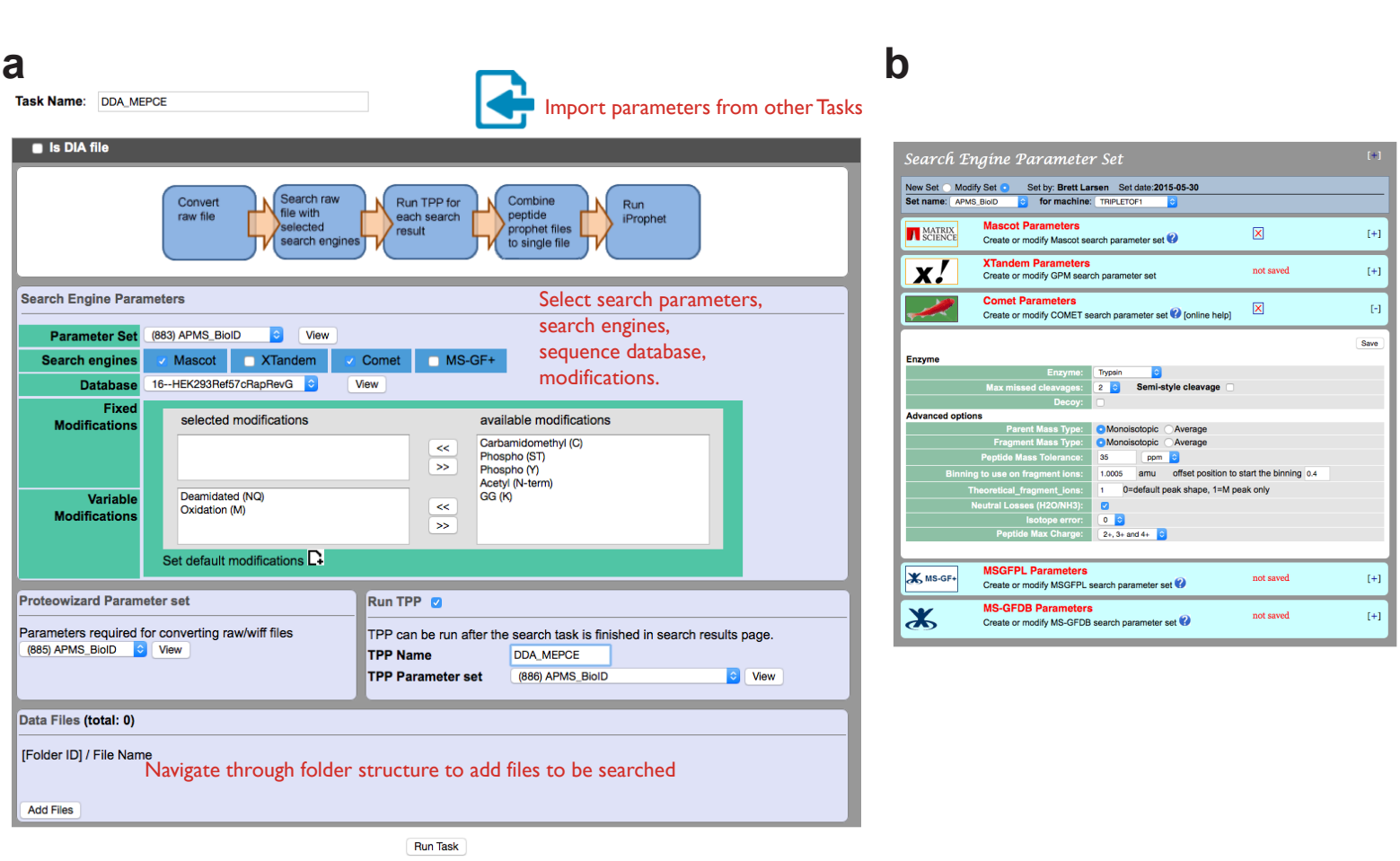
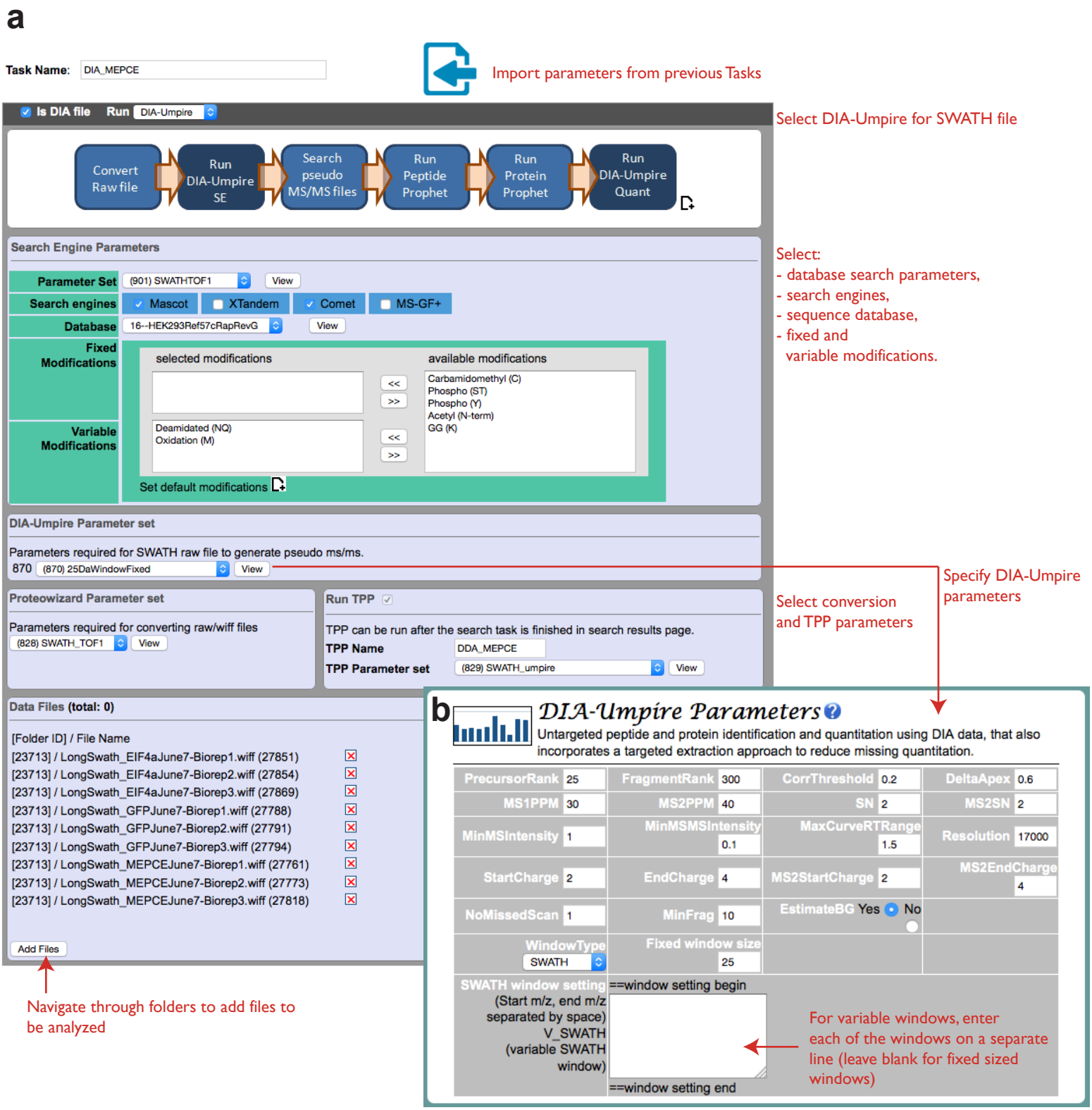


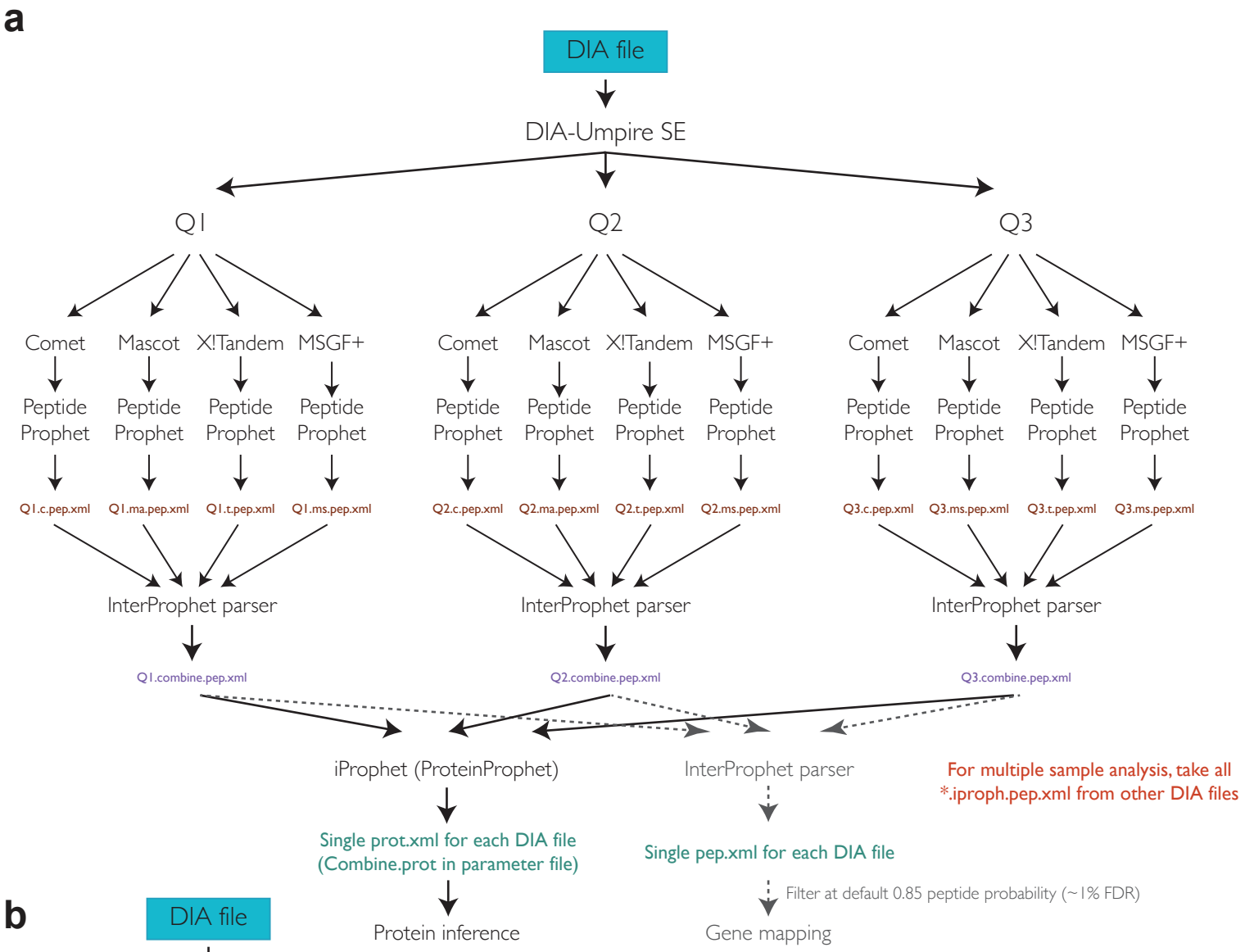

**Supplementary Fig. 1. Initiating searches for Data Dependent Acquisition.**

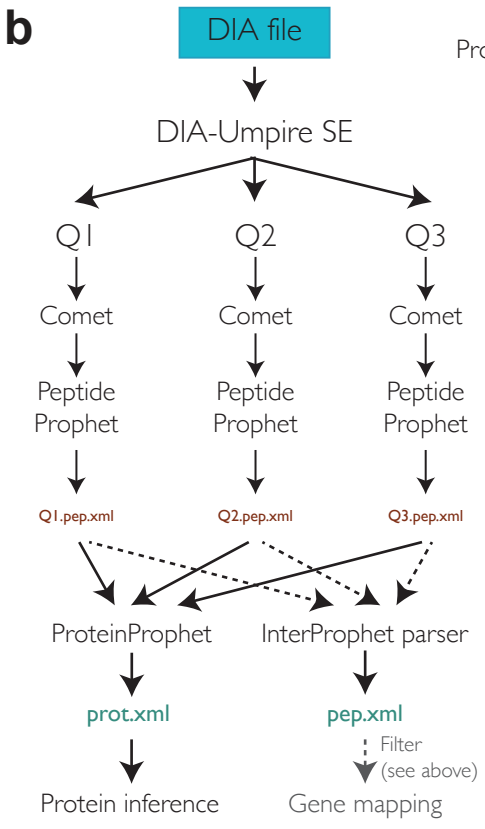
a) Search Task overview. Use drop-down menus and clickable options to select parameters and options or import from a previously-defined task. Generic parameter sets for each instrument and search engine can be selected (see panel b), the search engine(s) to be used specified, and the database and modifications chosen. c) To keep the interface clean and avoid human error, an administrator can define which sequence databases will be visible to users. d) An administrator can set the list of modifications which all users are allowed to use in setting up their tasks. Additional modifications can be added at any time by an administrator and made available for all users or for selected users. ProHits tracks all of these parameters: once a task has been started, these parameters cannot be further modified, though new files can be added to the search task.




**Supplementary Fig. 2. Running DIA-Umpire untargeted identification through the ProHits interface.**

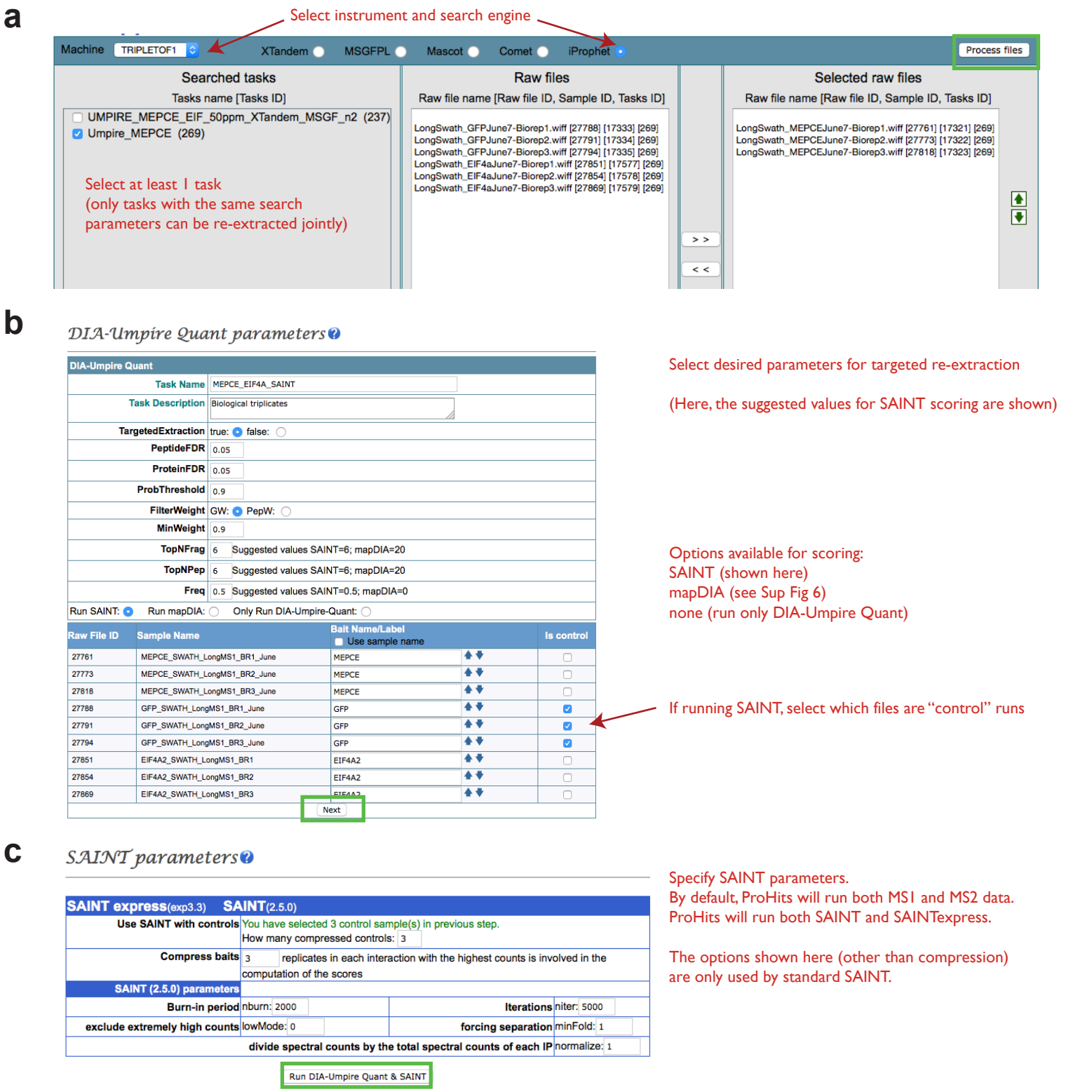
a) Select “DIA file” and “MS-Umpire”. In this Data Management module of ProHits, DIA-Umpire SE (Signal Extraction) generates pseudo-MS/MS spectra from precursor-fragment group data. These spectra can be searched using standard tools designed for DDA analysis, including the iProphet pipeline; note that DIA-Umpire SE extracts three quality tiers of precursor-fragment group data that are each searched and modeled separately (Supplementary Fig. 3). The final iProphet ProteinProphet result merges all quality groups for all search engines used. Currently, Comet, MSGF+, X!Tandem and Mascot are supported. Select the desired database search engines, protein database and search parameters that will be used to search the pseudo MS/MS spectra. b) Select the DIA-Umpire parameters. DIA-Umpire supports a variable window design (see Tsou et al., Nature Methods, 2015). In this case, specify each individual window’s boundaries. Leave blank if fixed windows are used.



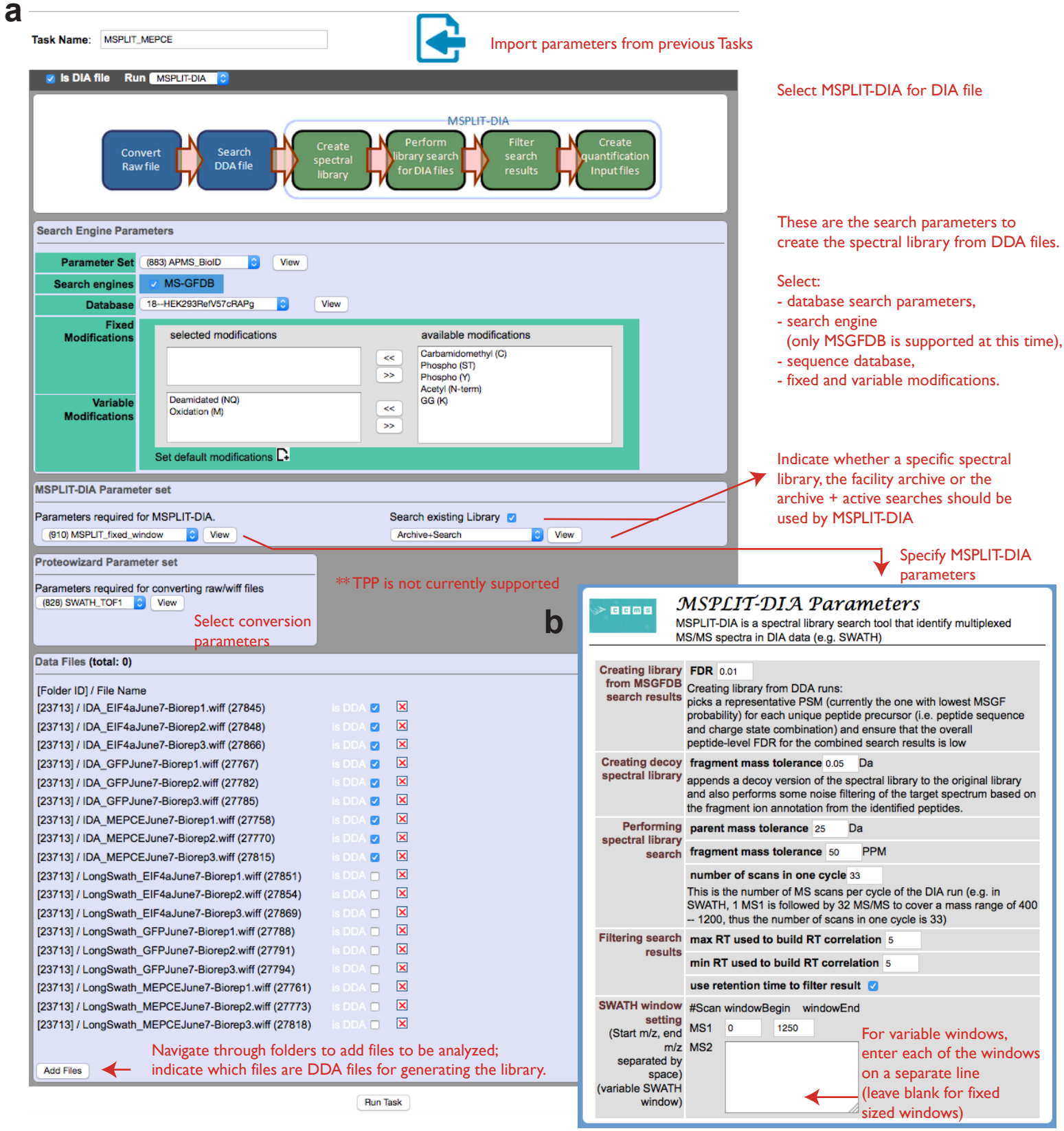

**Supplementary Fig. 3. Detailed process for combining search results in the DIA-Umpire SE module.**

a) iProphet pipeline. DIA-Umpire separates the precursor-fragment groups in three quality tiers. Q1 corresponds to MS1 clusters of 3 or more isotope peaks, Q2, to those with 2, and Q3 to those that have been detected only as clusters of at least 2 unfragmented precursor ion peaks in MS2. Because of the different ratios of correct versus incorrect identifications across these tiers, PeptideProphet is run separately for each of the tiers (see details in DIA-Umpire original paper by Tsou et al., Nature Methods 2015). To combine results from multiple search engines into one final report, searches using each search engine are first performed on individual quality tiers, and PeptideProphet is performed on those results. InterProphet parser is used to combine the results into a single pep.xml file for each of the quality tiers, and once again to combine across all search engines and all quality tiers (this is the version used for extraction of gene level information). iProphet (ProteinProphet) is then applied to these three combined pep.xml files to yield the final protein list. Note that the gene level viewer instead extracts information directly from the pep.xml to perform mapping at the gene level (grey); see Supplementary Fig. 6 for additional information. b) In addition to providing an iProphet view of the data, pep.xml and prot.xml files are generated for each search engine. In this case, PeptideProphet is run on the individual files, which are combined into a single pep.xml using the InterProphet parser (gene level information is extracted from the resulting file). ProteinProphet runs directly on the Q1, Q2 and Q3 pep.xml outputs to generate the prot.xml file.



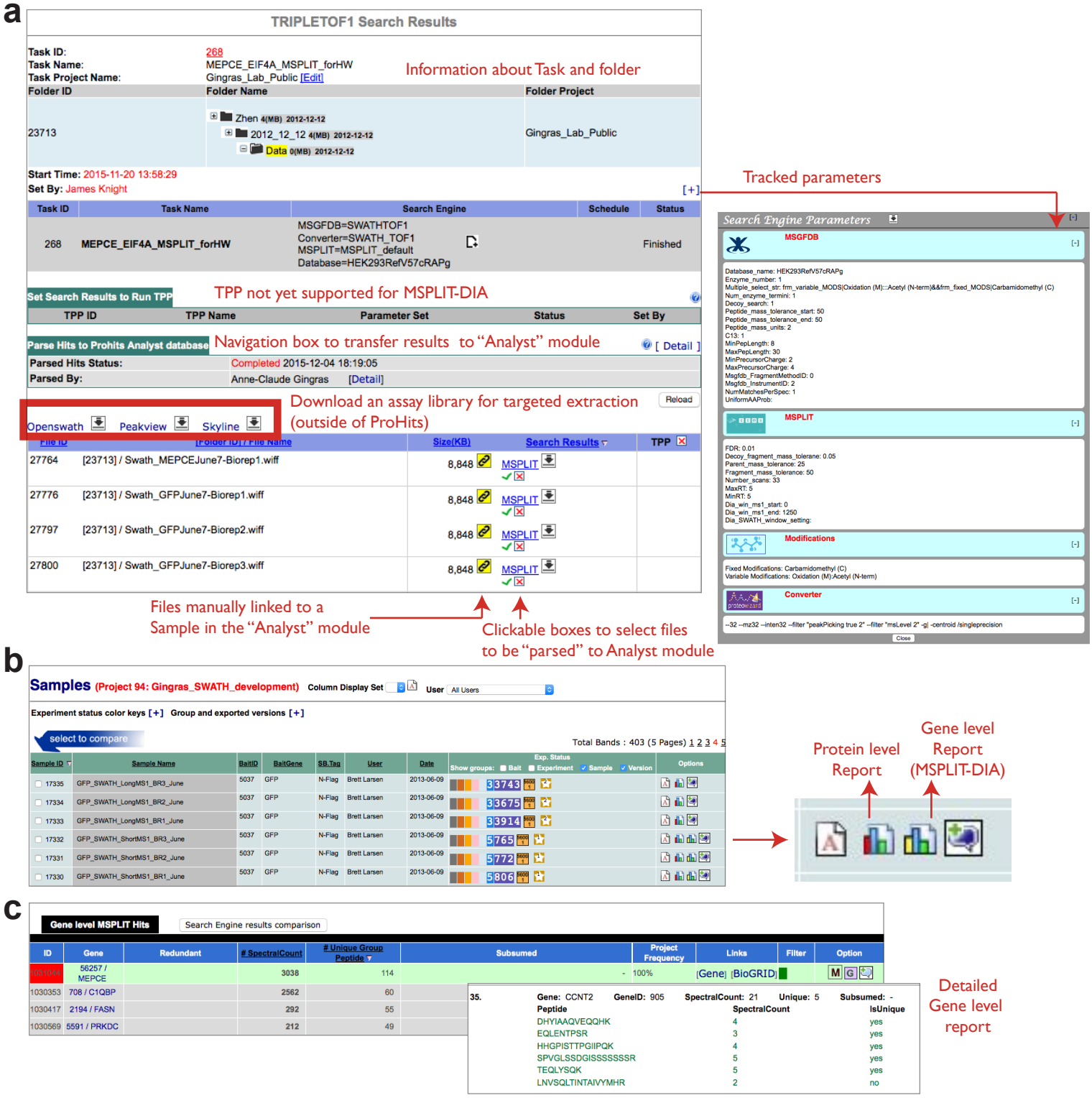

**Supplementary Figure 4. Running the semi-targeted re-extraction module of DIA-Umpire (DIA-Umpire Quant).**

a) To be able to perform targeted extraction using DIA-Umpire's Quant, a DIA-Umpire SE task must first have been run in the Data Management module. All files to be re-extracted jointly must have been searched with the same database and search parameters. Select the files for re-extraction. b) Specify the parameters for DIA-Umpire extraction, and whether you want to only run DIA-Umpire or to also analyze the data with SAINT (panel c) or mapDIA (Supplementary Fig. 7). If running SAINT, specify here which samples should be used as negative controls. Note that in SAINT, all samples listed with the same name (here, the gene name by default) will be combined as one SAINT bait (with n replicates): If a separation of the samples is desired, either click the "Use Sample Name" box, or manually force separation by giving each sample a different name. c) By default, ProHits will initiate the following SAINT tasks: SAINTexpress, intensity model, with both MS1 and MS2 quantification, and standard SAINT (currently SAINT 2.5.0), also with both MS1 and MS2 quantification. In each case, the controls can be compressed (to the highest values, a more stringent parameter for scoring), and the baits can also be compressed (meaning that for each prey, the x highest values will be considered). The parameters for standard SAINT can be specified. Each of these four SAINT variations is then associated with a SAINT report that can be viewed online or downloaded for further analysis.



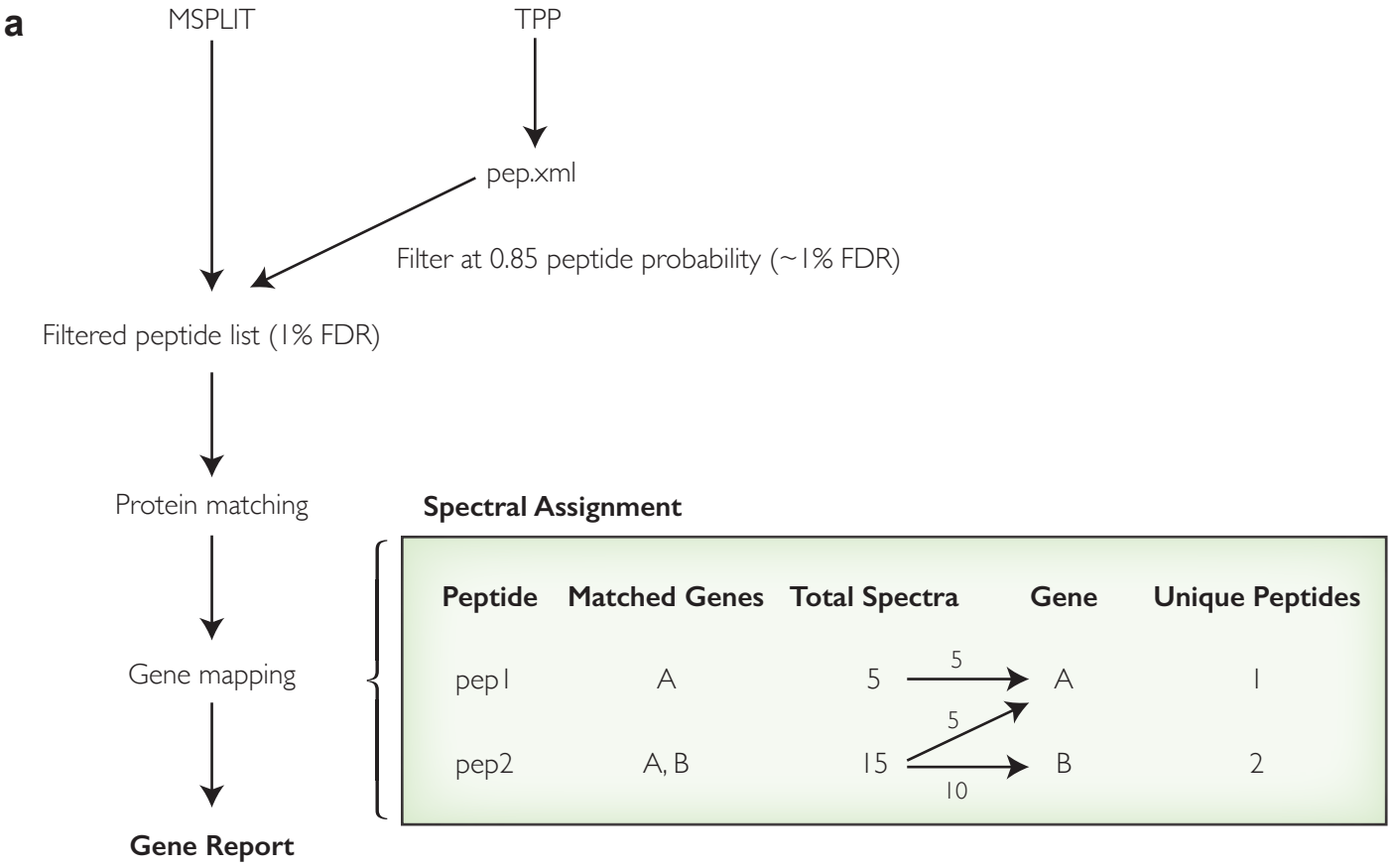

**Supplementary Fig. 5. Running MSPLIT-DIA through the ProHits interface.**

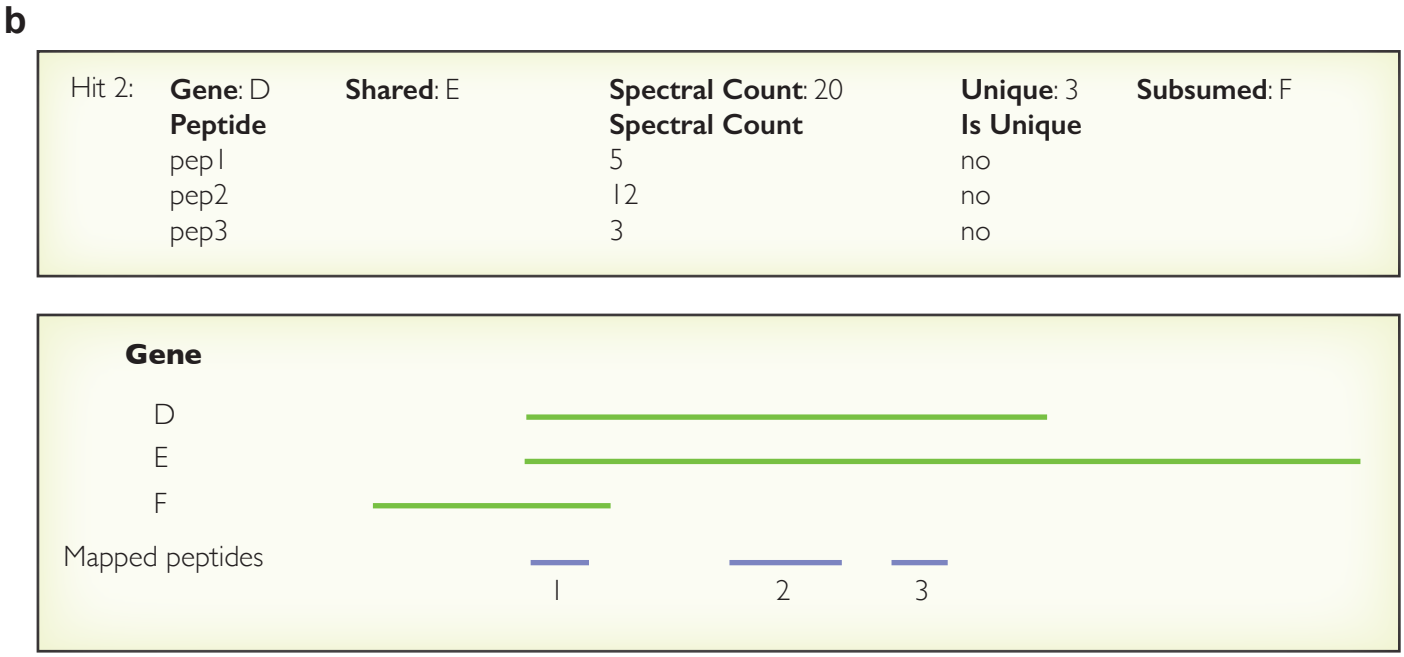
a) Select “DIA file” and “MSPLIT-DIA”. If using DDA files to create a library, select DDA search parameters (only MSGFDB is currently supported as a search engine). Add both DDA and DIA (SWATH) files to the list of files and select which of those are the DDA files that will be used for library generation. Existing libraries can also be searched: individual external libraries are listed (here Human_Swath_Atlas_v1 from the Aebersold group). In addition, ProHits automatically generates a spectral library from all previously-searched DDA files that it appends to externally generated libraries (“Archive”). These can be searched in isolation, or alongside the newly searched results. In all cases, ProHits will take the highest MSGFDB spectrum for spectral matching with MSPLIT-DIA. b) Select the parameters to be used for creating the spectral library from MSGFDB results, creating the decoy spectral library, performing the spectral library search and filtering the search results using retention time correlation. MSPLIT-DIA supports a variable window design (this was not specifically described in Wang et al., Nature Methods, 2015). In this case, specify each individual window’s boundaries. Leave blank if fixed windows are used.




**Supplementary Fig. 6. Results page for MSPLIT-DIA in the Data Management module.**

a) ProHits tracks all search parameters to facilitate reporting. The identification results for each file can be downloaded as a tab delimited file (download icon by each file name). Alternatively, the results can be parsed to the "Analyst" module of ProHits. ProHits also facilitates the generation through MSPLIT-DIA of "assay libraries" for targeted extraction by the popular tools OpenSWATH, PeakView and Skyline. b) Gene level reports (e.g. here for MSPLIT-DIA) are distinguishable from protein level reports through color-coded icons. c) Gene level views. Peptides are first matched to proteins and those identifications are then used for mapping peptides to genes (using Refseq for protein and gene IDs). For peptides shared between multiple genes, spectral counts are assigned exclusively to those genes with unique peptides in proportion to the existing evidence for those genes. If a peptide matches exclusively to genes that have no unique peptides, then spectral counts are divided equally between the genes. After all spectra have been assigned, counts for each gene are rounded to the nearest integer. The final report contains all genes to which peptides have been matched. For each gene, "shared genes" match the exact same set of peptides. "Subsumed genes" are also reported that include genes that match to a subset of the current gene's peptides and nothing outside of this set. In addition to total spectral counts, unique peptide counts are also reported for each gene. In the case of genes that match to the exact same set of peptides, this "unique" number will refer to the number of unique peptides for the corresponding group.




**Supplementary Fig. 7. Gene inference.**

a) Peptides are first matched to proteins and those identifications are then used for mapping peptides to genes (using Refseq for protein and gene IDs). For peptides shared between multiple genes, spectral counts are assigned exclusively to those genes with unique peptides in proportion to the existing evidence for those genes. If a peptide matches exclusively to genes that have no unique peptides, then spectral counts are divided equally between the genes. b) The final report contains all genes to which peptides have been matched. For each gene, "shared genes" are reported if they exist. These encompass genes that match to the exact same set of peptides, i.e. we have no evidence to determine which of the corresponding proteins is in the sample although the peptides tell us a least one is. "Subsumed genes" are also reported that include genes that match to a subset of the current gene's peptides and nothing outside of this set. In addition to total spectral counts, unique peptide counts are also reported for each gene. In the case of genes that match to the exact same set of peptides, this "unique" number will refer to the number of unique peptides for the corresponding group.

**a** *DIA-Umpire Quant parameters*

**DIA-Umpire Quant**

| | |
|---|---|
| Task Name | MEPCE_EIF4A_mapDIA |
| Task Description | |
| TargetedExtraction | true: ● false: ○ |
| PeptideFDR | 0.05 |
| ProteinFDR | 0.05 |
| ProbThreshold | 0.9 |
| FilterWeight | GW: ● PepW: ○ |
| MinWeight | 0.9 |
| TopNFrag | 20 Suggested values SAINT=6; mapDIA=20 |
| TopNPep | 20 Suggested values SAINT=6; mapDIA=20 |
| Freq | 0 Suggested values SAINT=0.5; mapDIA=0 |

Run SAINT: ○ Run mapDIA: ● Only Run DIA-Umpire-Quant: ○

| Raw File ID | Sample Name | Bait Name/Label (Use sample name) | Is control |
|---|---|---|---|
| 27788 | GFP_SWATH_LongMS1_BR1_June | GFP | ☐ |
| 27791 | GFP_SWATH_LongMS1_BR2_June | GFP | ☐ |
| 27794 | GFP_SWATH_LongMS1_BR3_June | GFP | ☐ |
| 27761 | MEPCE_SWATH_LongMS1_BR1_June | MEPCE | ☐ |
| 27773 | MEPCE_SWATH_LongMS1_BR2_June | MEPCE | ☐ |
| 27818 | MEPCE_SWATH_LongMS1_BR3_June | MEPCE | ☐ |
| 27851 | EIF4A2_SWATH_LongMS1_BR1 | EIF4A2 | ☐ |
| 27854 | EIF4A2_SWATH_LongMS1_BR2 | EIF4A2 | ☐ |
| 27869 | EIF4A2_SWATH_LongMS1_BR3 | EIF4A2 | ☐ |

Next

Select desired parameters for targeted re-extraction

(Here, the suggested values for mapDIA scoring are shown)

Options available for scoring:
- SAINT (see Sup. Figure 3)
- mapDIA (here)
- none (run only DIA-Umpire Quant)

Use arrows to organize order as desired

**b** *mapDIA parameters*

Link to the mapDIA user manual

http://xtandemserver.mshri.on.ca/thegpm-cgi/

**mapDIA** (version:2.0.5)

**Experimental design**: replicatedesign

REP design: For example, if the design is a time course experiment with 3 time points (t1,t2,t3) across 2 biological replicates (A,B), then the conditions are time points and thus the samples should be organized in the following order (t1-A, t1-B) (t2-A, t2-B) (t3-A, t3-B)

**Normalization**: none — none/TIS/rt 30

**Filter**: SDF 2; MIN_CORREL 0.2; MIN_FRAG_PER_PEP 3; MAX_FRAG_PER_PEP 5; MIN_PEP_PER_PROT 1

**Sample information**

| Group # | Raw file ID | LABEL | MIN_OBS |
|---|---|---|---|
| 1 | 27794\|27791\|27788 | GFP | 2 |
| 2 | 27818\|27773\|27761 | MEPCE | 2 |
| 3 | 27869\|27854\|27851 | EIF4A2 | 2 |

**DEPs**: MIN_DE 0.01 MAX_DE 0.99

**Protein_level**: MAX_PEP_PER_PROT 5

**Contrast matrix for group comparison**

| | 1 | 2 | 3 |
|---|---|---|---|
| 1 | ☐ | ☐ | ☐ |
| 2 | ☑ | ☐ | ☐ |
| 3 | ☑ | ☑ | ☐ |

Run DIA-Umpire Quant & mapDIA

Specify experimental design and parameters

Suggested parameters are listed by default

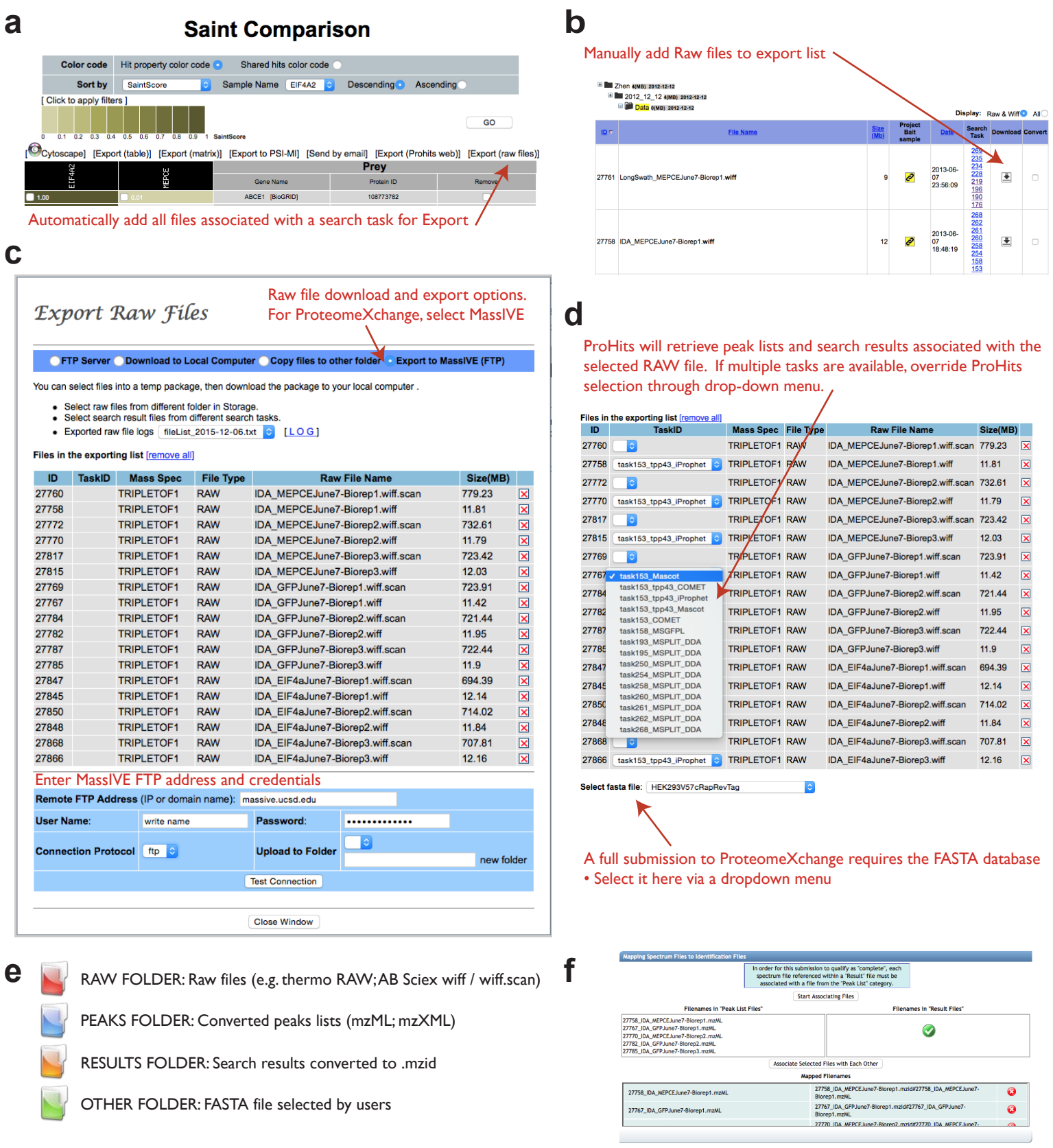
**c**

mapDIA "analysis output"

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Protein | nPeptide | nFragment | Label | Label2 | log2FC | score | FDR | log_oddsDE | log2FC_1 | log2FC_2 | log2FC_3 | nUp | nDown | mcr | min.mcr | max.mcr | PreyGeneID | PreyGene |
| 2 | Q13347 | 8 | 41 | 2 to 1 | EIF4A2/GFP | 2.33989 | 1 | 0 | 164.192 | 2.1362 | 2.42028 | 2.46321 | 3 | 0 | 0.756243 | 0.702197 | 0.785021 | EIF3I | 8668 |
| 3 | O75821 | 5 | 28 | 2 to 1 | EIF4A2/GFP | 2.31502 | 1 | 0 | 169.756 | 2.5796 | 1.94332 | 2.42214 | 3 | 0 | 1 | 1 | 1 | EIF3G | 8666 |
| 4 | O00303 | 4 | 15 | 2 to 0 | EIF4A2/MEPCE | 2.23984 | 1 | 0 | 54.1806 | 2.03122 | 2.6019 | 2.08639 | 3 | 0 | 0.864434 | 0.594655 | 0.999517 | EIF3F | 8665 |
| 5 | P60842 | 7 | 32 | 2 to 1 | EIF4A2/GFP | 2.0113 | 1 | 0 | 67.3016 | 1.53731 | 2.43739 | 2.0592 | 3 | 0 | 0.599081 | 0.514769 | 0.714154 | EIF4A1 | 1973 |
| 6 | P22626 | 3 | 12 | 2 to 1 | EIF4A2/GFP | 1.99311 | 1 | 0 | 69.5806 | 1.92326 | 1.90409 | 2.15198 | 3 | 0 | 0.81223 | 0.775 | 0.884661 | HNRNPA2B1 | 3181 |
| 7 | P62753 | 4 | 20 | 2 to 1 | EIF4A2/GFP | 1.79617 | 1 | 0 | 134.034 | 1.97417 | 1.69194 | 1.72239 | 3 | 0 | 0.969364 | 0.942671 | 0.983032 | RPS6 | 6194 |
| 8 | O00571 | 7 | 39 | 2 to 1 | EIF4A2/GFP | 1.7518 | 1 | 0 | 121.191 | 1.85427 | 1.9021 | 1.49904 | 3 | 0 | 0.83772 | 0.75 | 0.964453 | DDX3X | 1654 |
| 9 | P60842 | 5 | 22 | 2 to 0 | EIF4A2/MEPCE | 1.67667 | 1 | 0 | 52.961 | 1.48226 | 1.39583 | 2.15192 | 3 | 0 | 0.599081 | 0.514769 | 0.714154 | EIF4A1 | 1973 |

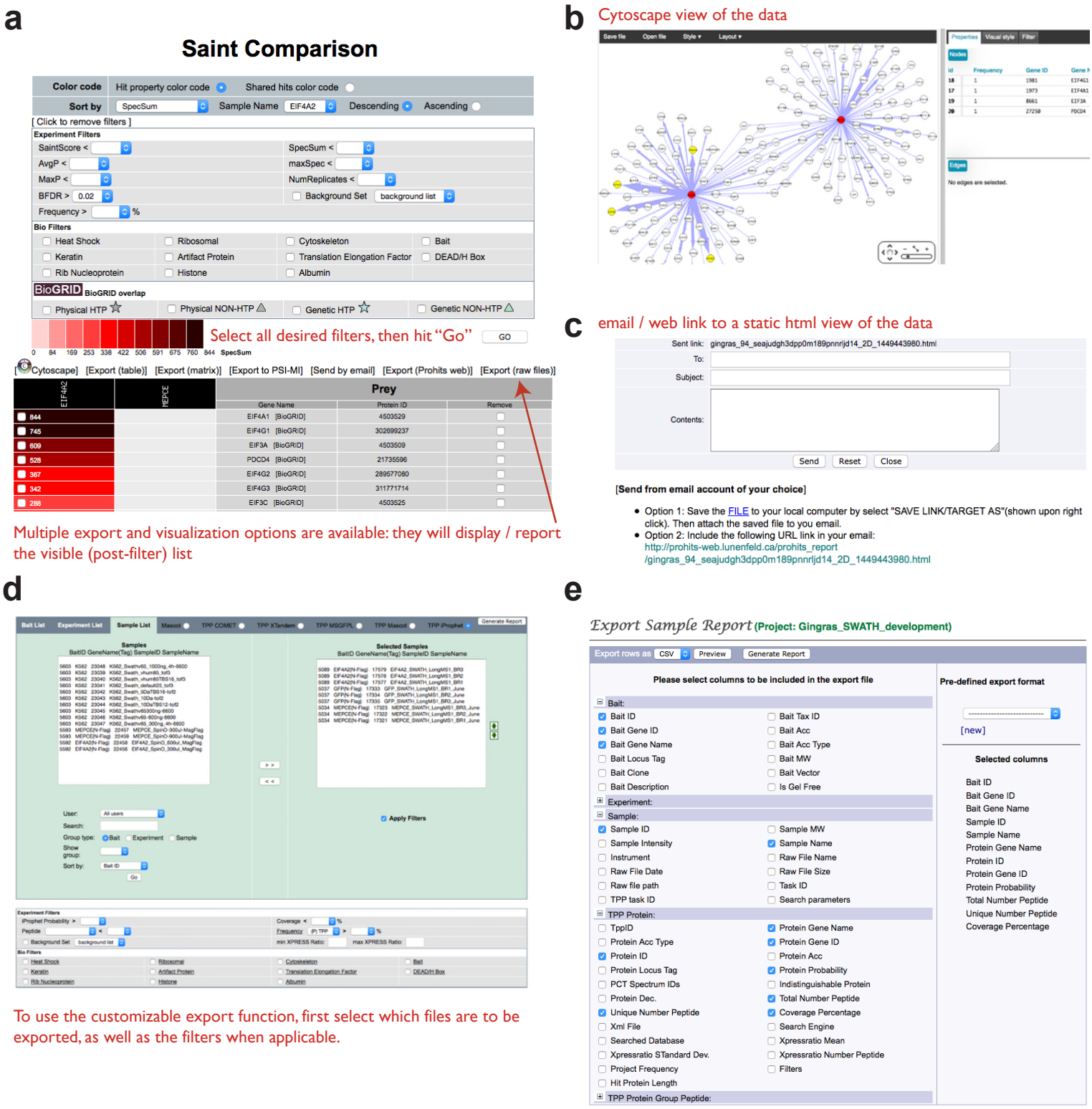
**Supplementary Fig. 8. Running mapDIA from DIA-Umpire through ProHits.**

a) When initiating the DIA-Umpire semi-targeted extraction, select to run mapDIA. Use arrows to organize the data as desired to accommodate experimental design. b) Select mapDIA parameters (see linked mapDIA user manual for detailed explanation of the experimental design and parameters. c) mapDIA returns a folder with several files, including this "analysis output" file which contains quantitative information and statistical parameters. Note that this can be used as an input for visualization tools (e.g. Fig. 2) that are currently run outside of ProHits.




**Supplementary Fig. 9. ProHits facilitates the deposition of data in ProteomeXchange via MassIVE.**

Files to be exported to MassIVE can be selected as a batch from the SAINT Comparison page (a), or manually added to the export list by clicking the download buttons located by individual files or folders in Data Management or individual Samples in the Analyst module (b). All files will be added to a “Export Raw files” list in Data Management (c). To export the data for complete ProteomeXchange submission at MassIVE, select the appropriate button, and login using your credentials. d) ProHits will retrieve both the Peaks list and the search results associated with the selected Raw files: ProHits uses by default the “Search Task” for which the results were parsed to the Analyst; this can be overridden by the user through the dropdown menu selection. For iProphet results, ProHits uses the TPP converters and renames the files for direct recognition in MassIVE. Complete submissions also require deposition of the FASTA sequence database that can be selected via the dropdown menus. e) ProHits generates at MassIVE subdirectories containing the Raw, Peak List, Results and FASTA files. They are named according to the MassIVE nomenclature and can easily be selected. f) MassIVE should recognize automatically all the files and associate them.




**Supplementary Fig. 10. Selected visualization and export options in ProHits.**

a) Multiple export and visualization options are available from any SAINT Comparison page. First select all desired filters and apply them by pressing "Go". Manual removal of selected hits can also be performed. The final view of the data can be downloaded as a table or matrix format, converted to the PSI-MI format, directly exported to the quantitative interaction proteomics repository prohits-web.lunenfeld.ca, viewed as a Cytoscape figure (c), or emailed in a static html format with functional links (d). Additional export options are also available, e.g. that allow for the selection of any file (d) and download of selected values (e).