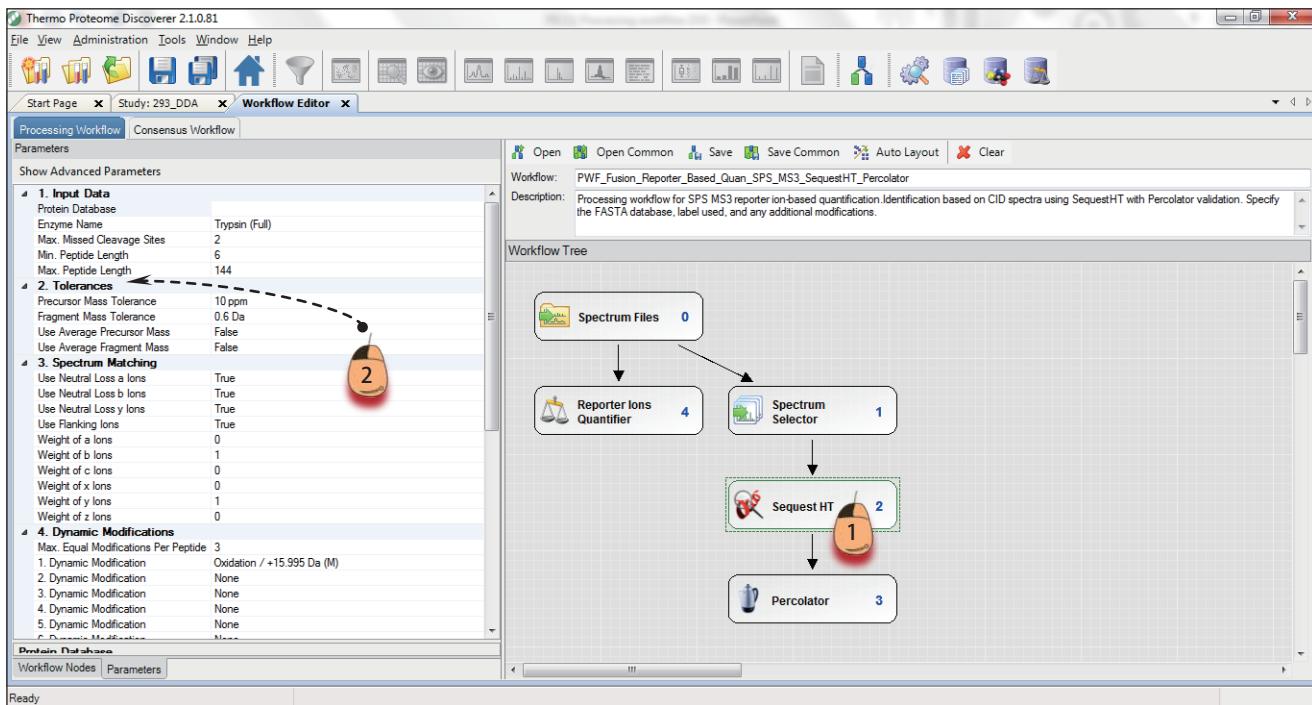


参数设置

1. Peak Integration: 确定从原始谱图中提取报告离子谱峰时的允许质量偏差和提峰方法，通常使用默认的 20 ppm 和“Most Confident Centroid”即可。
2. Scan Event Filters: 指定采集定量谱图时所用的质量检测器、谱图级数和碎裂方法。对于三级 SPS TMT 定量，设置为 Orbitrap 质量检测器、MS₃ 谱图、HCD 碎裂和碎裂能量从 0~1000 即可。



当进行三级 SPS TMT 定量时，Spectrum Selector-Scan Event Filters 参数请按照上图红色方框中设置即可。



Thermo Proteome Discoverer 2.1.0.81

File View Administration Tools Window Help

Start Page Study: 293_DDA Workflow Editor

Processing Workflow Consensus Workflow

Parameters

Show Advanced Parameters

1. Input Data

- Protein Database: Trypsin (Full)
- Enzyme Name: Trypsin (Full)
- Max. Missed Cleavage Sites: 2
- Min. Peptide Length: 6
- Max. Peptide Length: 144

2. Tolerances

- Precursor Mass Tolerance: 10 ppm
- Fragment Mass Tolerance: 0.6 Da
- Use Average Precursor Mass: False
- Use Average Fragment Mass: False

3. Spectrum Matching

- Use Neutral Loss a Ions: True
- Use Neutral Loss b Ions: True
- Use Neutral Loss y Ions: True
- Use Flanking Ions: True
- Weight of a Ions: 0
- Weight of b Ions: 1
- Weight of c Ions: 0
- Weight of x Ions: 0
- Weight of y Ions: 1
- Weight of z Ions: 0

4. Dynamic Modifications

- Max. Equal Modifications Per Peptide: 3
- 1. Dynamic Modification: Oxidation / +15.995 Da (M)
- 2. Dynamic Modification: None
- 3. Dynamic Modification: None
- 4. Dynamic Modification: None
- 5. Dynamic Modification: None
- 6. Dynamic Modification: None

5. Dynamic Modifications (peptide terminus)

- 1. N-Terminal Modification: None
- 2. N-Terminal Modification: None
- 3. N-Terminal Modification: None
- 4. C-Terminal Modification: None
- 5. C-Terminal Modification: None
- 6. C-Terminal Modification: None

6. Dynamic Modifications (protein terminus)

- 1. N-Terminal Modification: Acetyl / +42.011 Da (N-Terminus)
- 2. N-Terminal Modification: None
- 3. N-Terminal Modification: None
- 4. C-Terminal Modification: None
- 5. C-Terminal Modification: None
- 6. C-Terminal Modification: None

7. Static Modifications

- Peptide N-Terminus: TMTplex / +229.163 Da (Any N-Terminus)
- Peptide C-Terminus: None
- 1. Static Modification: Carbamidomethyl / +57.021 Da (C)
- 2. Static Modification: TMTplex / +229.163 Da (K)
- 3. Static Modification: None
- 4. Static Modification: None
- 5. Static Modification: None
- 6. Static Modification: None

Workflow Nodes Parameters

Workflow Tree

Open Open Common Save Save Common Auto Layout Clear

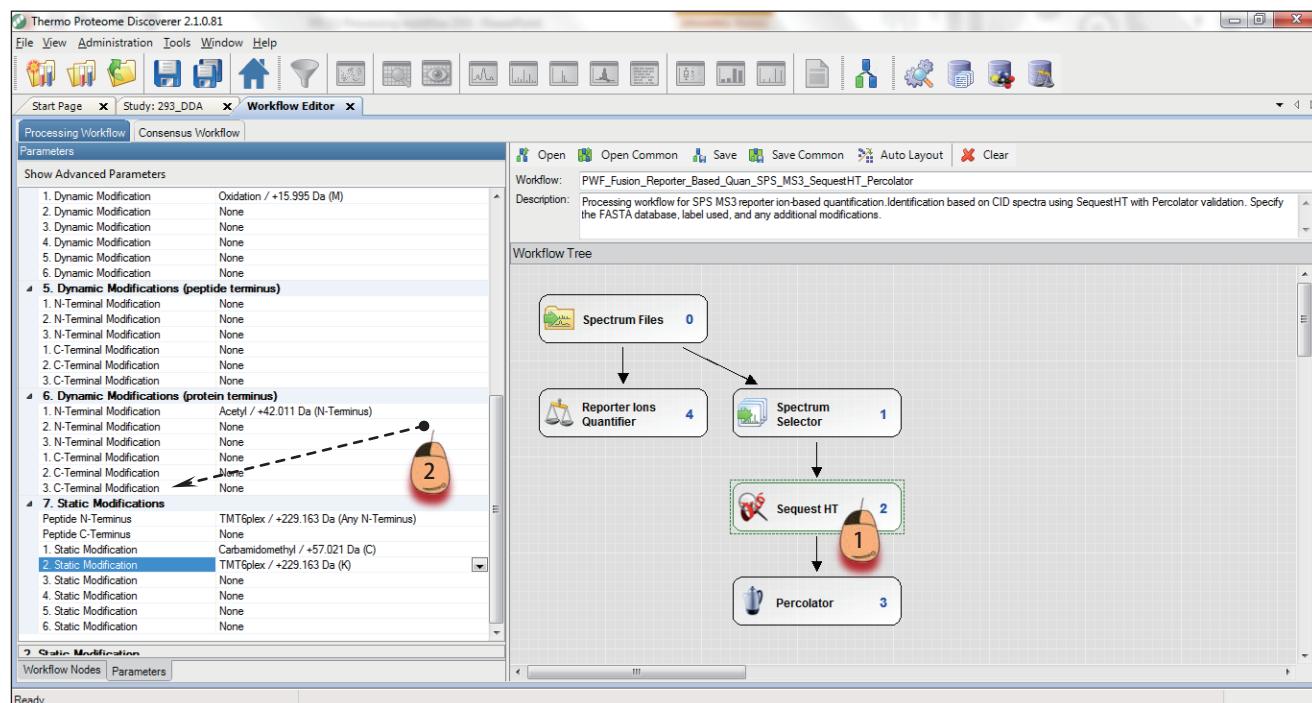
Workflow: PWF_Fusion_Reportor_Based_Quan_SPS_MS3_SequestHT_Percolator

Description: Processing workflow for SPS MS3 reporter ion-based quantification. Identification based on CID spectra using SequestHT with Percolator validation. Specify the FASTA database, label used, and any additional modifications.

```

graph TD
    A[Spectrum Files 0] --> B[Reporter Ions Quantifier 4]
    A --> C[Spectrum Selector 1]
    B --> D[Sequest HT 2]
    C --> D
    D --> E[Percolator 3]
  
```

当使用三级 SPS TMT 定量时, Sequest HT 模块将离子阱采集的二级数据用于定性搜索, 故 Sequest HT-Tolerances-Fragment 要设置为离子阱采集数据的质量精度范围, 通常在 0.6~0.8 Da 即可。



Thermo Proteome Discoverer 2.1.0.81

File View Administration Tools Window Help

Start Page Study: 293_DDA Workflow Editor

Processing Workflow Consensus Workflow

Parameters

Show Advanced Parameters

1. Dynamic Modification: Oxidation / +15.995 Da (M)

2. Dynamic Modification: None

3. Dynamic Modification: None

4. Dynamic Modification: None

5. Dynamic Modification: None

6. Dynamic Modification: None

5. Dynamic Modifications (peptide terminus)

1. N-Terminal Modification: None

2. N-Terminal Modification: None

3. N-Terminal Modification: None

1. C-Terminal Modification: None

2. C-Terminal Modification: None

3. C-Terminal Modification: None

6. Dynamic Modifications (protein terminus)

1. N-Terminal Modification: Acetyl / +42.011 Da (N-Terminus)

2. N-Terminal Modification: None

3. N-Terminal Modification: None

1. C-Terminal Modification: None

2. C-Terminal Modification: None

3. C-Terminal Modification: None

7. Static Modifications

Peptide N-Terminus: TMTplex / +229.163 Da (Any N-Terminus)

Peptide C-Terminus: None

1. Static Modification: Carbamidomethyl / +57.021 Da (C)

2. Static Modification: TMTplex / +229.163 Da (K)

3. Static Modification: None

4. Static Modification: None

5. Static Modification: None

6. Static Modification: None

Workflow Nodes Parameters

Workflow Tree

Open Open Common Save Save Common Auto Layout Clear

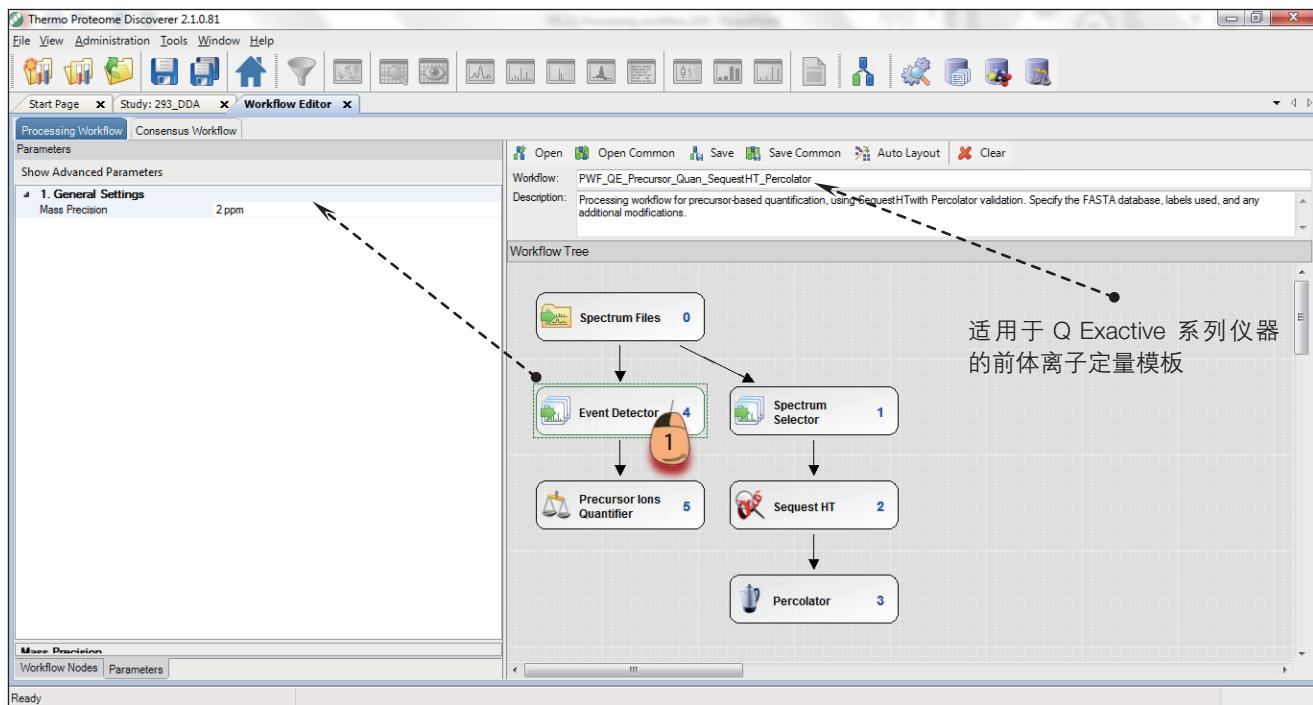
Workflow: PWF_Fusion_Reportor_Based_Quan_SPS_MS3_SequestHT_Percolator

Description: Processing workflow for SPS MS3 reporter ion-based quantification. Identification based on CID spectra using SequestHT with Percolator validation. Specify the FASTA database, label used, and any additional modifications.

```

graph TD
    A[Spectrum Files 0] --> B[Reporter Ions Quantifier 4]
    A --> C[Spectrum Selector 1]
    B --> D[Sequest HT 2]
    C --> D
    D --> E[Percolator 3]
  
```

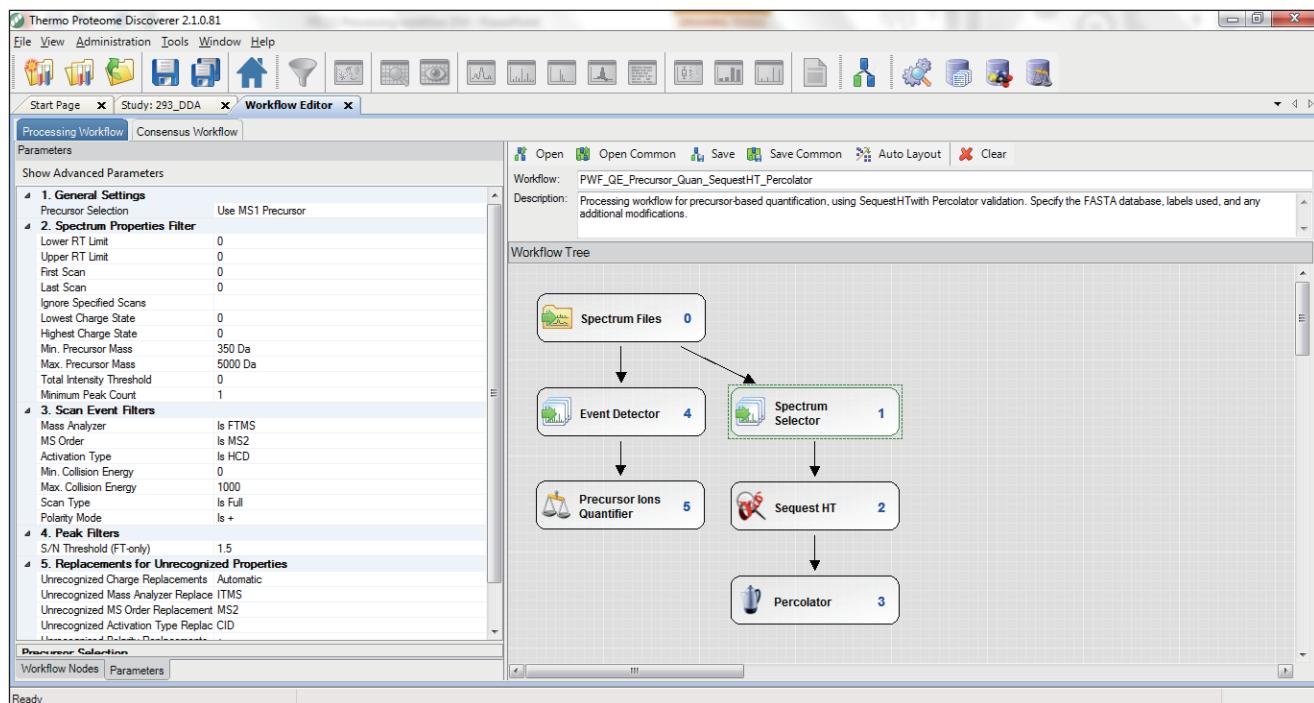
当使用三级 SPS TMT 定量时, Sequest HT 模块将离子阱采集的二级数据用于定性搜索, 此时需要将 TMT 标签设置为肽段 N 末端和赖氨酸侧链上的固定修饰。



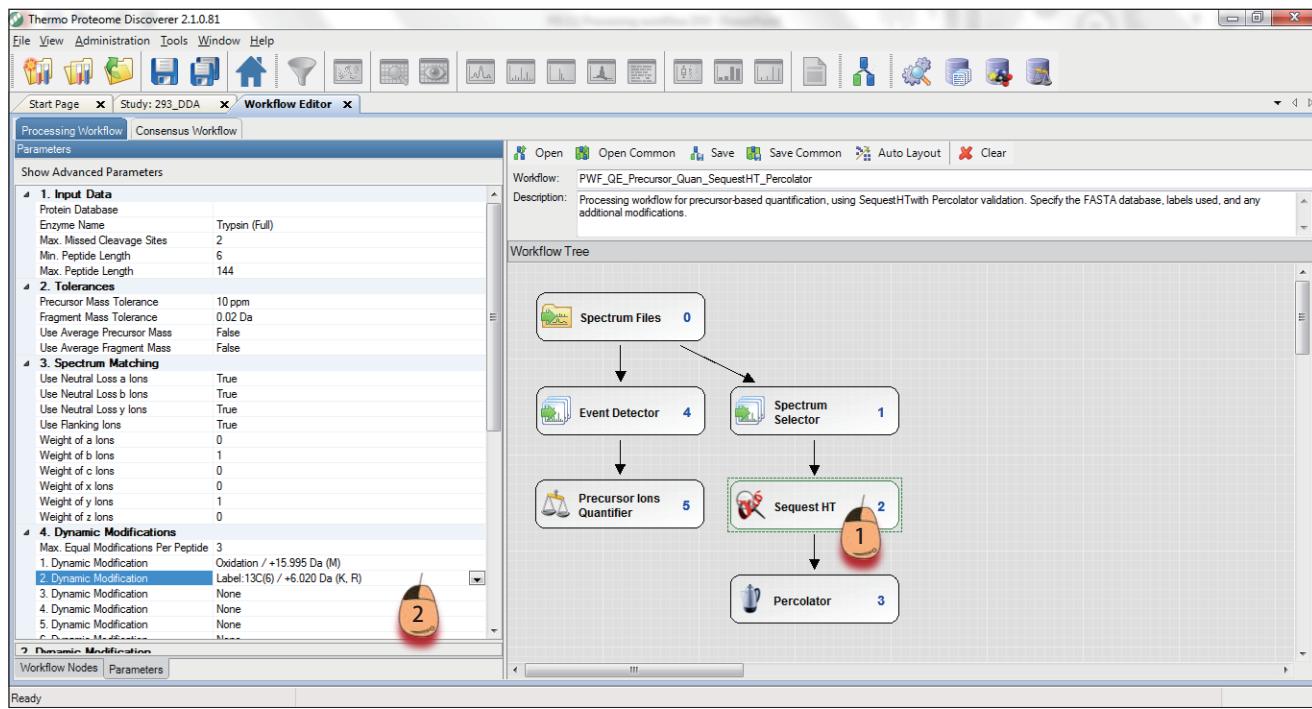
参数设置

当在 Q Exactive 系列仪器上进行基于前体离子峰面积定量实验（如 SILAC、¹⁸O 标记等）时，请调用此内置模板；LTQ-Orbitrap 和 Fusion 系列仪器二级使用 HCD 碎裂并以 Orbitrap 采集数据时，同样也可以使用此模板。

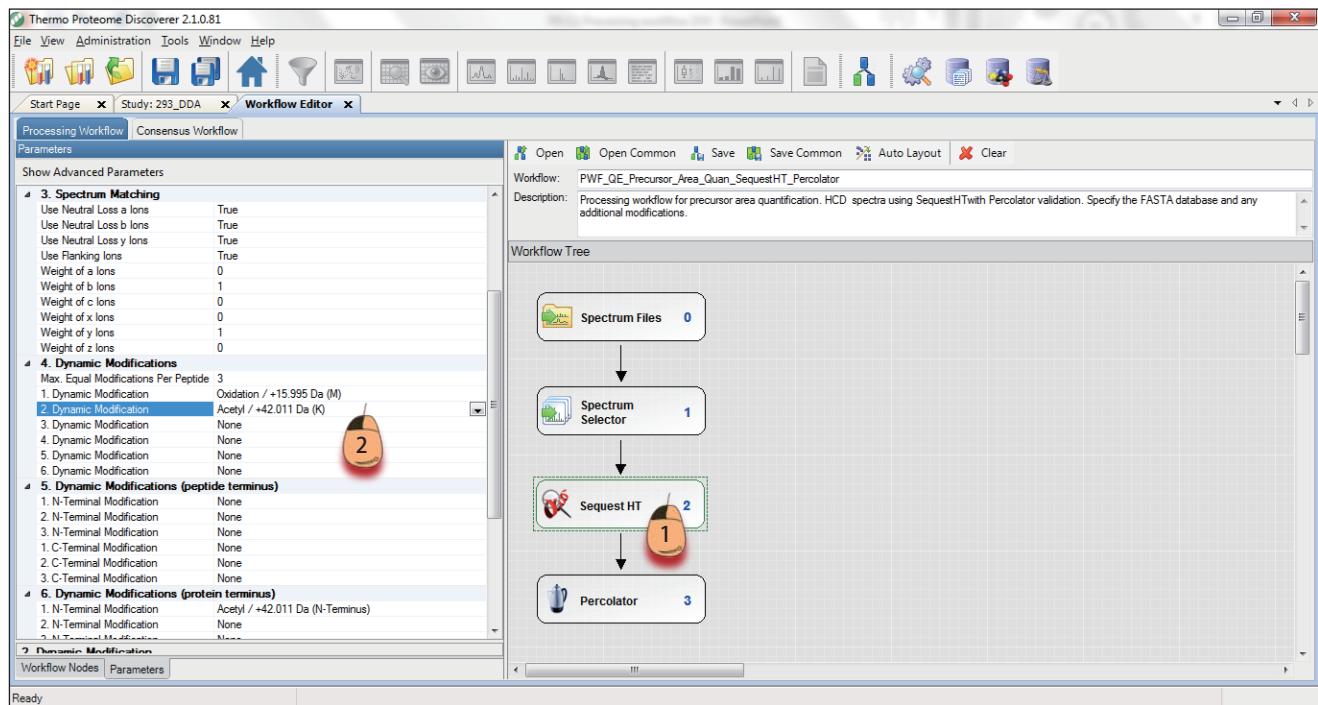
该模板中的 Event Detector 模块的作用为当从原始谱图中提取前体离子的色谱质谱流出峰时，要确保连续谱图内检测到的前体离子的精确质量的标准偏差在设定容忍范围内，通常保持默认的 2 ppm 即可。



Spectrum Selector 相关的参数设置请参照 Q Exactive 系列仪器定性实验参数。

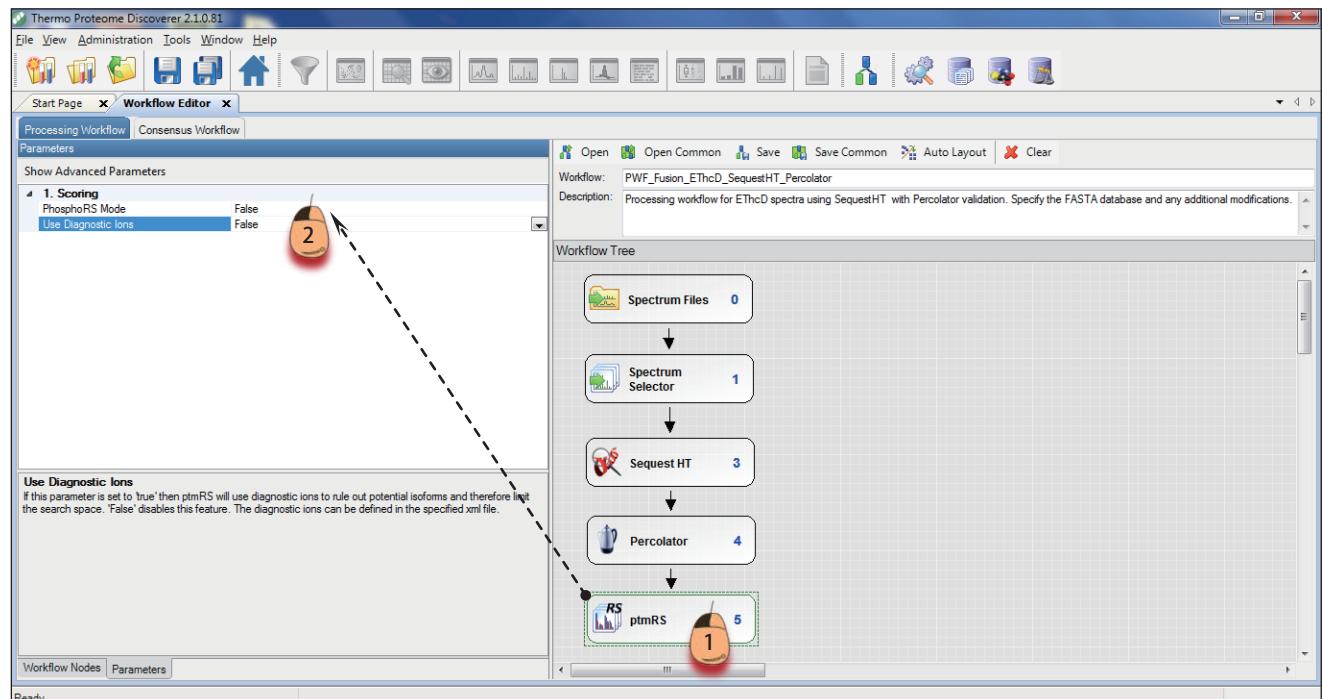


当进行基于前体离子峰面积定量（如 SILAC、¹⁸O 标记等）的实验时，会使用稳定同位素标记的氨基酸替代天然同位素组成的氨基酸，所以在 Sequest HT 模块中要将该标记设置为可变修饰。以 SILAC 实验为例，上图中以 K6R6（标记组中的 K, R 两种氨基酸分别比天然同位素组成的氨基酸质量增加了 6 Da）为例展示了如何设置该修饰。

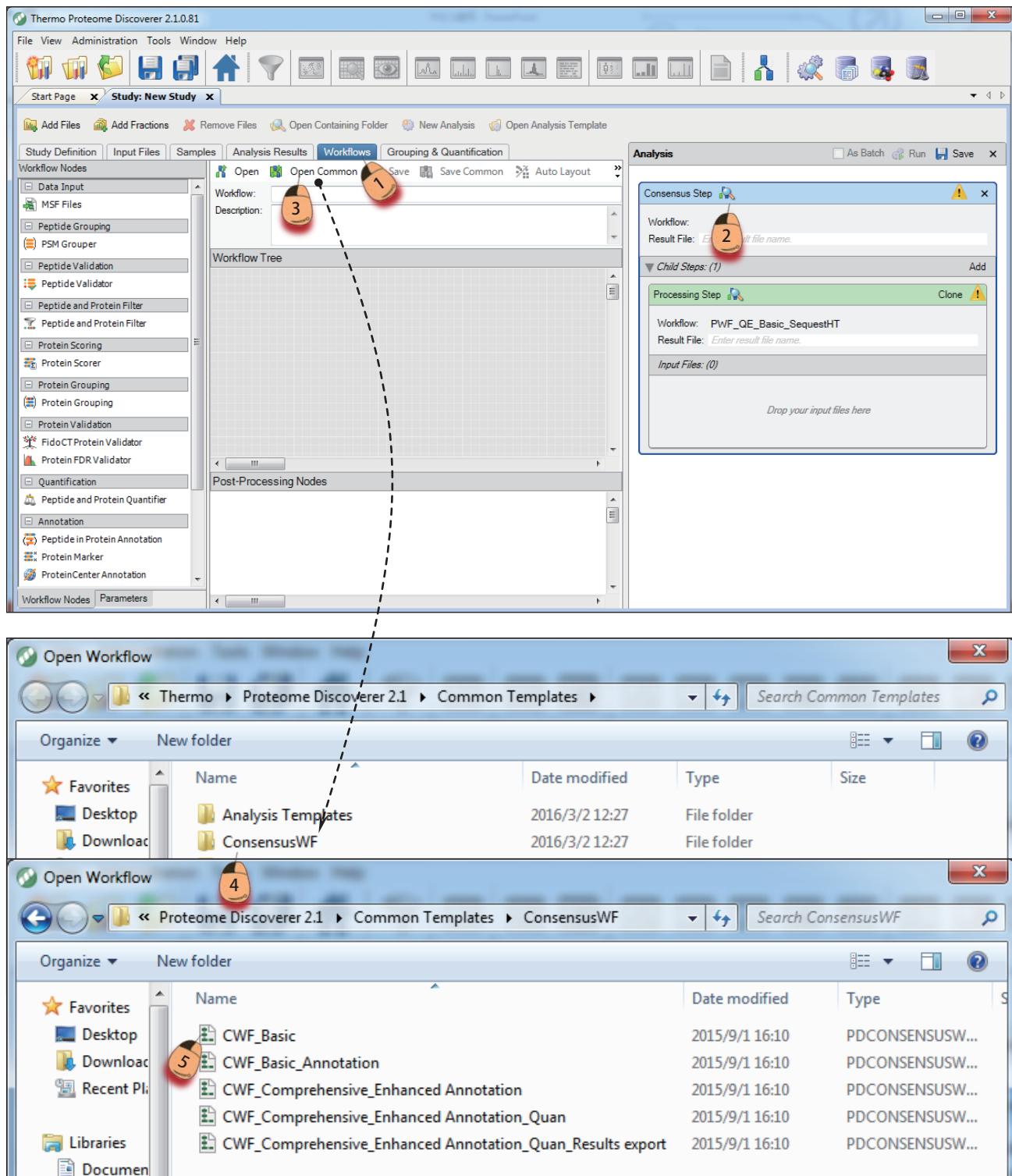


进行翻译后修饰鉴定时，一级、二级离子质量精度等参数请根据所使用的质量分析器设置，碎裂方式也请根据具体实验进行设置，具体参数值请参考相关定性实验设置；根据需要鉴定的翻译后修饰发生的位置（氨基酸侧链、肽段/蛋白末端）分别将其设置到特定的可变修饰（Dynamic Modification）中。

上图以赖氨酸侧链上发生的乙酰化为例，展示了如何进行翻译后修饰的设置。

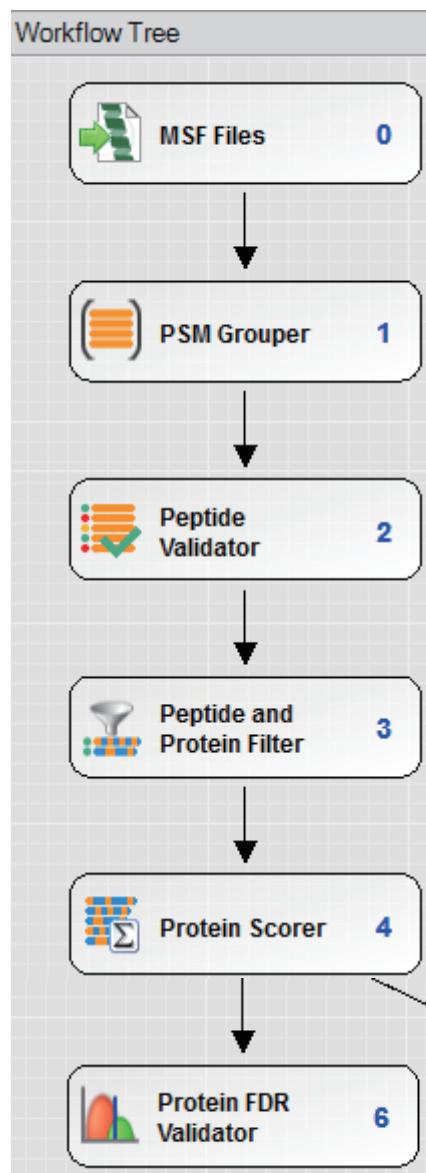


如需使用 PTMRS 模块对翻译后修饰位点可能性打分，请参考上图的流程。该模块中“PhosphoRS Mode”若设置为 True，意味着仅对磷酸化位点可能性进行打分；如果设置为 False，则会对所有可变修饰位点都进行打分。如果有该修饰特有的诊断离子，可以将“Use Diagnostic Ions”设置为 True，并添加这些诊断离子。如果没有 PTMRS 模块，需安装应用程序“Thermo Proteome Discoverer ThirdParty Components”，该程序可以从 <https://portal.thermo-brims.com/> 上下载。



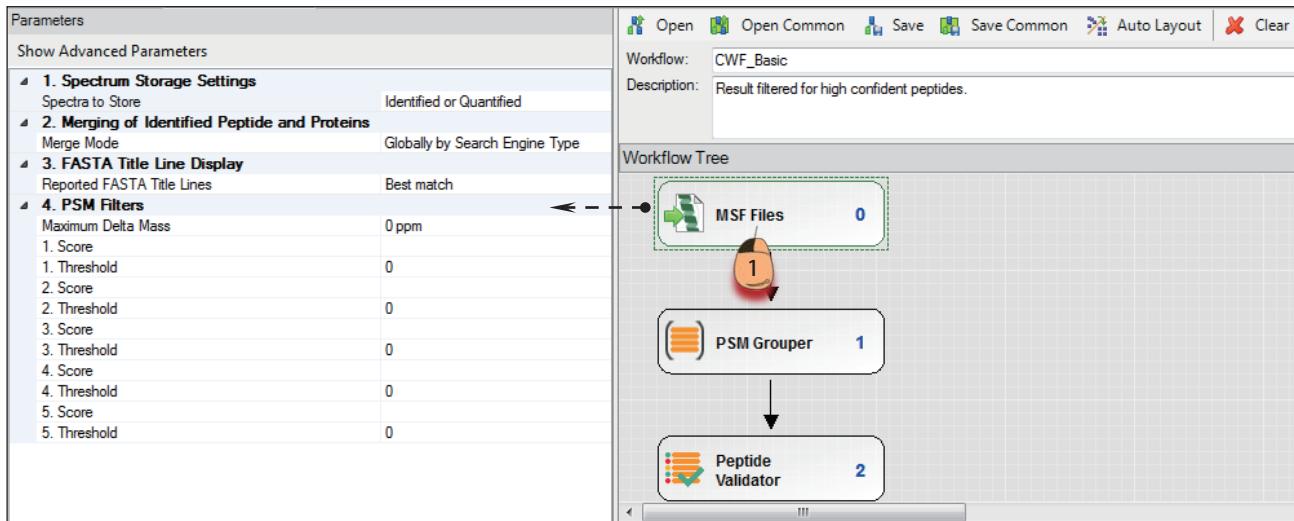
Consensus workflow 可从 PD 默认的模板中选择，根据需要再进行修改，打开步骤按照“鼠标”提示进行；第 1 步：在打开的 study 中点击 workflow；第 2 步：点击右侧 Consensus step 中的放大镜；第 3 步：点击 open common 即可打开 PD 中已存储的默认模板；第 4 步：点击 Consensus WF；第 5 步：选取所需的 Consensus 模板，CWF_Basic 模板最常用，同时可在 CWF_Basic 基础上进行修改。

CWF_Basic 模板中各模块意义说明



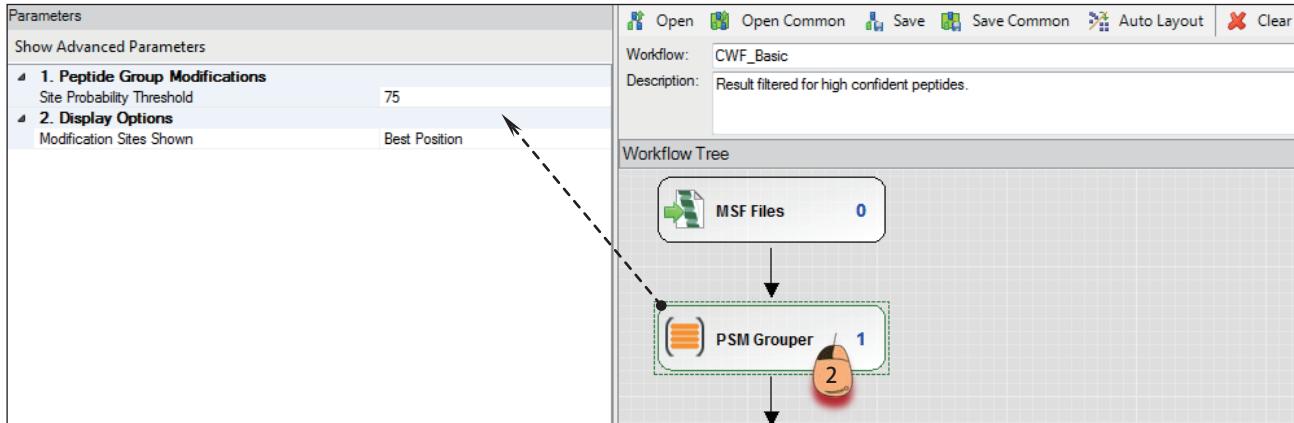
- MSF Files: 数据导入模块，自动将原始数据 processing 后产生的“Thermo Mass Spec Format”文件导入
- PSM Grouper: PSM 组装模块，将具有相同肽段序列（包括不同修饰）的 PSM 组装到一起
- Peptide Validator: 利用 PEPs、q-value 和 FDR 算法来验证肽段可信度
- Peptide and Protein Filter: 根据不同分数阈值过滤肽段和蛋白
- Protein scorer: 根据过滤后的 PSM 及肽段计算蛋白分数
- Protein Grouping: 对鉴定到的所有蛋白进行分组，生成最后结果中的 Protein Group
- Protein FDR Validator: 根据 decoy protein 的鉴定结果计算蛋白的假阳性率

每个模块详细参数设置见下页



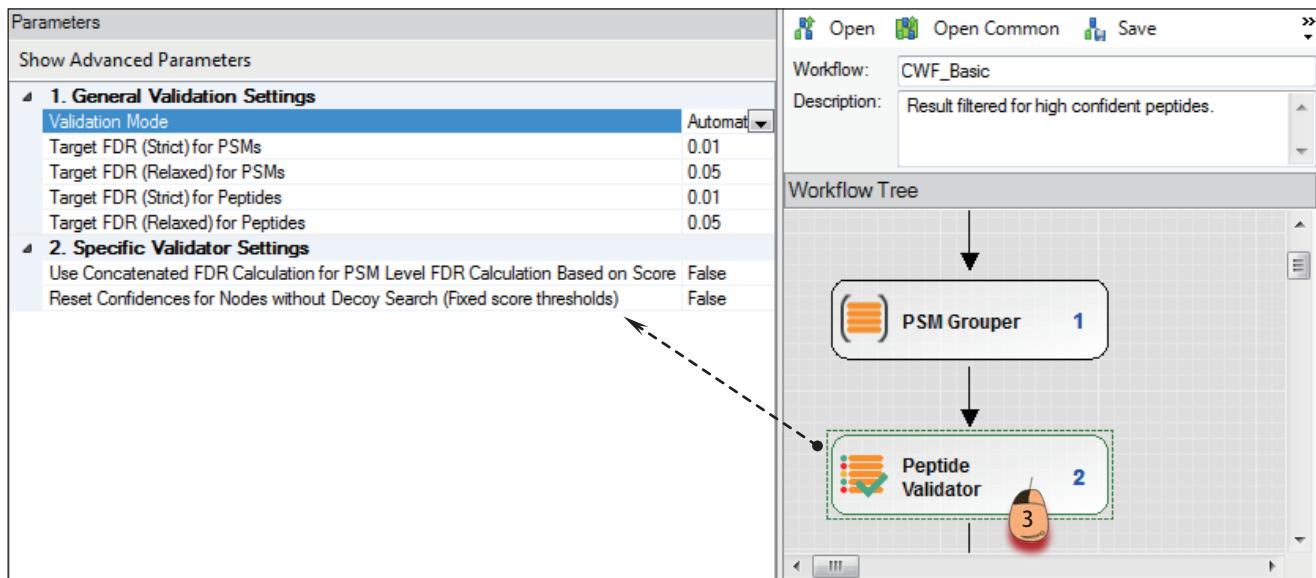
参数设置

1. 谱图保留设置：选择将要在 .pdResult 文件中保留的谱图，默认设置即可（只保留有鉴定或定量信息的谱图）
2. 鉴定结果的合并：根据搜索引擎进行结果合并（当用到多种搜索引擎时，此项起作用），当用 spectral counts 定量时选不合并，每个文件鉴定到的 PSMs 分别列出。
3. 鉴定到的蛋白以 FASTA 中的命名呈现：默认为呈现蛋白在 FASTA 中匹配最好的名称。
4. PSM 过滤：由于 Percolator 已经对 PSM 进行了过滤，所以此处建议不再进行 PSM 的过滤（默认为 0，代表不进行过滤）。



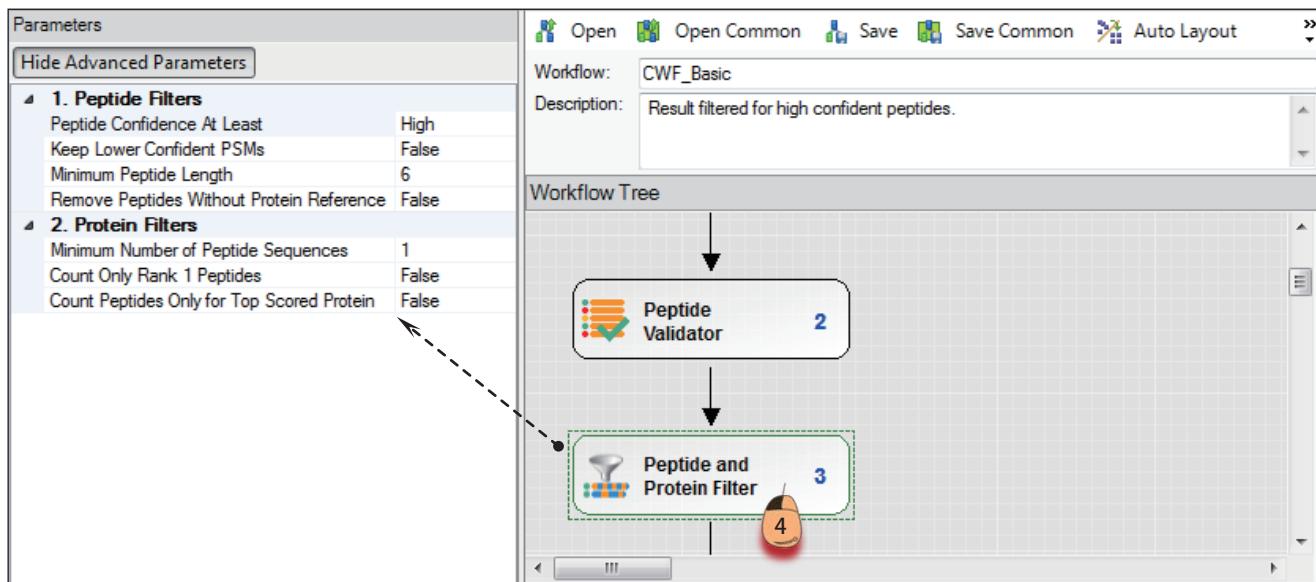
参数设置

1. 修饰肽段组装：对于修饰的肽段，其修饰位点的得分值大于 75 可信度较高，将其保留并组装成相应的修饰肽段。（分值范围：0-100）。
2. 修饰位点的呈现：默认认为在结果中呈现可信度最高的修饰位点。



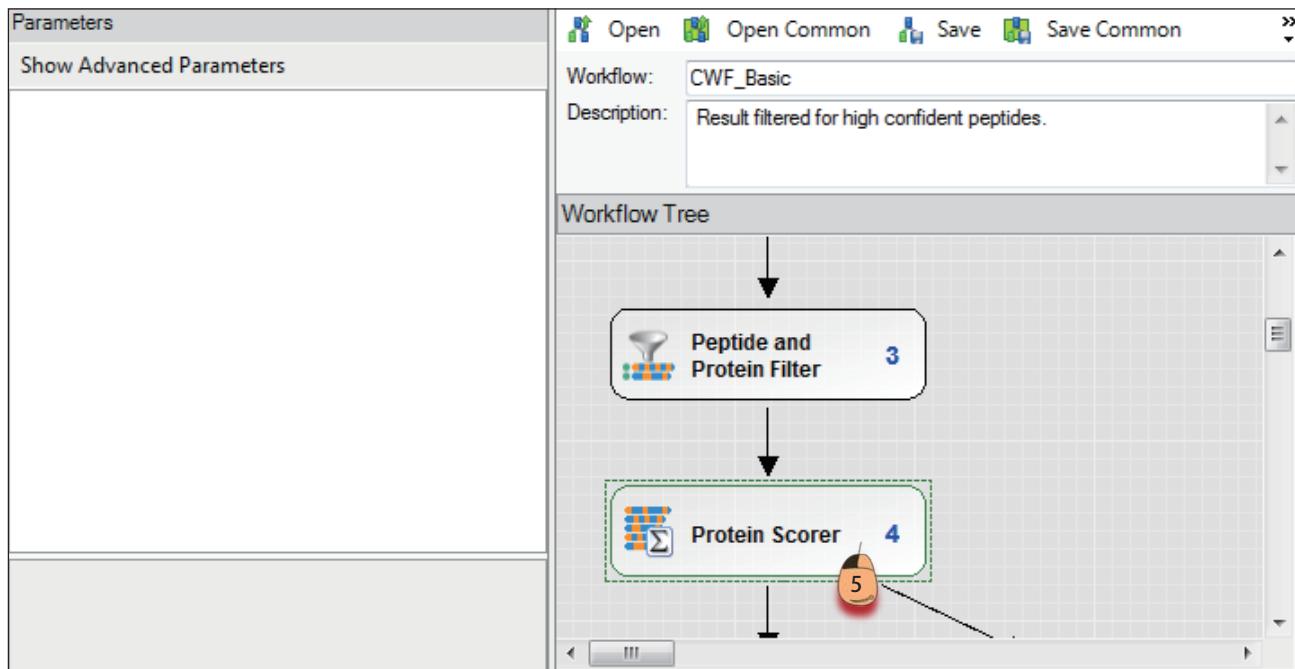
参数设置

- 肽段验证: 其中包含四种验证模式, 默认为 Automatic(控制肽段水平的错误率), 首先根据 Percolator 产生的 PEPs、q-value 和 FDR 设定的阈值将 PSM 的可信度归属为 High、Medium, 进而将肽段的可信度归属为 High、Medium;
- 具体验证设置: a) 默认设置为 False: PSM 的 FDR 计算是将 target 和 decoy 分开进行计算; b) 默认设置为 False: 对于没有进行 decoy 搜索的数据, 按照分数进行卡值。



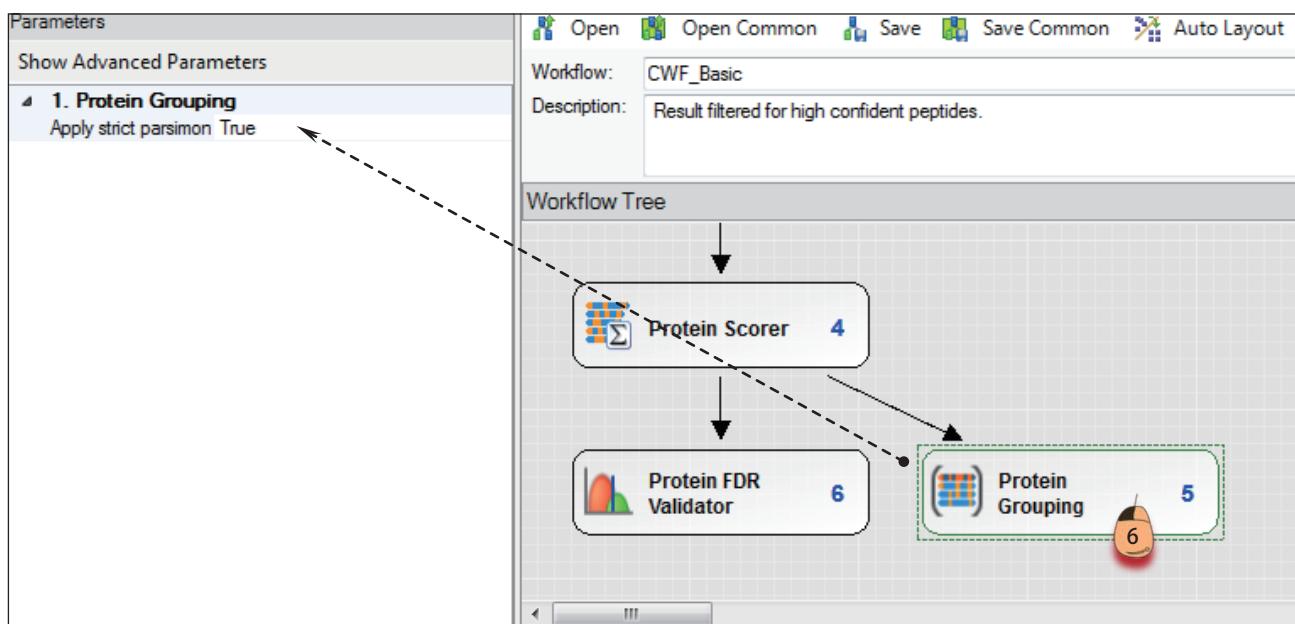
参数设置

- 肽段过滤: a) 根据肽段可信度过滤, 默认为 high; b) 比 a 中肽段可信度低的 PSM 是否保留, 默认为 False 不保留, 只保留与肽段可信度相同或更高的 PSM; c) 设置允许最小的肽段长度; d) 对没有匹配到蛋白的肽段和 PSM 是否去除, 默认为 False;
- 蛋白过滤: a) 最少的肽段数目, 默认为 1; b) 是否只记录排名为 1 的肽段, 默认为 False; c) 是否只记录归属到得分最高的蛋白对应的肽段, 默认为 False。



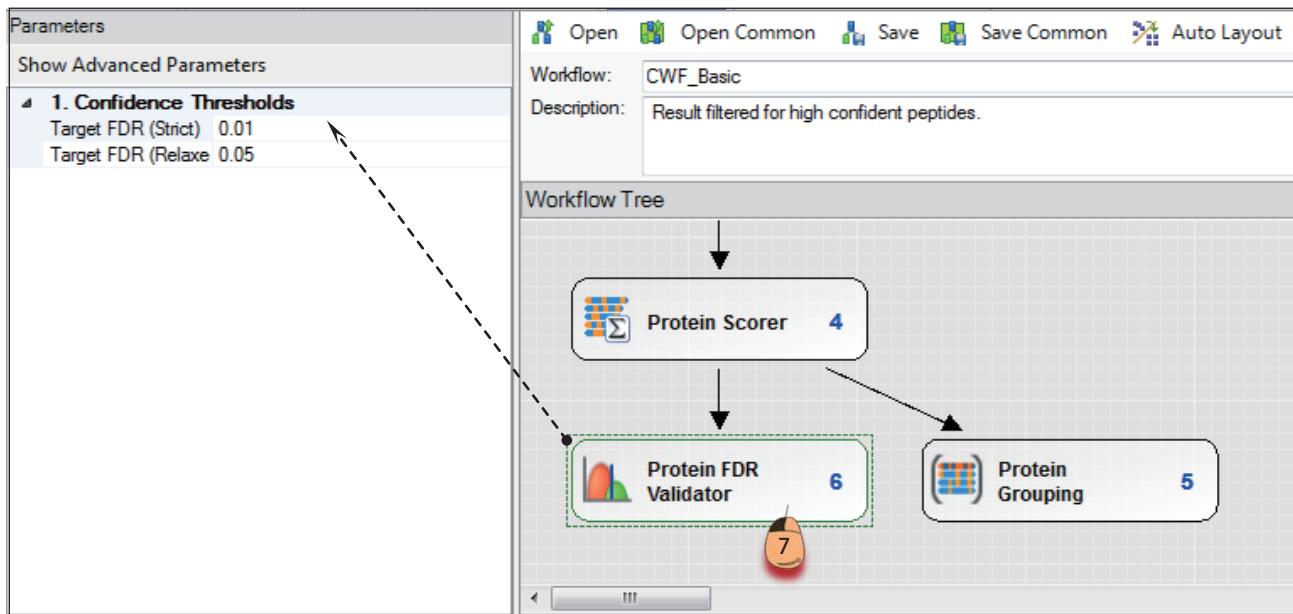
参数设置

此处没有任何需要用户设置的参数



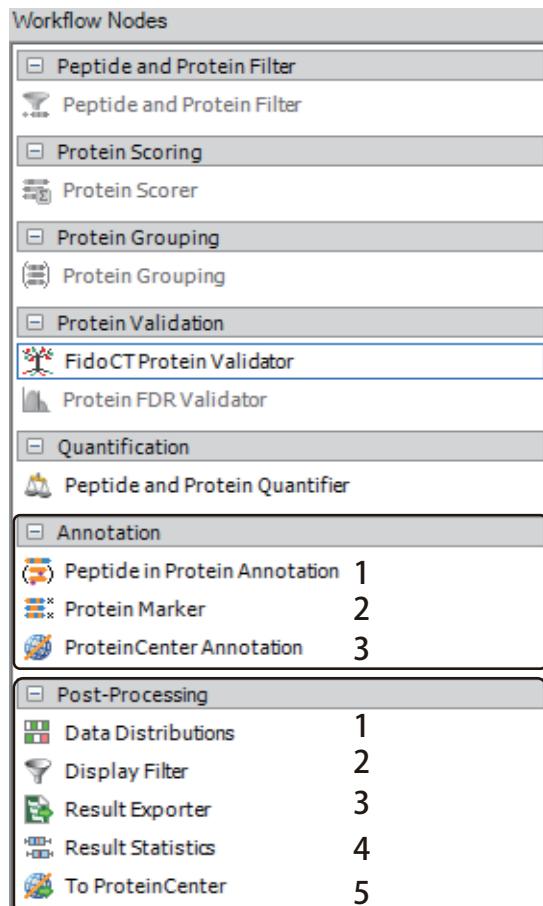
参数设置

蛋白分组: 默认为严格的解析原则即可



参数设置

蛋白假阳性率验证：假阳性率 = Number of decoy proteins/Number of target proteins (最终以 q-value 作为评价标准)，FDR (strict) <0.01 时 Protein confidence 为 high; FDR (relaxed) 0.05-0.01 时 Protein confidence 为 medium。备注：在使用此 node 时，前一个 node 需为 Protein scorer; 另一个用于蛋白假阳性率验证的模块 FidoCT Protein Validator 比 Protein FDR Validator 更加严格，得到的蛋白数目更少，因此谨慎使用。



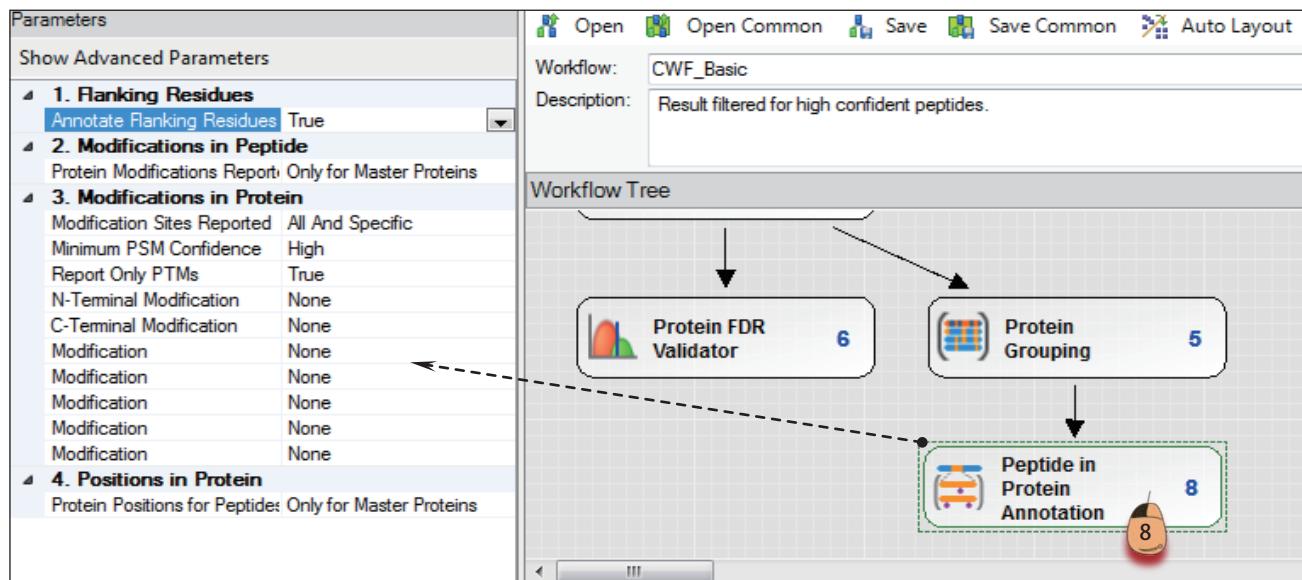
可选添加其他 Node: Annotation & Post-Processing

Annotation:

1. 注释肽段在蛋白中的具体位置、肽段前后连接氨基酸残基、蛋白序列中修饰位点及类型的信息。
2. 当添加多个 FASTA 文件时，此 Node 起作用，对来自不同 FASTA 的蛋白进行标注。
3. 对鉴定到的蛋白进行 GO、Pfam 和 Gene ID 注释（一年保修内可免费使用此功能，使用此模块时需要联网）。

Post-Processing:

1. 注释蛋白在不同文件中的分布情况，当对蛋白或肽段进行定量时，可在不同样品中呈现 ratio 或 area 分布，同时表达量的高低会用不同的颜色展示。
2. 设置过滤条件，进行结果展示。使用默认设置时，结果将自动呈现 master protein。
3. 将鉴定结果自动导出 txt 文件。
4. 对结果进行统计学分析。
5. 上传结果到 ProteinCenter。



参数设置

1. 默认为 true，呈现肽段前后连接氨基酸残基。

Peptides			
	Checked	Confidence	Annotated Sequence
46	<input type="checkbox"/>		[K].NVPQVVNVQELK.[N]
47	<input type="checkbox"/>		[K].RPLDDGVGNQLGALVHQR.[T]
48	<input type="checkbox"/>		[.].MMGHRPVLVLSQNTK.[R]
49	<input type="checkbox"/>		[K].NNLCPSGSNIISNLFK.[E]
50	<input type="checkbox"/>		[R].FSPDGELYASGSEDGTLR.[L]

2. 对于肽段的修饰，只呈现该肽段在 master protein 中的修饰信息，且 PSM 的 confidence 为 High。

Peptides					
	Checked	Confidence	Annotated Sequence	Modifications	Modifications in Master Proteins
46	<input type="checkbox"/>		[K].NVPQVVNVQELK.[N]		
47	<input type="checkbox"/>		[K].RPLDDGVGNQLGALVHQR.[T]		
48	<input type="checkbox"/>		[.].MMGHRPVLVLSQNTK.[R]	1×Acetyl [N-Term]	P49368 1×Acetyl [N-Term]
49	<input type="checkbox"/>		[K].NNLCPSGSNIISNLFK.[E]	1×Carbamidomethyl [C4]	

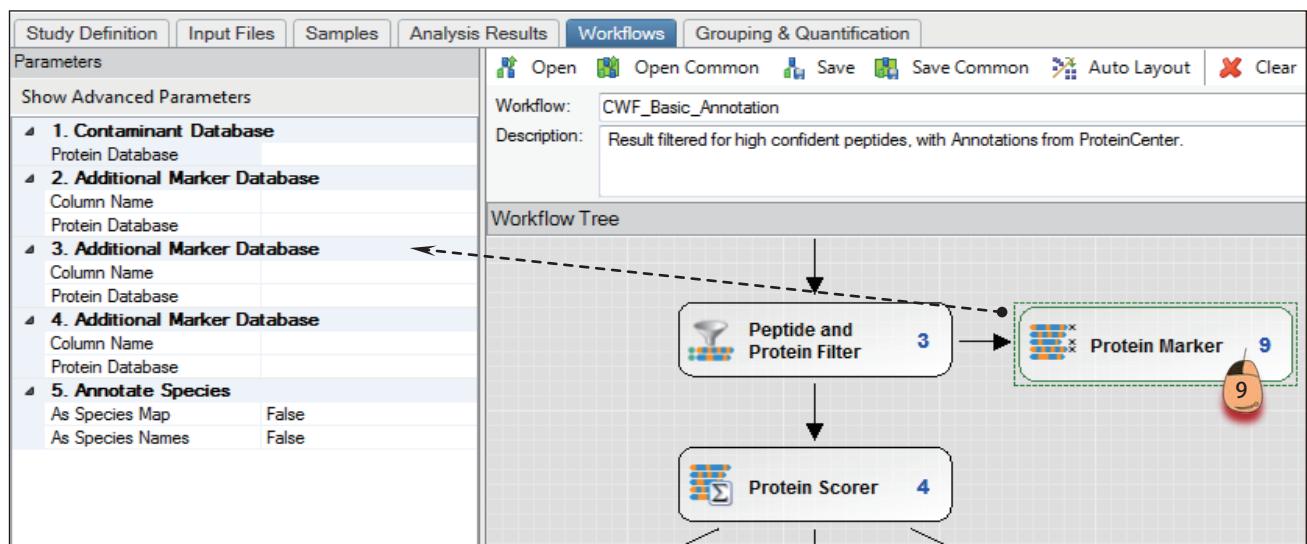
3. 对于蛋白的修饰，呈现该蛋白的所有修饰位点信息。

Proteins					
	Checked	Confidence	Annotated Sequence	Modifications	Modifications in Master Proteins
46	<input type="checkbox"/>		[K].NVPQVVNVQELK.[N]		
47	<input type="checkbox"/>		[K].RPLDDGVGNQLGALVHQR.[T]		
48	<input type="checkbox"/>		[.].MMGHRPVLVLSQNTK.[R]	1×Acetyl [N-Term]	P49368 1×Acetyl [N-Term]
49	<input type="checkbox"/>		[K].NNLCPSGSNIISNLFK.[E]	1×Carbamidomethyl [C4]	

4. 呈现肽段在蛋白中的具体位置：默认为只呈现肽段在 master protein 中的位置。

Peptides (filtered) Protein Groups Peptide Groups PSMs MS/MS Spectrum Info Result Statistics													
#	Checked	Confidence	Annotated Sequence	Modificati	Modificatio	Quality PEP	Quality q-value	# Protein Groups	# Proteins	# PSMs	Master Protein Accessions	Positions in Master Proteins	
46	<input type="checkbox"/>		[K]NVPQVVNVQELK[N]			2.76e-05	0	1	1	1	P35658	P35658 [1118-1129]	
47	<input type="checkbox"/>		[K]RPLDDGVGNQLGALVHQR[T]			1.35e-10	0	1	1	1	Q96124	Q96124 [58-75]	
48	<input type="checkbox"/>		[J]MMGHRPVVLVLSQNTK[R]	1xAcetyl	P49368 1:	2.36e-09	0	1	1	2	P49368	P49368 [1-15]	
49	<input type="checkbox"/>		[K]NNLCPGNSNIISNLFK[E]	1xCarba		1.62e-07	0	1	1	1	P60033	P60033 [172-187]	
50	<input type="checkbox"/>		[R]FSPDGELYASGSEDTLR[L]			1.97e-09	0	1	1	1	Q9Y3F4	Q9Y3F4 [273-290]	
51	<input type="checkbox"/>		[R]QYTSPPEIDAQQLQAEK[I]			5.17e-08	0	1	1	2	Q13442	Q13442 [16-31]	

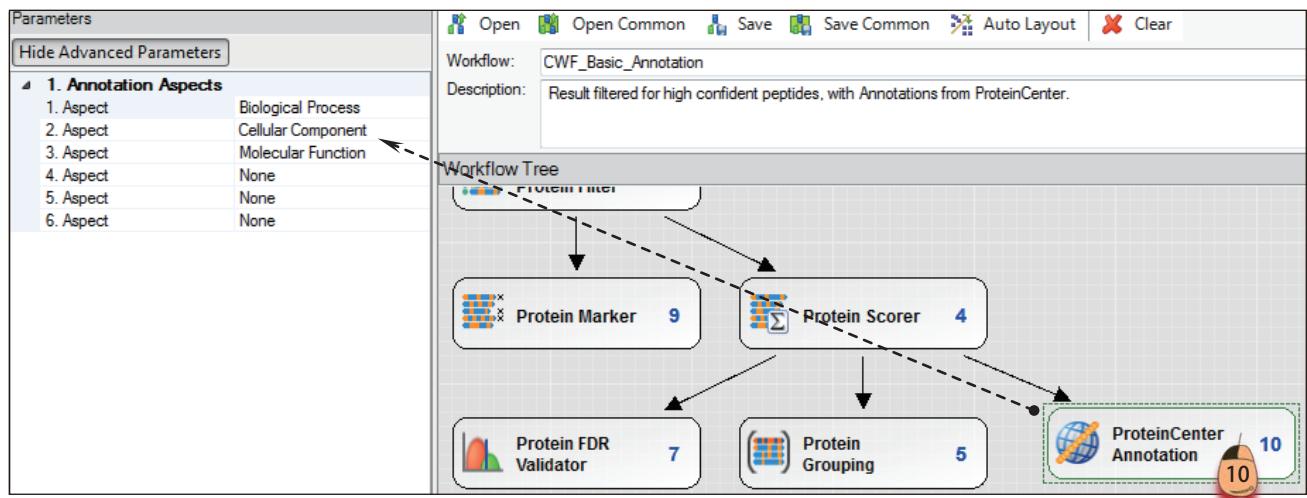
Protein Marker: 鉴定和定量蛋白的来源会根据添加的蛋白库进行标注 (当添加多个 Fasta 数据库时, 必须使用此 node)。



参数设置

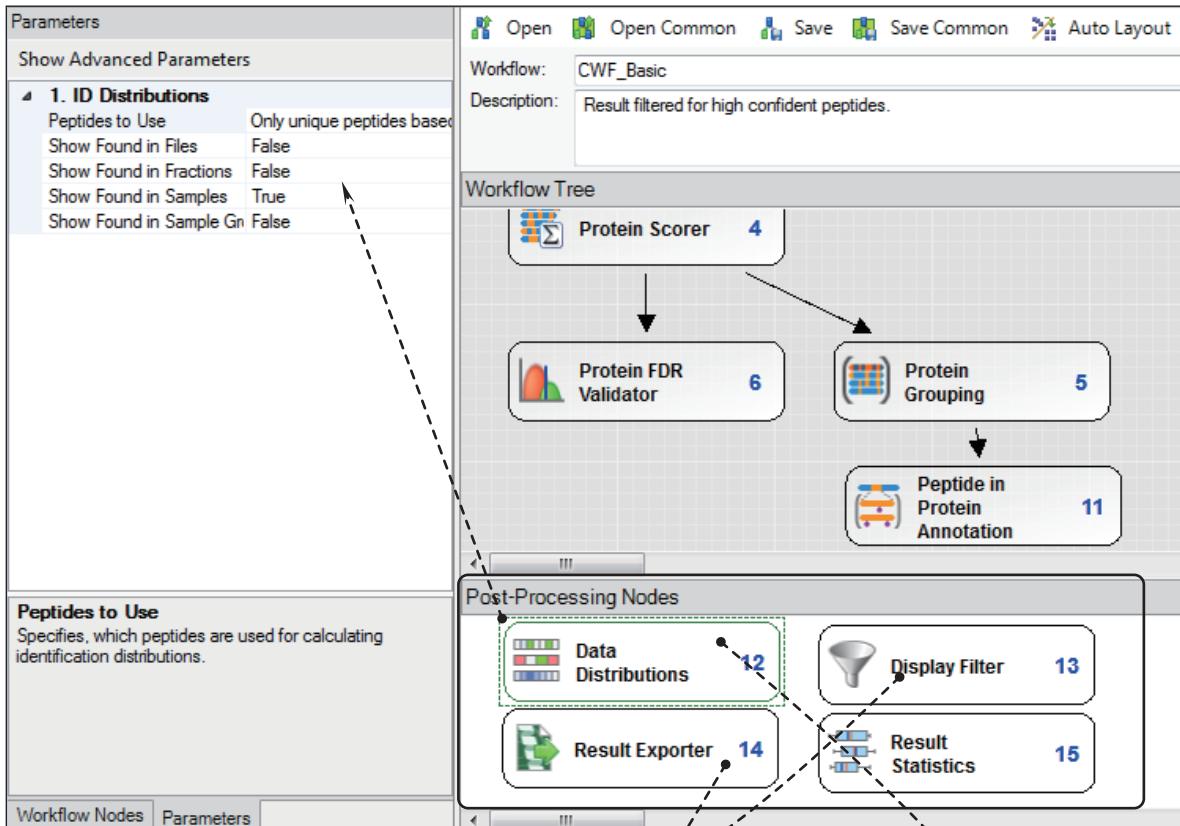
1. 将污染蛋白库 (包含角蛋白、胰酶等常见污染蛋白) 添加到 PD 的 Fasta 数据库中, 在此即可选择相应的污染蛋白库, 在结果中将展示哪些蛋白是污染蛋白。
- 2-4: 可添加多个蛋白数据库, 并且进行相应的命名, 结果中将展示蛋白的相应来源。
5. 默认为不展示物种的名称。

Protein Center Annotation: 蛋白功能注释 (一年保修内可免费使用此功能)



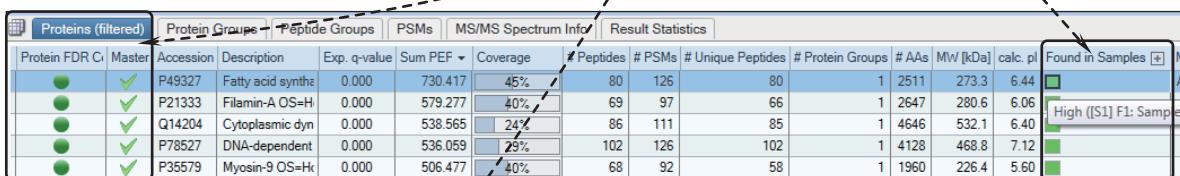
1. 默认为对蛋白参与的生物过程、细胞定位及分子功能进行注释, 同时也可以在 Protein Center 中下载其他注释功能。

Post-Processing Nodes:



参数设置

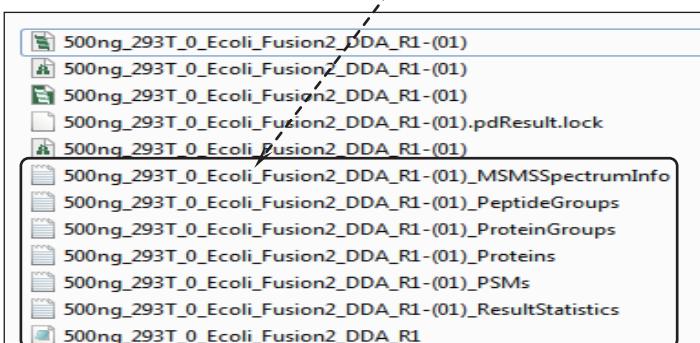
1. 当多个文件一起检索时, 使用 Data Distributions 可呈现蛋白在不同文件中的分布情况, 使用默认参数即可。



Proteins (filtered)		Protein Groups	Peptide Groups	PSMs	MS/MS Spectrum Info	Result Statistics
Protein FDR C	Master	Accession	Description	Exp. q-value	Sum PEF	Coverage
●	✓	P49327	Fatty acid synth	0.000	730.417	45%
●	✓	P21333	Filamin-A OS=H	0.000	579.277	40%
●	✓	Q14204	Cytoplasmic dyn	0.000	538.565	24%
●	✓	P78527	DNA-dependent	0.000	536.059	29%
●	✓	P35579	Myosin-9 OS=Hk	0.000	506.477	40%

2. Display Filter: 使用默认筛选设置, 打开结果时呈现 master protein。

3. Result Exporter: 默认同时将结果导出 txt 格式。



根据实际需求添加不同的 Post-processing node

4. 对搜库结果做统计学分析。

Parameters

Show Advanced Parameters

1. ID Distributions

Peptides to Use Only unique peptides based on sequence

Show Found in Files False

Show Found in Fractions False

Show Found in Samples True

Show Found in Sample Groups False

Workflow: CWF_Basic

Description: Result filtered for high confident peptides.

Workflow Tree

```

    graph TD
        4[Protein Scorer] --> 6[Protein FDR Validator]
        4 --> 5[Protein Grouping]
        5 --> 11[Peptide in Protein Annotation]
    
```

Post-Processing Nodes

```

    graph TD
        12[Data Distributions] --- 13[Display Filter]
        12 --- 14[Result Exporter]
        12 --- 15[Result Statistics]
    
```

Workflow Nodes Parameters

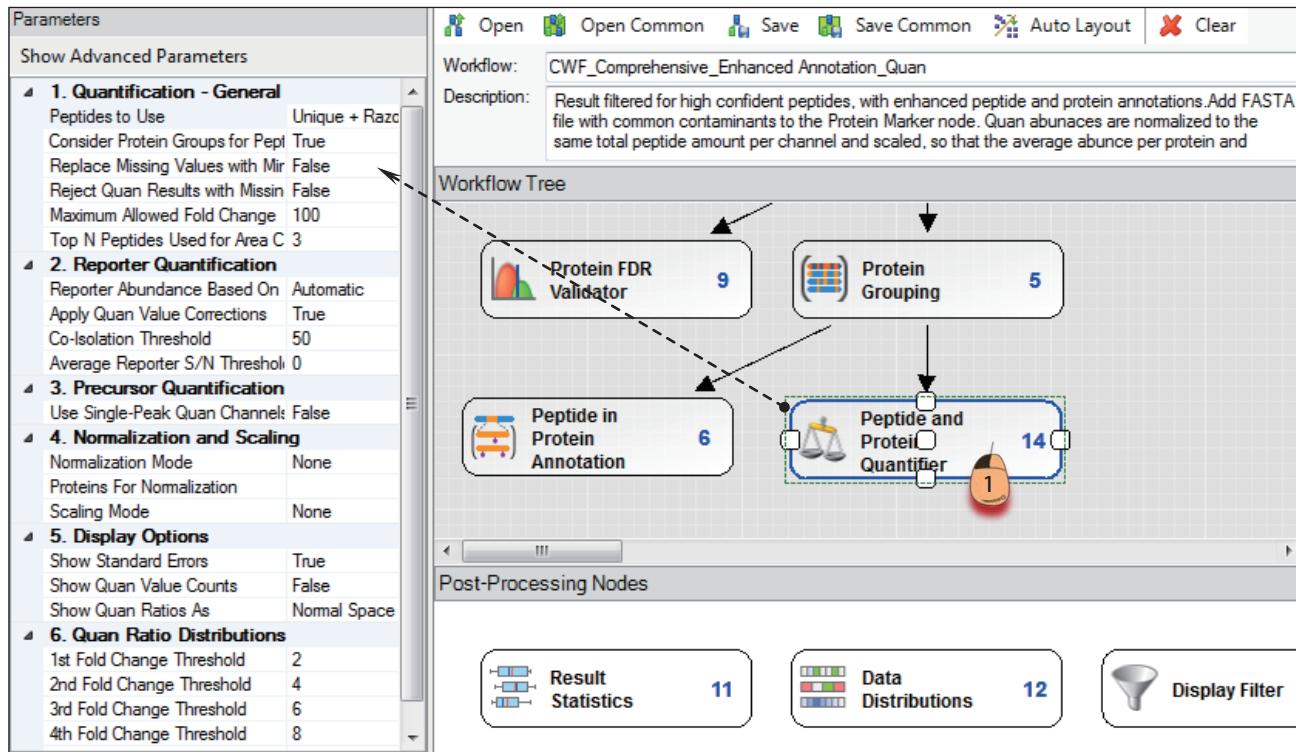
Peptides to Use

Specifies, which peptides are used for calculating identification distributions.

Result Statistics

	Name	Arithmetic Mean	Median	SD	IQR	CV	Sum	Minimum	Maximum	Count
1	Proteins - # AAs	613.2805	445	792.19445	461	1.29173	2704567	51	34350	4410
2	Proteins - # Decoy Proteins	22.09002	0	74.46881	1	3.37115	97417	0	450	4410
3	Proteins - # Peptides	5.08503	3	6.77298	5	1.33194	22425	1	102	4410
4	Proteins - # Peptides Sequest HT	5.08503	3	6.77298	5	1.33194	22425	1	102	4410
5	Proteins - # Protein Groups	1.04036	1	0.51311	0	0.4932	4588	0	10	4410
6	Proteins - # Protein Unique Peptides	4.22653	2	6.39848	4	1.51388	18639	0	102	4410
7	Proteins - # PSMs	7.24059	3	11.63628	7	1.60709	31931	1	126	4410
8	Proteins - # PSMs Sequest HT	7.24059	3	11.63628	7	1.60709	31931	1	126	4410
9	Proteins - # Unique Peptides	4.63741	2	6.45829	5	1.39265	20451	0	102	4410
10	Proteins - calc. pl	7.1685	6.9	1.66128	2.68	0.23175	31613.07	3.78	12.56	4410
11	Proteins - Coverage	15.27534	9.67742	15.17514	18.99195	0.99344	67364.23984	0.09094	89.62963	4410
12	Proteins - Coverage Sequest HT	15.27534	9.67742	15.17514	18.99195	0.99344	67364.23984	0.09094	89.62963	4410
13	Proteins - Exp. q-value	0.00521	0	0.01727	0.00028	3.31447	22.97641	0	0.10204	4410
14	Proteins - MW [kDa]	68.49985	49.9245	88.32608	51.37125	1.28943	302084.342	6.009	3813.652	4410
15	Proteins - Score Sequest HT	26.41234	9.5065	46.72529	25.20708	1.76907	116478.40733	0	540.68394	4410
16	Proteins - Sum PEP Score	32.43239	12.92064	52.2757	33.29005	1.61184	143026.86192	0.72515	730.41677	4410
17	Protein Groups - # Peptides	5.54093	3	7.14849	6	1.29012	20712	1	102	3738
18	Protein Groups - # Proteins	1.22739	1	1.55952	0	1.27059	4588	1	71	3738
19	Protein Groups - # PSMs	7.65195	3	11.80724	8	1.54304	28603	1	128	3738
20	Protein Groups - # Unique Peptides	5.18941	3	6.81759	5	1.31375	19398	1	102	3738

定量 node: 无论是基于母离子定量还是报告离子定量都需将 Peptide and Protein Quantifier node 加在 Protein Grouping node 后, 其他 node 参数同之前定性参数设置。



参数设置

1. 定量通用设置 (母离子及报告离子定量均适用) :

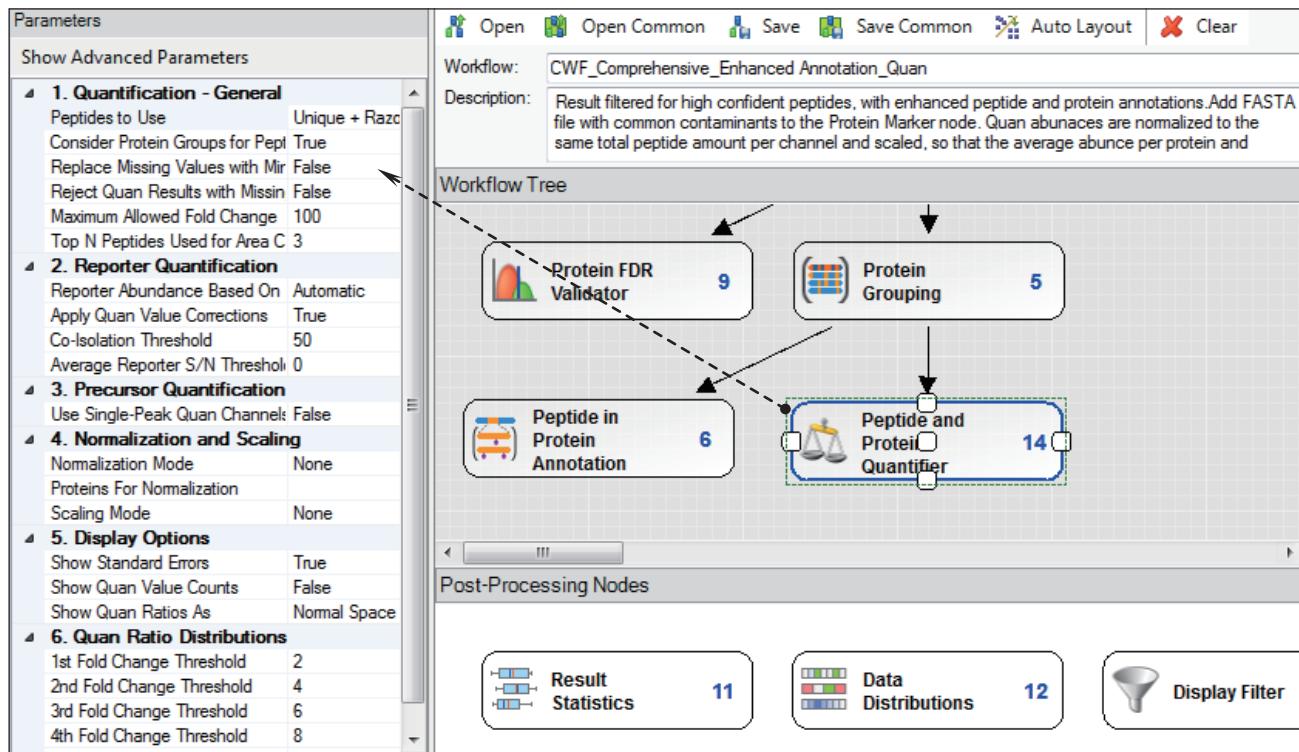
a) 用于定量的肽段: 大多数软件利用 Unique peptide 和 Razor peptide 做定量; b) 对于 Unique peptide 要考虑是否专属于一个 protein group, 默认为 true; c) 对于没有定量信息的蛋白是否使用最小值去替代, 默认为 False; d) 是否拒绝带有 missing channel 的定量结果, 默认 False; e) 允许最大的差异倍数, 默认 100; 选取 Top3 的肽段用于蛋白母离子峰面积定量; (使用默认参数即可)。

2. 报告离子定量:

a) 报告离子定量依据: 默认为 automatic, 当所有报告离子有信噪比时, 用信噪比定量, 否则用强度定量 (共分为信噪比、intensity 两种定量模式); b) 是否对报告离子定量比值进行校正 (TMT/ITRAQ 标记试剂的天然同位素丰度会影响报告离子的定量准确性), 当选择 True 时, 需要在定量方法中添加校正因子 (校正因子可以在标记试剂的说明书中查阅, TMT10 标记时, 必须使用校正因子); c) 允许最大的共隔离的干扰离子所占比例, 默认为 50%; 当设为 100% 时则认为不排除共隔离离子; d) 报告离子平均信噪比的阈值, 不严格要求定量结果时可设置为 0, 建议设为 10。

3. 母离子定量 (SILAC、二甲基化标记) :

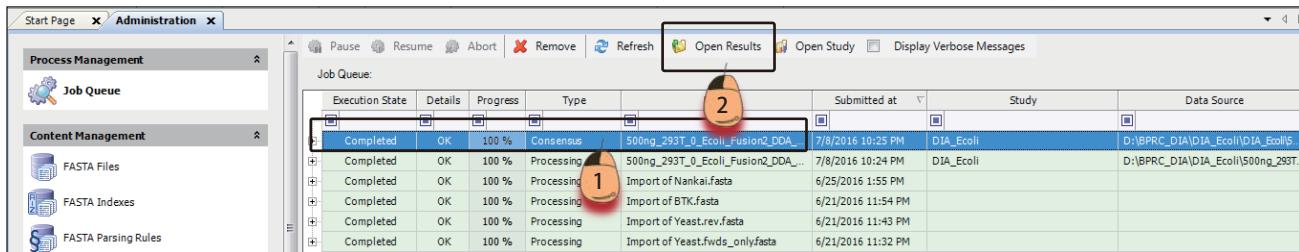
当只提取到一个同位素峰时 (不是一组同位素峰), 是否用于定量, 默认值 False。



- 定量结果的归一化与标尺化: a) 是否使用归一化对结果进行校正, 一般可以使用 “Total Peptide Amount” 进行归一化校正, 如果理论混合比例不为 1, 无需归一化校正, 可以选择 “None” ; b) 仅当 “Normalization Mode” 设置为 “Specific Protein Amount” 时使用, 需指定特定的 FASTA 数据库用于归一化, 该 FASTA 文件需要预先添加; c) 对定量结果进行标尺化, 有利于结果的热图展示, 可以选择 “On Channels Average (Per File)” , 保证每个文件中每个蛋白和肽段所有 channel 的平均值都为 100, 当选择 “On Control Channels Avg. (Per File)” 时, 保证每个文件中每个蛋白和肽段的 control channel 的平均值都为 100。
- a) 是否呈现标准误差, 默认 True; b) 是否呈现用于定量的 PSM 数目, 默认 False; c) 定量结果呈现形式: 常规数值或者 \log_2 值, 也可同时展示。
- PD 用不同的颜色强调不同的变化倍数, 变化倍数可自行定义 (如下图展示) :

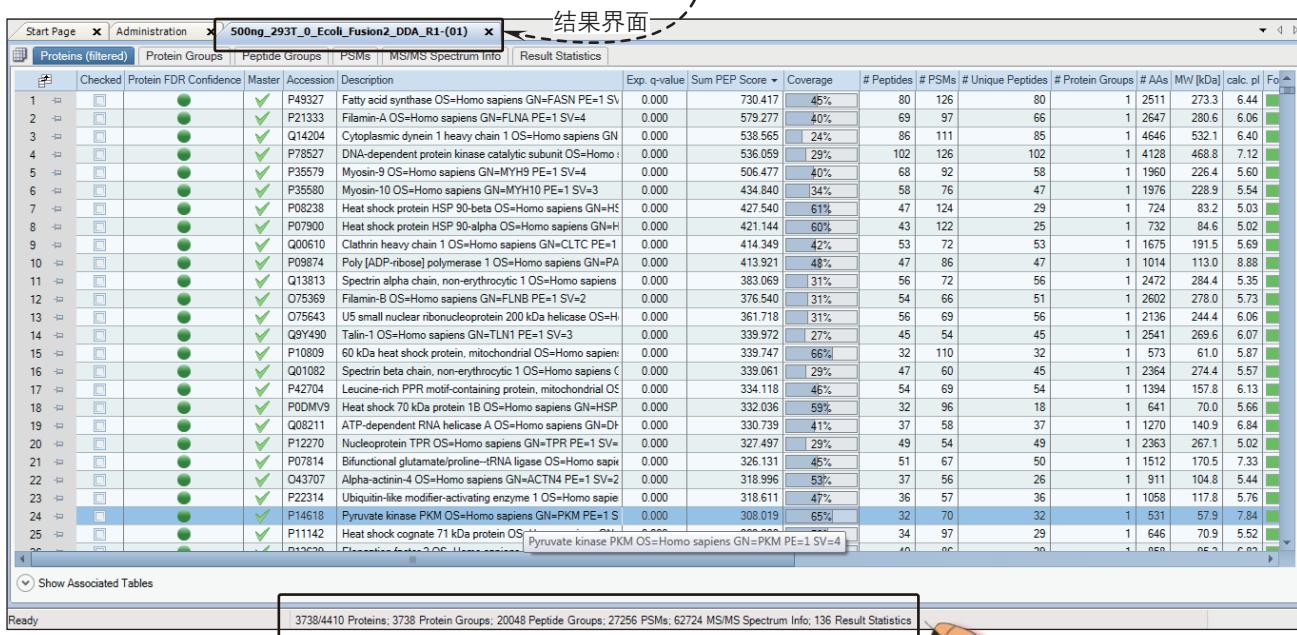
Highlight colors of the *n*th Fold Change Threshold parameters

Highlight color > value	Highlight color < 1/value
1st	1st
2nd	2nd
3rd	3rd
4th	4th
5th	5th



1. 在“Job Queue”中可双击目标文件的“Consensus”即可呈现结果界面。

2. 单击选中目标文件，再点击“Open Results”，也会呈现结果界面。



3. 结果界面的最下方对此结果进行统计：

Result Statistics: 只有在 consensus step 中 post-analysis 模块下添加“Result statistics”后在结果中才会出现，此处代表有 136 个简单的数据统计结果

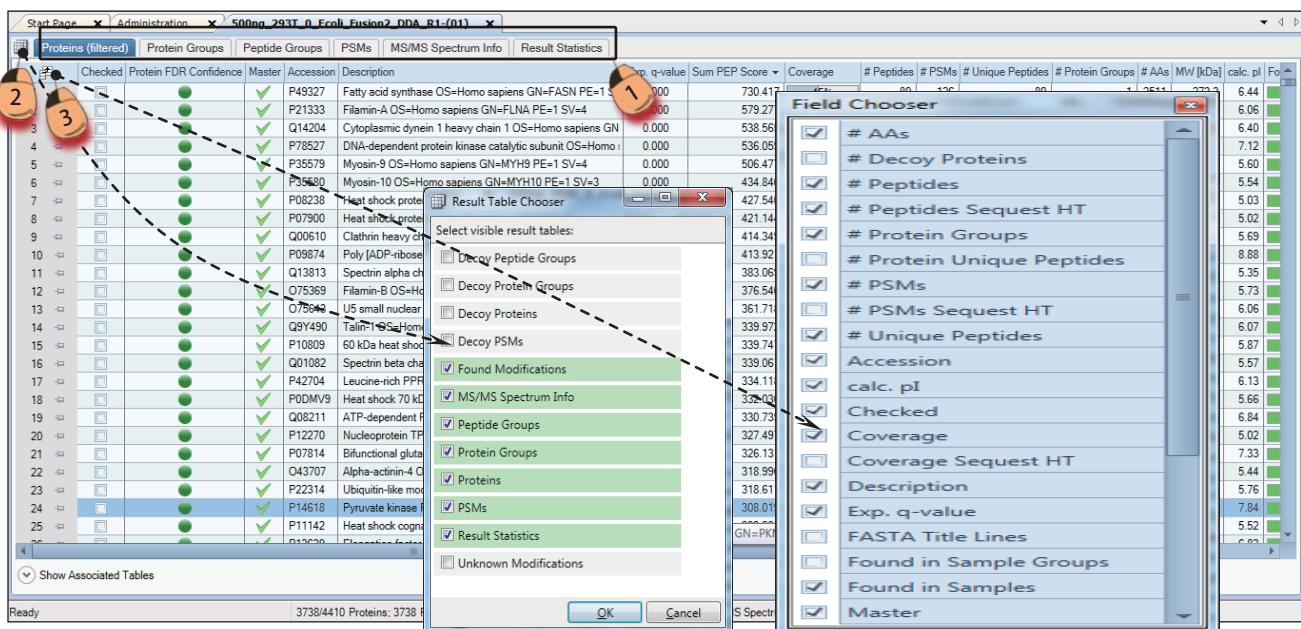
MS/MS Spectrum Info: 原始文件中所有通过 Spectrum Selector 模块筛选后送入 processing workflow 进行搜库的二级谱图数目

PSMs (Peptide Spectrum Match)：匹配到肽段的二级谱图数目

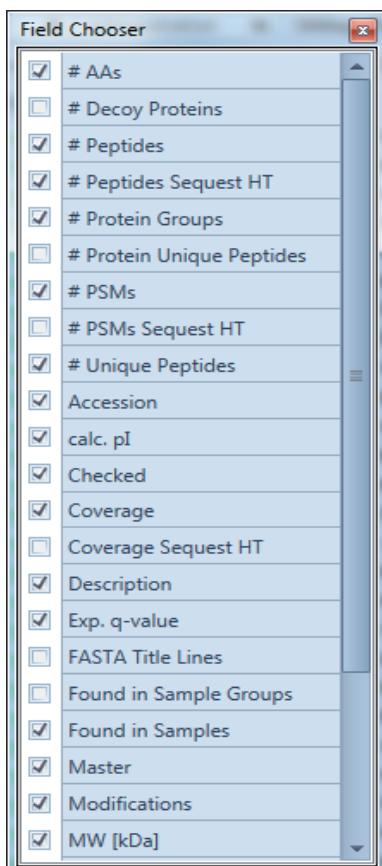
Peptide Groups: 将鉴定到的 PSM 进行组装后，得到非冗余的鉴定肽段数目

Protein Groups: 将鉴定到相同肽段的蛋白进行组装后，得到非冗余的鉴定蛋白数目

Proteins: 鉴定到的蛋白数目（注：如果在 consensus step 中 post-analysis 模块下添加“Display Filter”则对 Proteins 进行过滤，呈现 Proteins groups/Proteins 的比值，如图中 3738/4410）



- 结果中包含 Proteins、Protein Groups、Peptide Groups、PSMs、MS/MS Spectrum Info.、Result Statistics 等常用界面，根据需要进行相应查看。
- 点击左上图标，根据需要选择呈现其他信息界面。
- 在每个界面中点击左上图标，可根据需要选择待呈现的信息。



标题含义

#AAs: 氨基酸数目

#Decoy Proteins: 反库中鉴定到的蛋白数目

#Peptides: 鉴定到不同肽段的总数目

#Peptides Sequest HT: 使用 Sequest HT 搜索后鉴定到不同肽段的总数目 (注: 如只用 Sequest HT 搜索, # Peptides Sequest HT = # Peptides)

#Protein Groups: 组装到不同蛋白质组的数目, 一般为 1

#Protein Unique Peptides: master protein 含有的特异性肽段的数目

#PSMs: 肽段匹配到二级谱图的数目

#PSMs Sequest HT: 使用 Sequest HT 搜索后鉴定到肽段匹配到二级谱图的数目 (注: 如只用 Sequest HT 搜索, #PSMs Sequest HT = #PSMs)

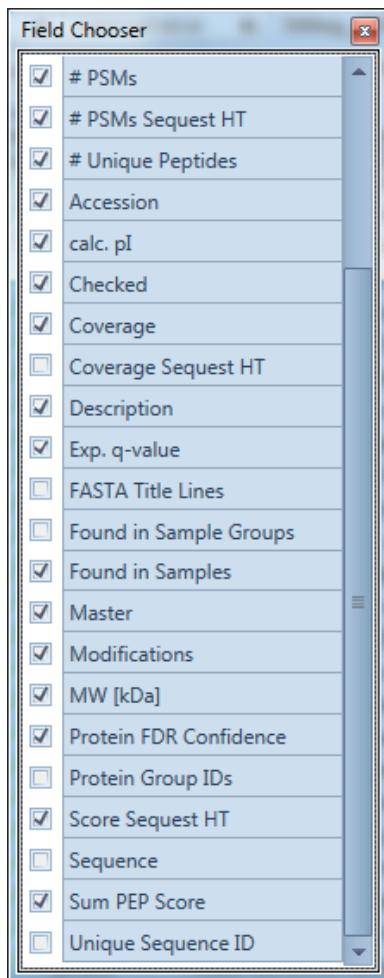
#Unique Peptides: protein group 中含有的特异性肽段数目, 一般大于等于 "#Protein Unique Peptides" 中的数目

Accession: 不同蛋白特征性的编号

calc.pl: 蛋白等电点

Checked: 可根据需要将感兴趣蛋白进行挑选并做标记, 进而导出目标蛋白

Coverage: 蛋白序列中氨基酸覆盖度



Description: 蛋白名称的详细描述

Exp.q-value: 当使用 Mascot 搜索时, 其匹配度用 q 值来代表, 值越小代表可信度越高

FASTA Title Lines: 蛋白在相应 FASTA 中的描述

Found in Sample Groups: 不同组分文件合并搜索时, 显示蛋白在不同文件组中的鉴定情况 (如需展示, 需要在 Consensus 中 Post-analysis 模块下添加 data distribution 模块)

Found in Samples: 不同文件合并搜索时, 展示蛋白在不同文件中的鉴定情况 (如需展示, 需要在 Consensus 中 Post-analysis 模块下添加 data distribution 模块), 比 “Found in Sample Groups” 常用

Master: 在 protein group 中标注是否为主要蛋白

Modifications: 显示蛋白鉴定到的修饰 (如需展示, 需要在 Consensus 中添加 “peptide in protein annotation”)

MW (kDa): 蛋白分子量

Protein FDR Confidence: 蛋白可信度

Protein Group IDs: 在 Protein Group 页面的编号

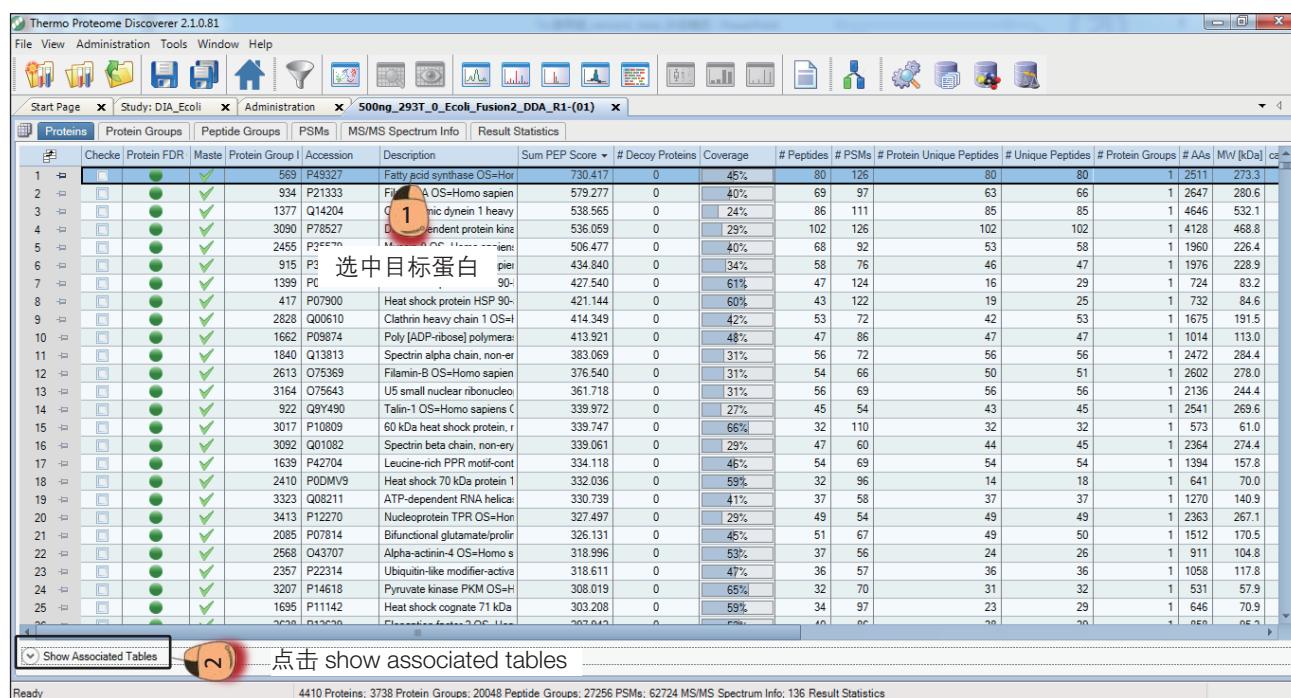
Score Sequest HT: Sequest HT 对蛋白匹配度的打分, 是肽段分数的加和, 分数越高可信度越高

Sequence: 展示蛋白氨基酸序列

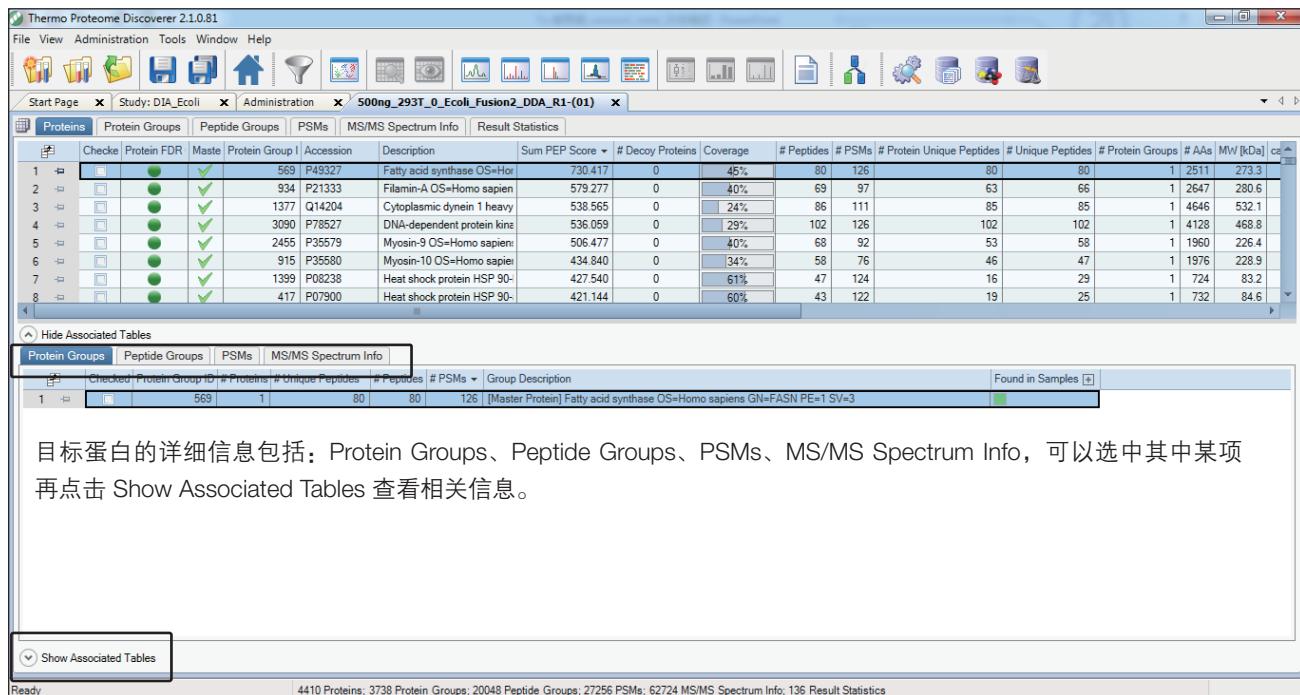
Sum PEP Score: 所有 PSM 的 PEP 值负对数的加和

Unique Sequence ID: 蛋白特有的序列号, 便于后续数据处理 (根据序列号去除冗余蛋白)

查看目标蛋白的详细信息

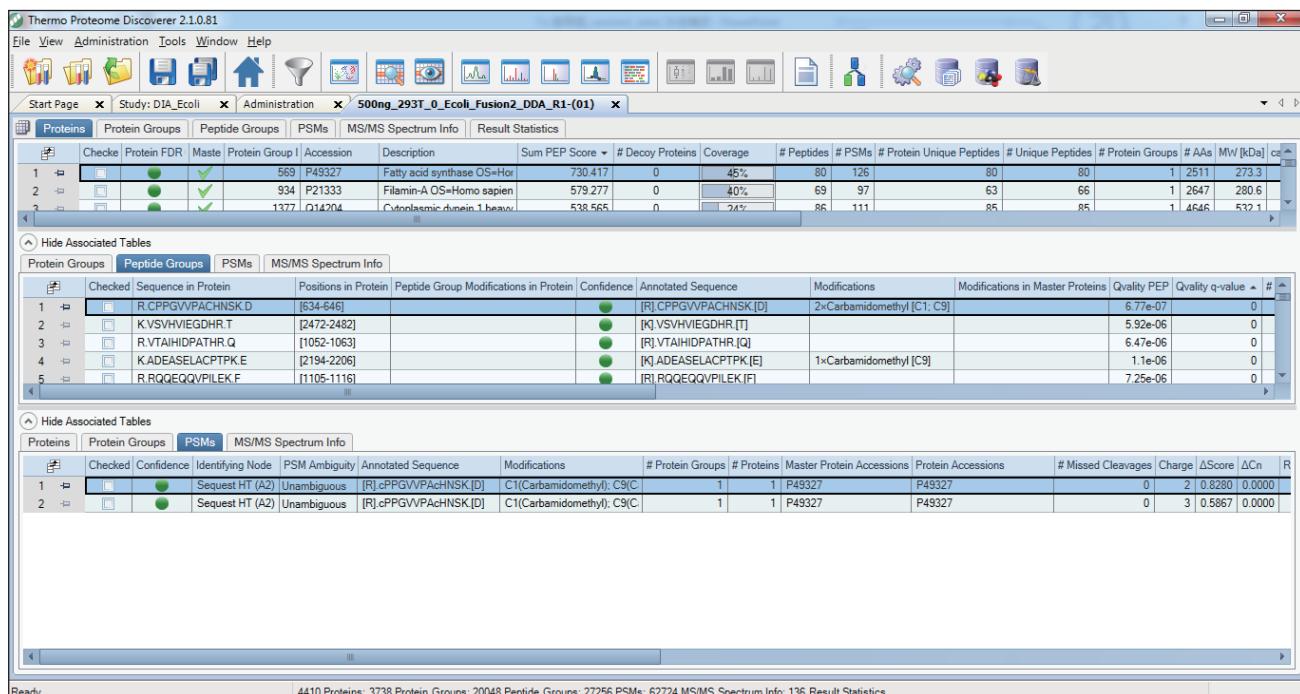


点击 show associated tables

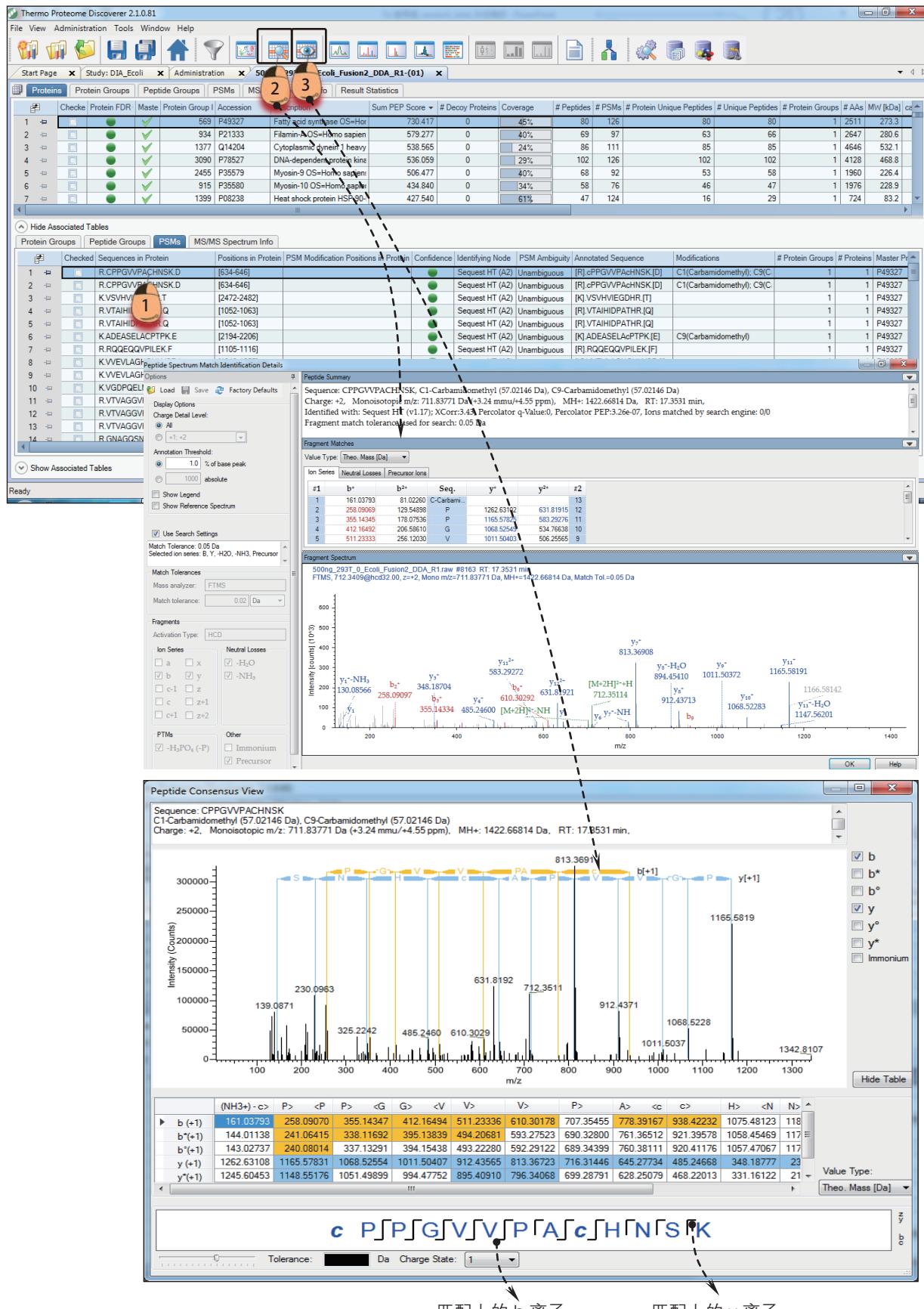


目标蛋白的详细信息包括：Protein Groups、Peptide Groups、PSMs、MS/MS Spectrum Info，可以选中其中某项再点击 Show Associated Tables 查看相关信息。

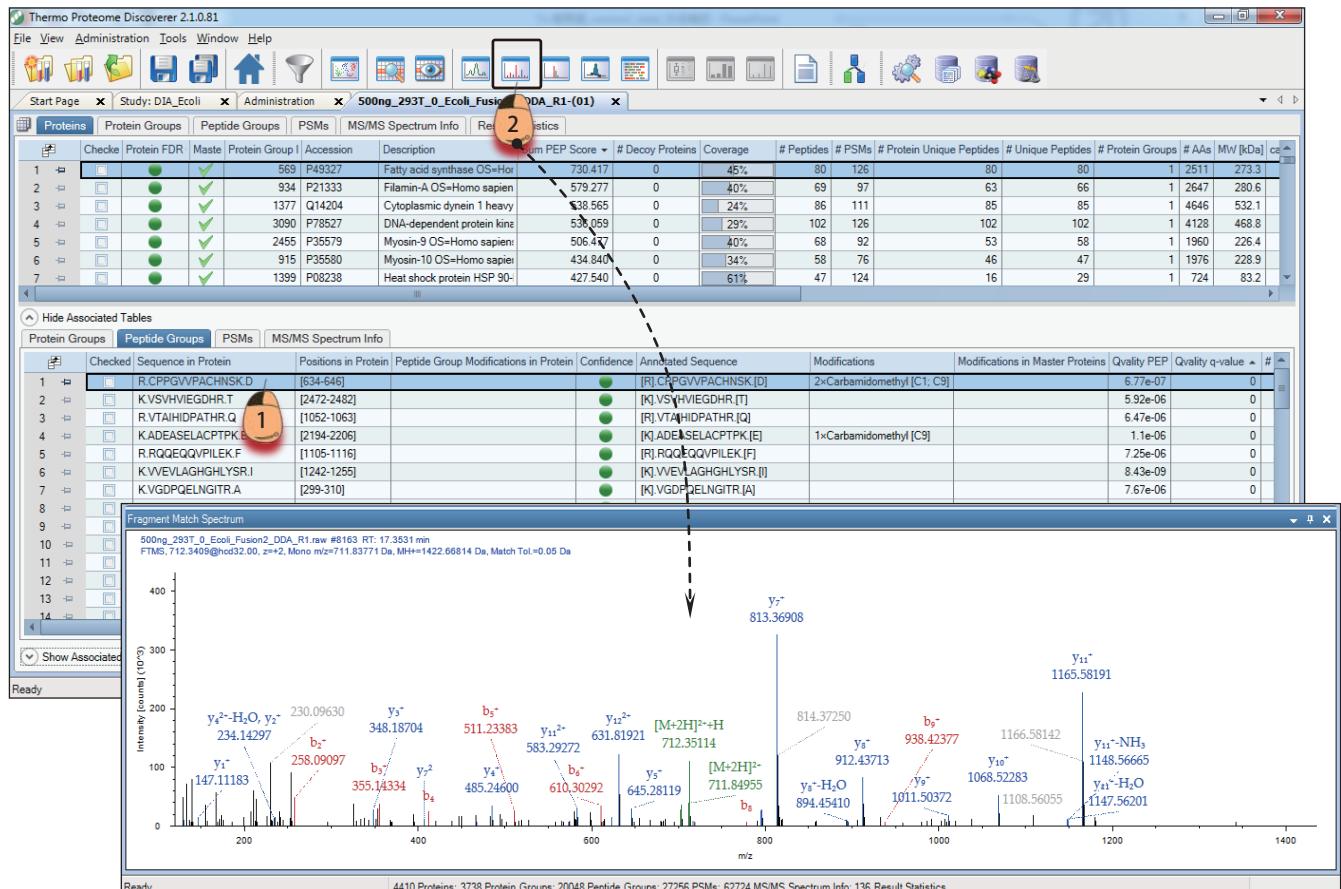
例如：查看选中蛋白的 Peptide Groups 中的目标肽段，同时可以查看目标肽段的 PSM 信息。



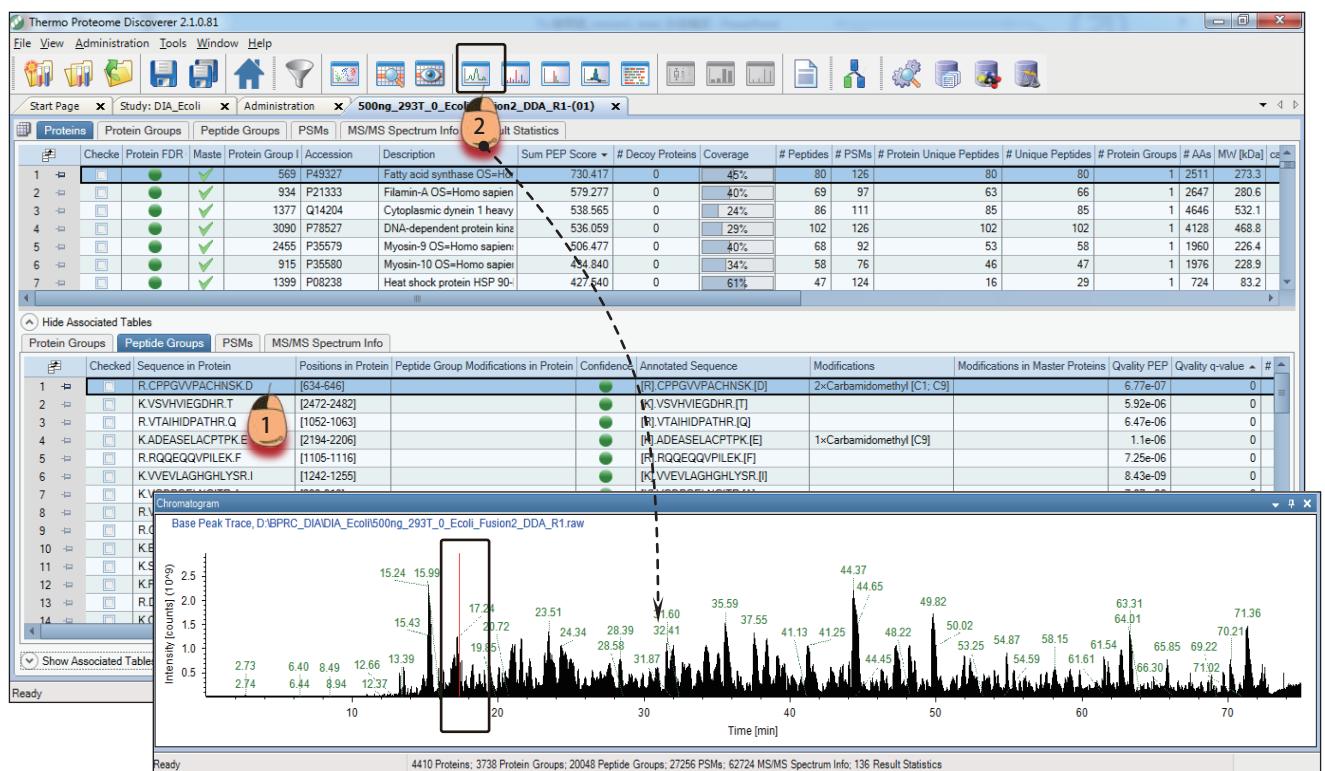
展示目标肽段的碎片离子



展示目标肽段的碎片离子

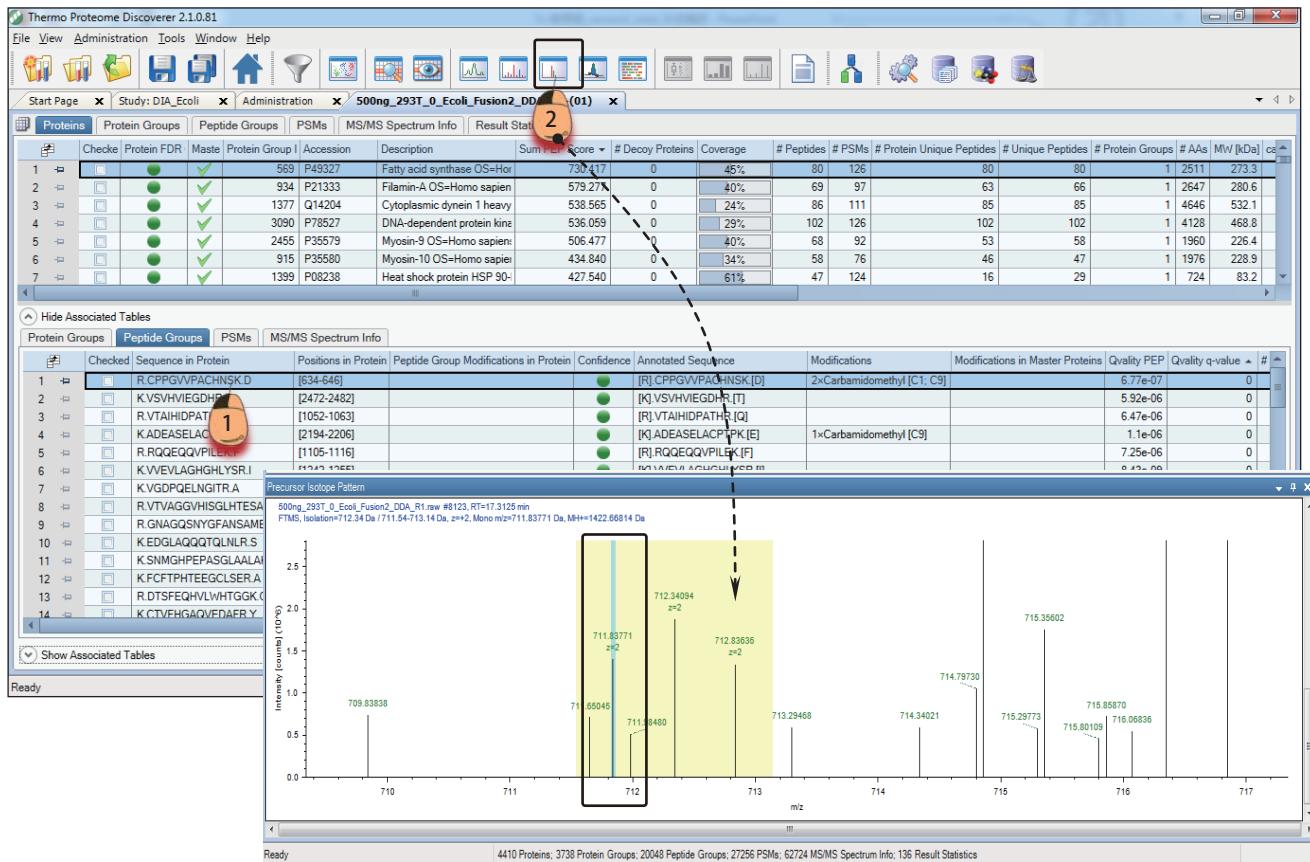


展示目标肽段在色谱图中的位置



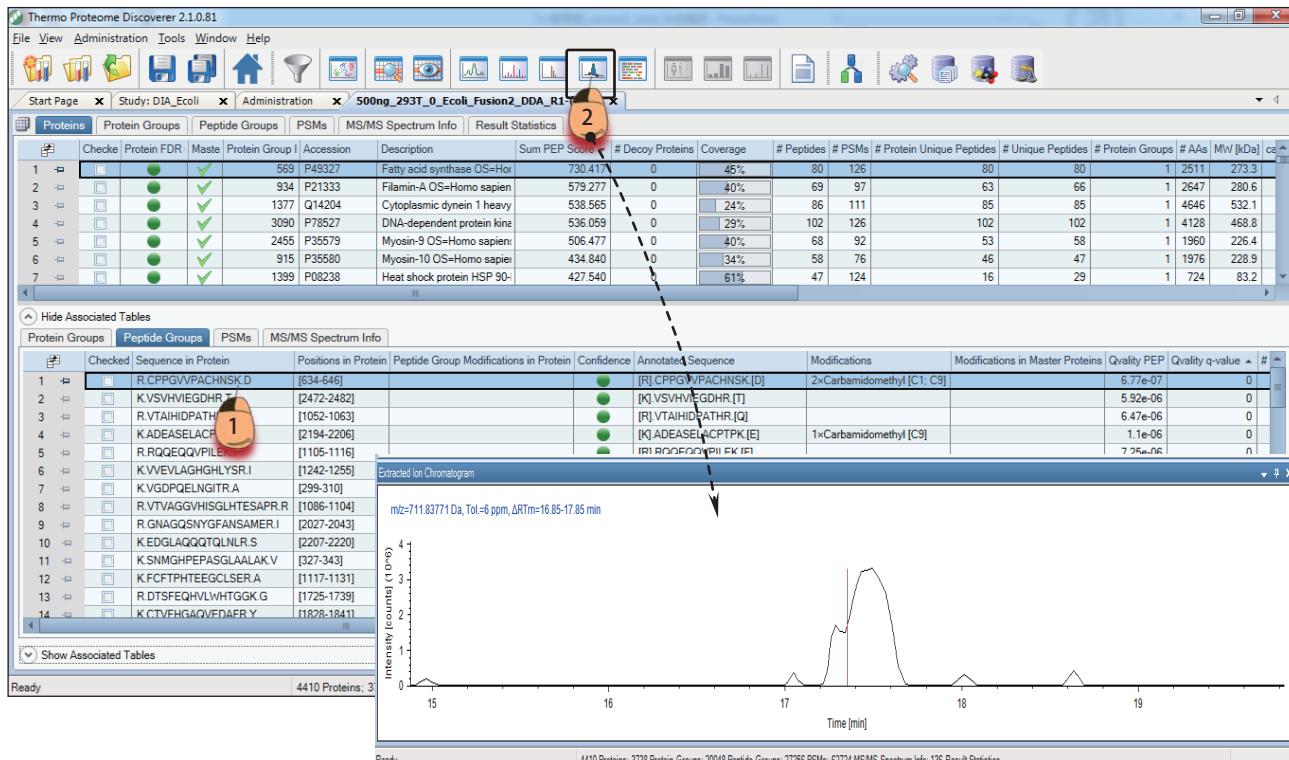
目标肽段在总离子流色谱图中的位置

展示目标肽段的母离子同位素峰分布



PD 软件匹配到目标肽段的母离子及其同位素峰分布

展示目标肽段的提取离子流色谱图



不同蛋白的氨基酸序列比较

选中待比较的不同蛋白, 例如选中第 1 和第 2 两个不同蛋白

彩色字体：代表该段序列对应的肽段被鉴定到，绿色字体表示肽段可信度为 high，黄色字体为 medium，红色字体为 low；蛋白序列有差异的部分以高亮背景显示，同时选择的蛋白序列最多不能超过 10000 个氨基酸。

過濾篩選

展示蛋白序列的详细信息

2

1

选中目标蛋白

Coverage: 45.08%

Found Modifications:

A: Acetyl (Nterm)
C: Carbamidomethyl (C)

Sequence:

1 A H E R V V I A G M S G H P E S E N L Q R P W D N L L G V D M V T D D O R R W P A G I L Y G L F R P S C K L K D L S P F D A S T F G V M P R Q C A R T H D P Q R L L L E V I T E A I V D G S I N P D S L

101 C P G T H T G W W V G V S G S E T S E A L S R D P T E V G Y S M V G C R A M H N R L S F F D F R G P S I A L D T A C S S I I M A L Q N A Y Q A I H S Q C P A A I V G I N V L L K P N T S V F

201 L R G H M L S P E G T C K A F D T A G N G V C R S E C O V V A V L T H Y S I A R R Y V A T I L Q H T H Y D G F R B Q C V T F F S C D I Q E Q L I R S I V Q S A C V A P E S T E V I E A H G T O T V K G

301 D P C B L I N G I T T I N 11 21 31 41 51 61 71 81 91

401 H E I L S P R T P P P 41

501 T Q R I C G C L S L M 51

601 L A A Y H R Q C I K E A R H L P P A M A V G L S W E E C K Q C P P O V P V P A C H N S H D T V T I S G P Q A P V F E F V E Q L R P E Q V P A K E V R T G M A F R H S Y M H A I A P P I L Q E L K 61

701 V I P R E P K P R S A P H L S T S I P R A Q M H S S L A R T S A Z Y V N V N N L V S P V L Q R A L M H V P E H A V N V L E I A P H A L L Q A V L K P G L M T S C T I P L M Q X D R S 71

801 G R H H L S C I D A N P H A L F P P V E F P A P R O P T L I S P L I N H C S L A M D V P A A E D F P H G S O S P S A A I Y N O T S S E S P O H V L D M T L D G R V L F P A G I Y L S I W K T L A 81

901 P A L G L G V E O L P V V F E D V V Y L N A T I L P K T C T V S I L E V R L L E A S P A F E S V E N G N L V U S G K Y O W D D P D P R L F D H P E S P I N P T E P L F L A Q V Y K E L R L R G D V 91

OK Help

展示数据结果统计图

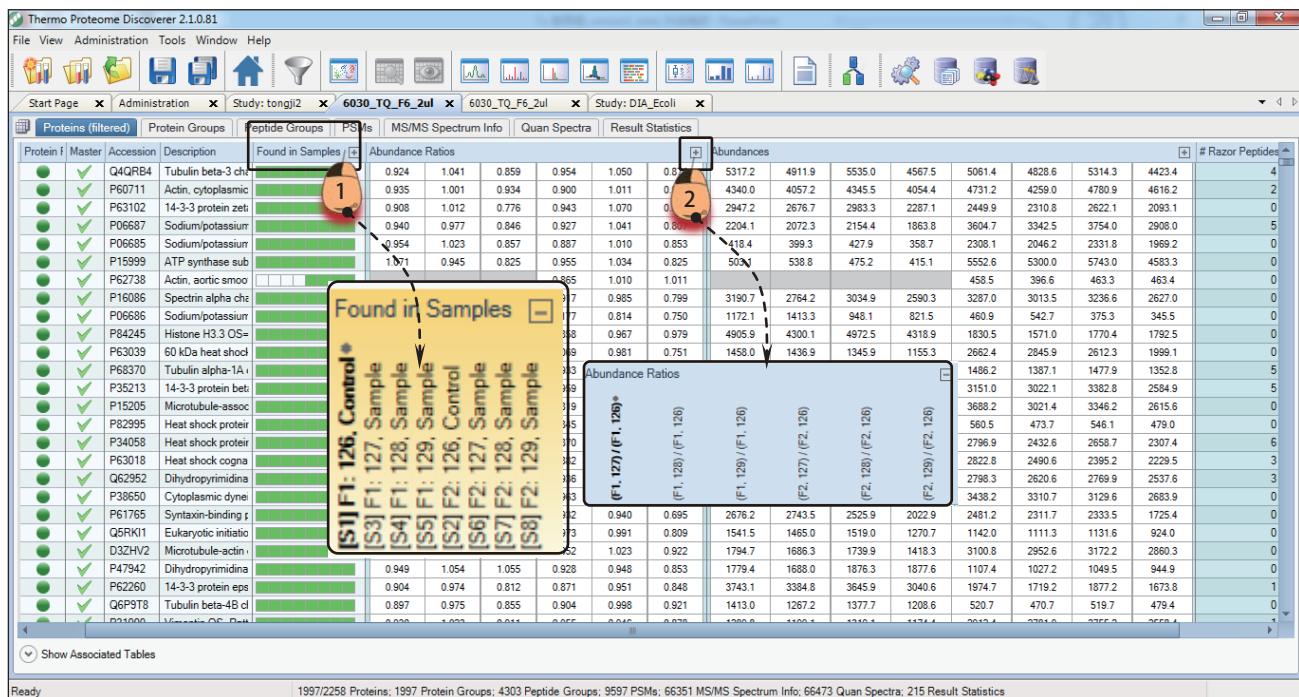
1

2

可分别进行简单的散点图、柱状图、韦恩图的绘制

以 TMT 定量为例

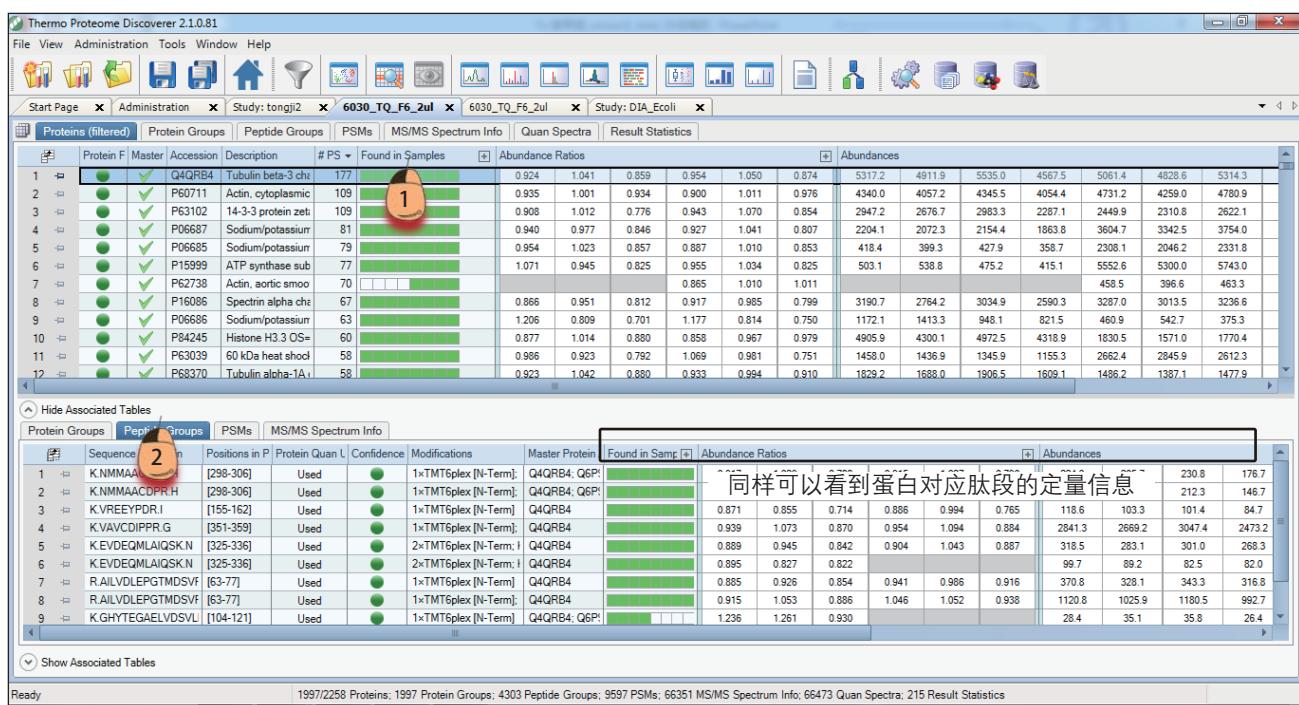
展示报告离子定量结果



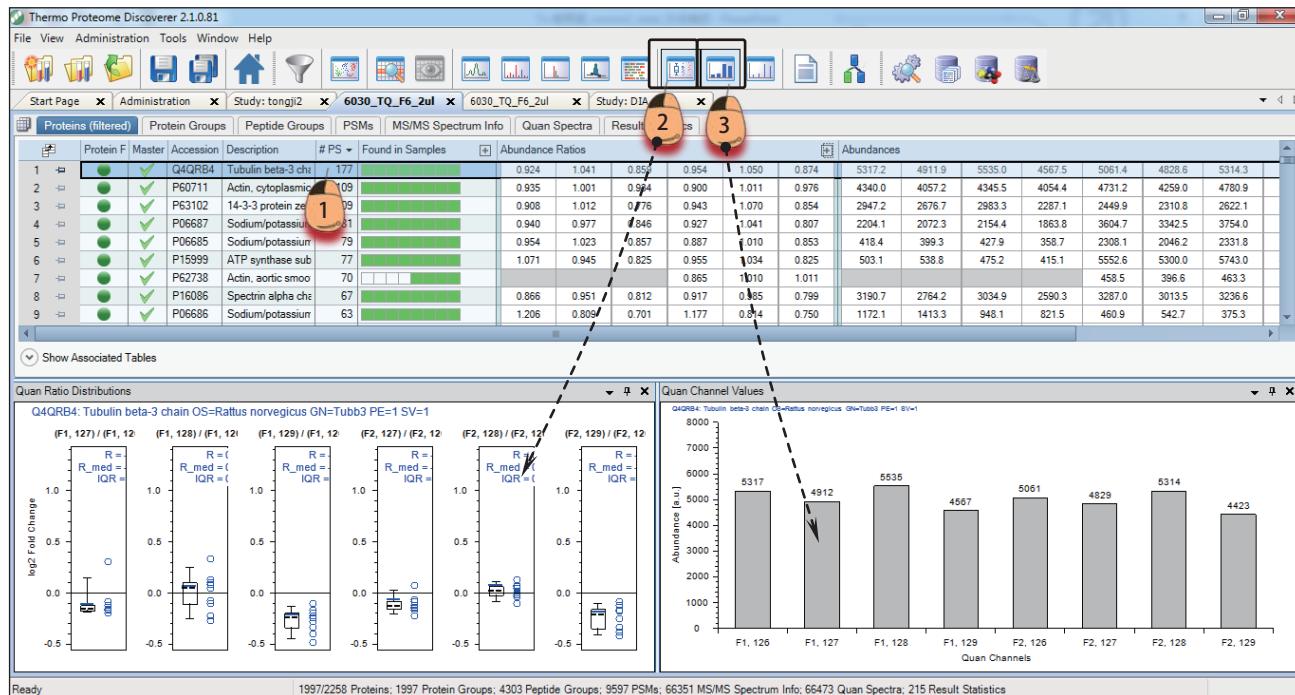
Found in Samples: 显示蛋白在不同文件中不同 channel 的鉴定情况, 点击“+”显示文件名称及不同 channel

Abundance Ratios: 根据定量方法中的设置, 显示不同 channel 间的比值, 点击“+”显示定量比值来源

Abundance: 蛋白在不同 channel 中的信噪比 (S/N) 或强度 (intensity)。

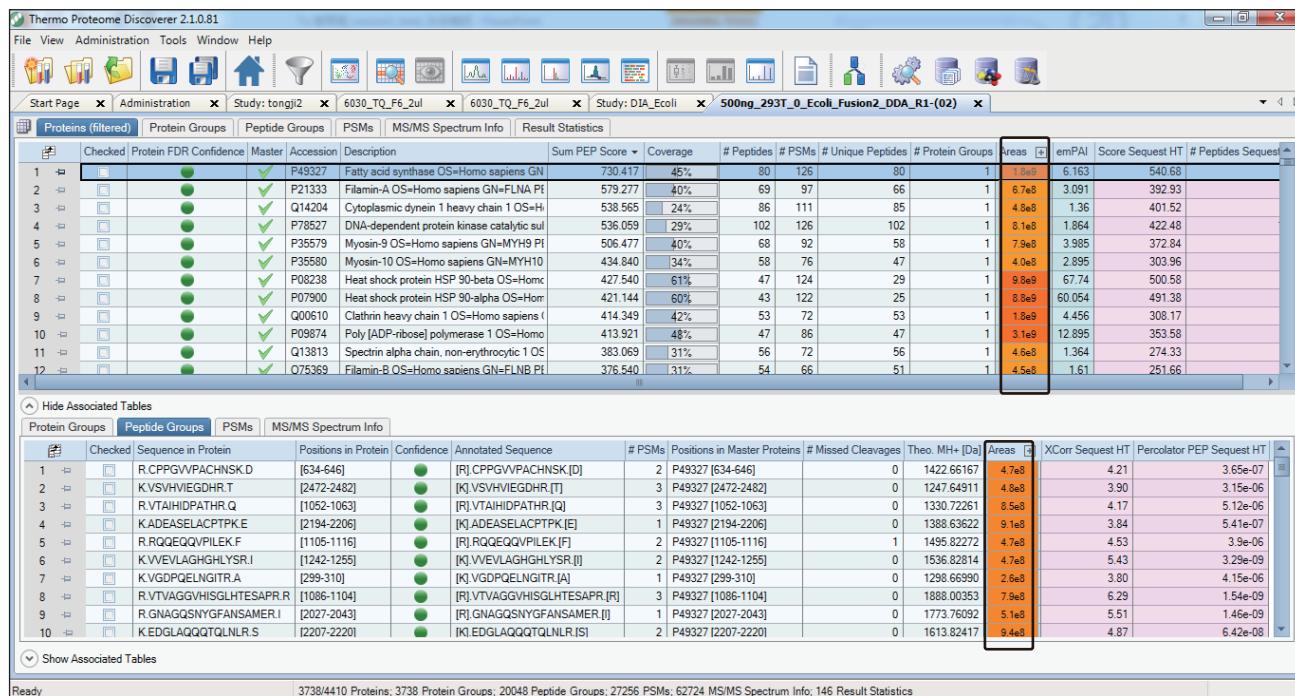


定量比值分布



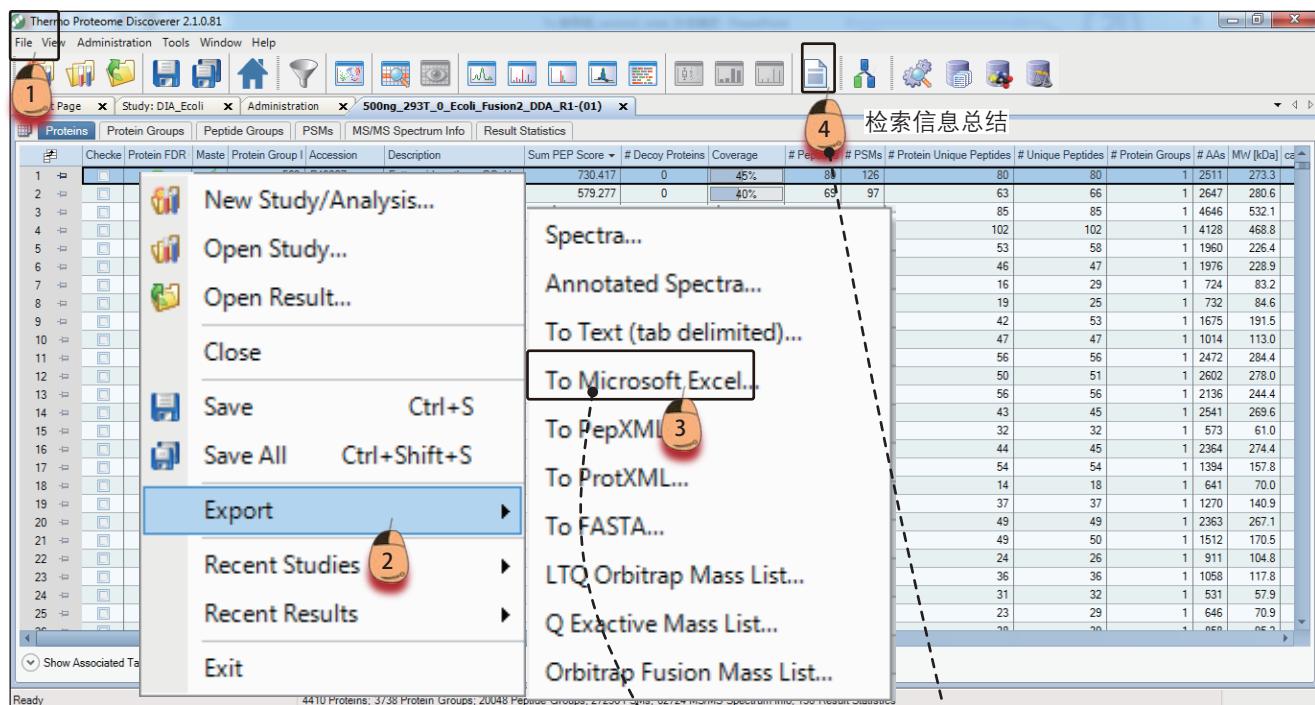
1. 选中目标蛋白。
2. 展示不同 channel 的比值分布图。
3. 展示不同 channel 定量结果。

展示母离子定量结果



根据 processing workflow 和 consensus workflow 中的母离子定量设置, 结果中呈现蛋白及相应肽段的峰面积; 其中 emPAI 值代表蛋白在一个样品中的绝对量, 值越大蛋白含量越高。

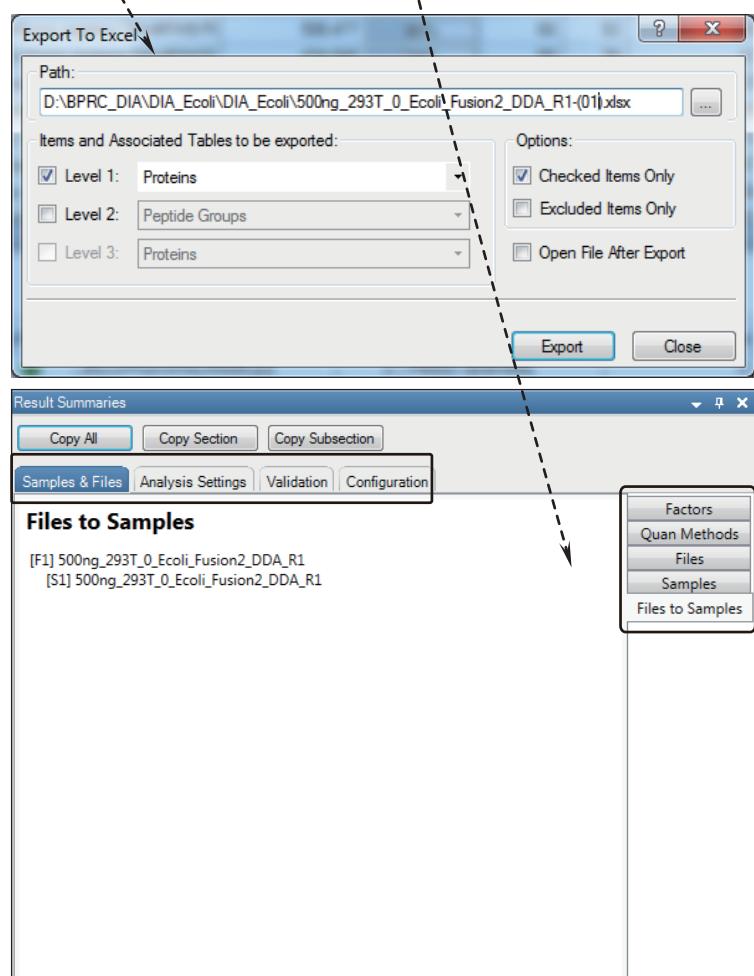
结果导出及检索信息总结



PD 可导出多种形式的文件，根据实际需要进行选择，例如选择输出 Excel，可导出蛋白及对应的肽段信息

注：PD 2.0 与 2.1 输出结果界面有较大区别，请以具体软件版本为准

可查看当前结果的详细搜索参数设置





赛默飞官方微信

赛默飞世尔科技（中国）有限公司

上海（中国总部）
上海浦东新金桥路27号7号楼

广州
广州东风中路410-412号
时代地产中心2405-2406, 3001-3004

北京
北京市安定门东大街28号
雍和大厦西楼7层

沈阳
沈阳市沈河区惠工街10号
卓越大厦3109室

香港
香港新界沙田，沙田乡事会路138号
新城市中央广场第一座九楼911-915室

武汉
武汉东湖高新技术开发区
高新大道858号A7楼

全国服务热线：800 810 5118 400 650 5118 (支持手机用户) 官方网站：www.thermofisher.com

Thermo
SCIENTIFIC

A Thermo Fisher Scientific Brand