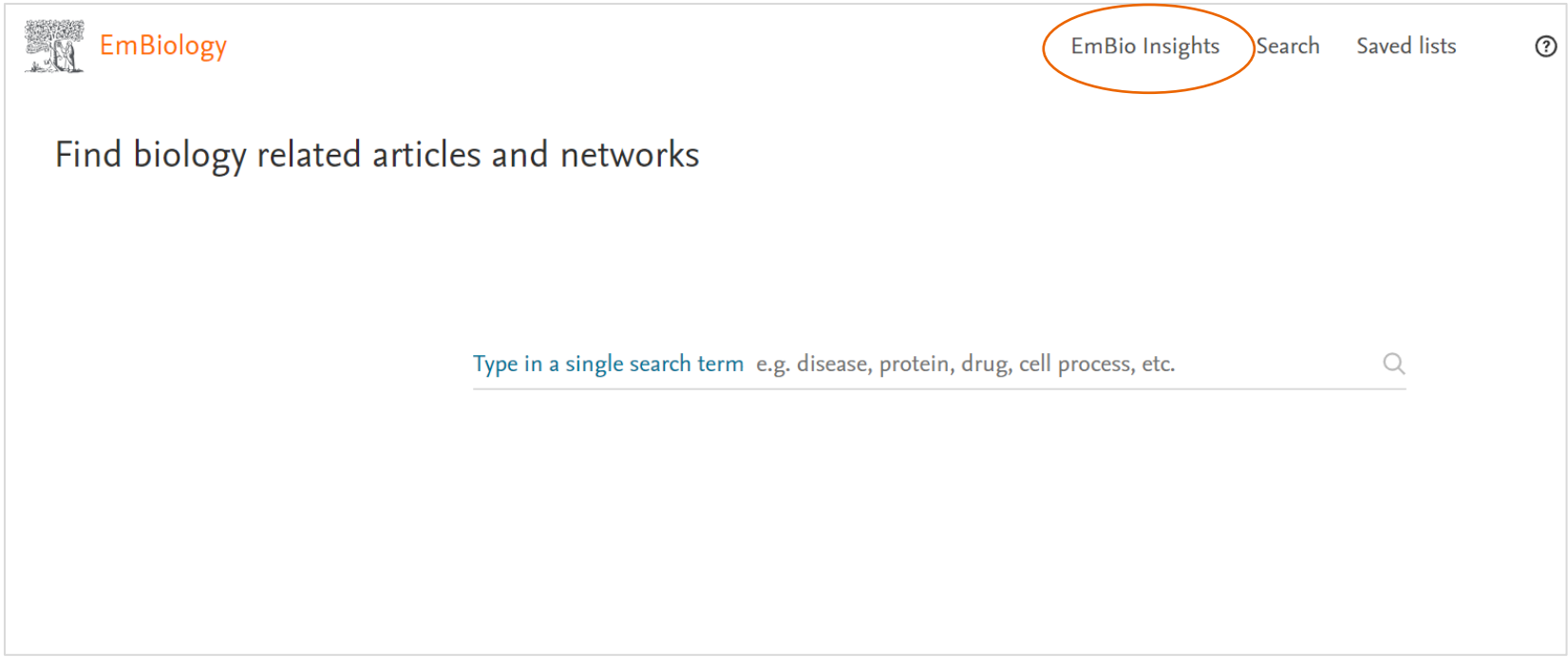
EmBiology Insights guide, including NEW filter options

(Refer to pages 9-12 for new filter workflow)

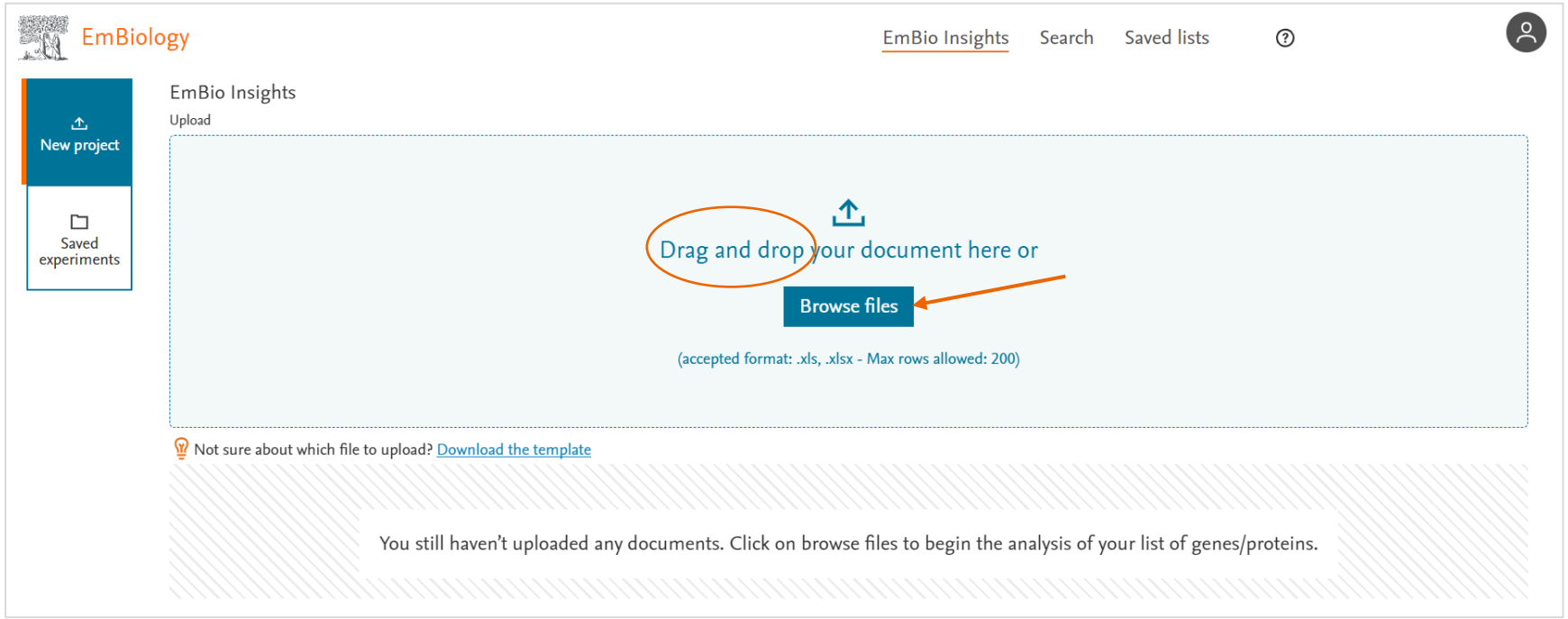
March 2024

Select EmBio Insights to upload and find information about multiple genes/proteins*



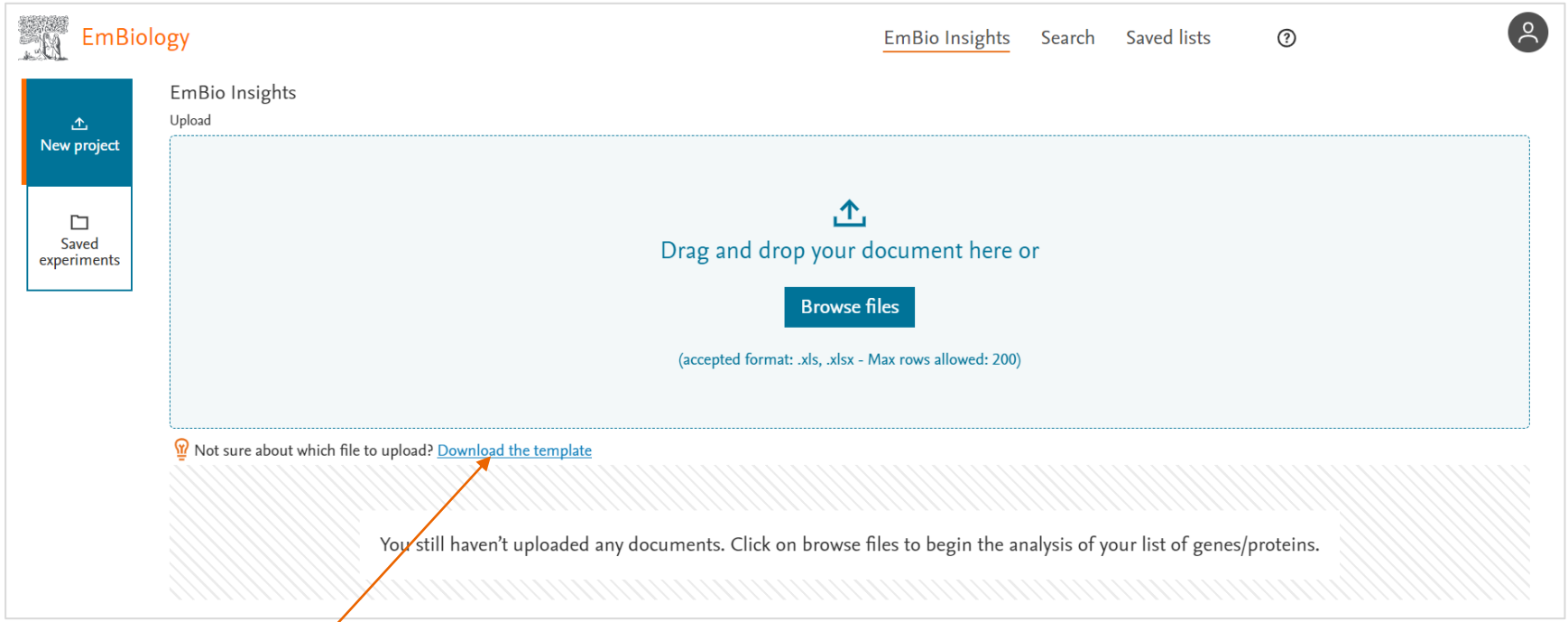
* Genes and proteins fall into the same concept type in EmBiology – for the remainder of this guide, they will be referred to as Proteins

Uploading list of proteins



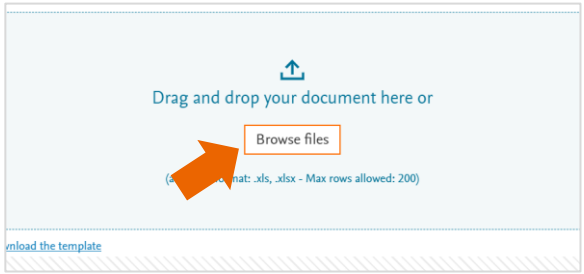
Drag files into the centre of the page or upload documents by clicking 'Browse files'
You can upload a list of up to 200 proteins in an xls or xlsx format

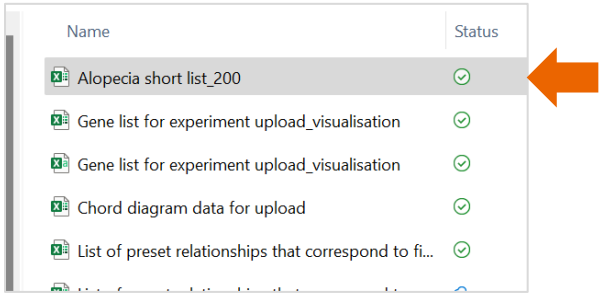
Uploading list of proteins

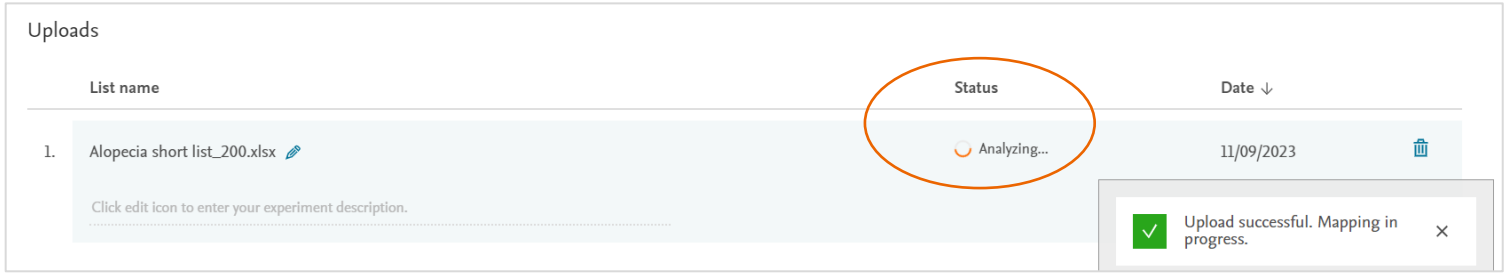


See an example of what the file should look like by clicking 'Download the template'

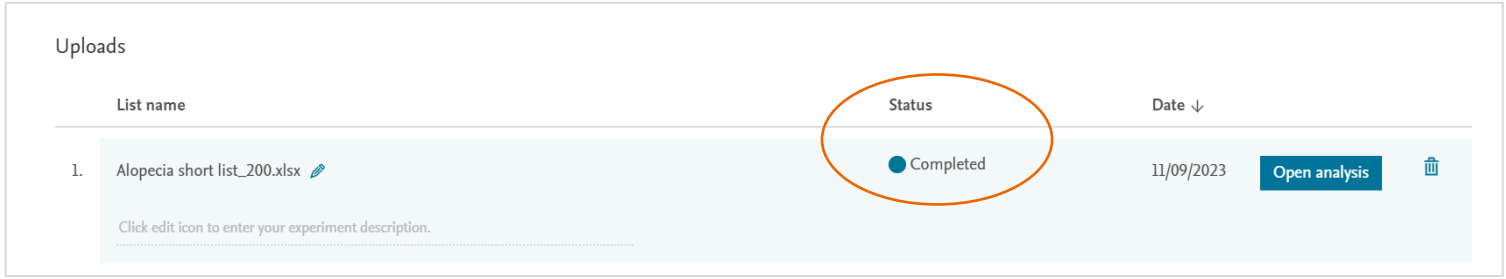
E.g., browse files to upload and map to the database







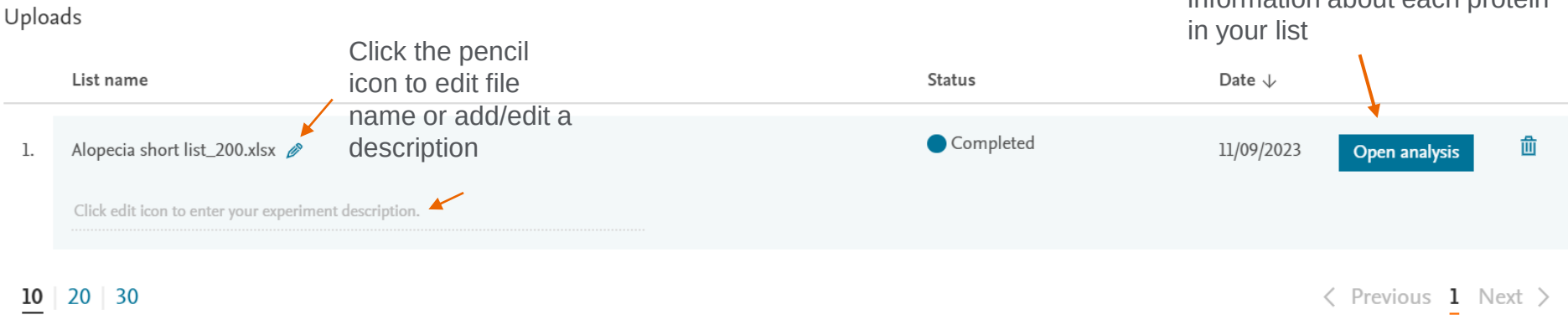
Information is shown that indicates a successful upload. '**Analyzing**' status indicates the proteins in the list are being mapped to the database

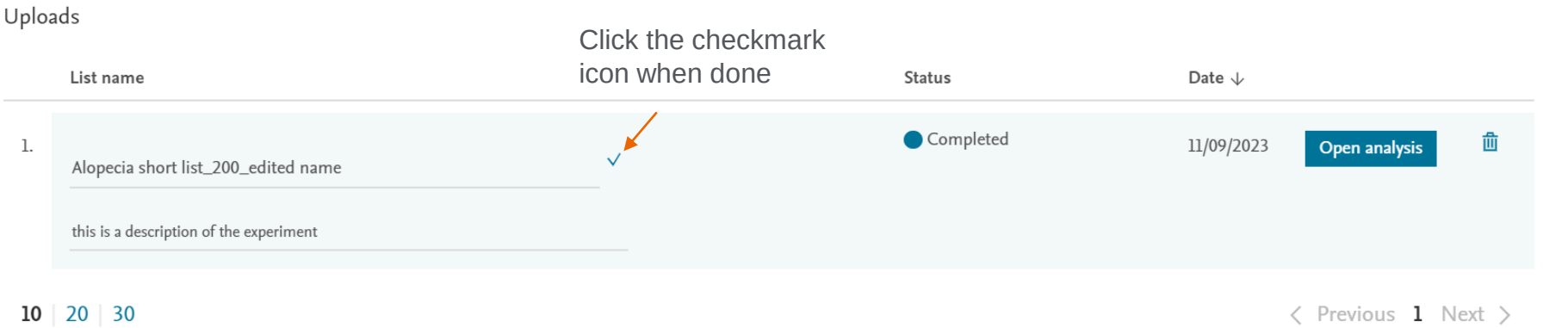


'**Completed**' status indicates mapping is done.
Select 'Open analysis' to proceed to the next step

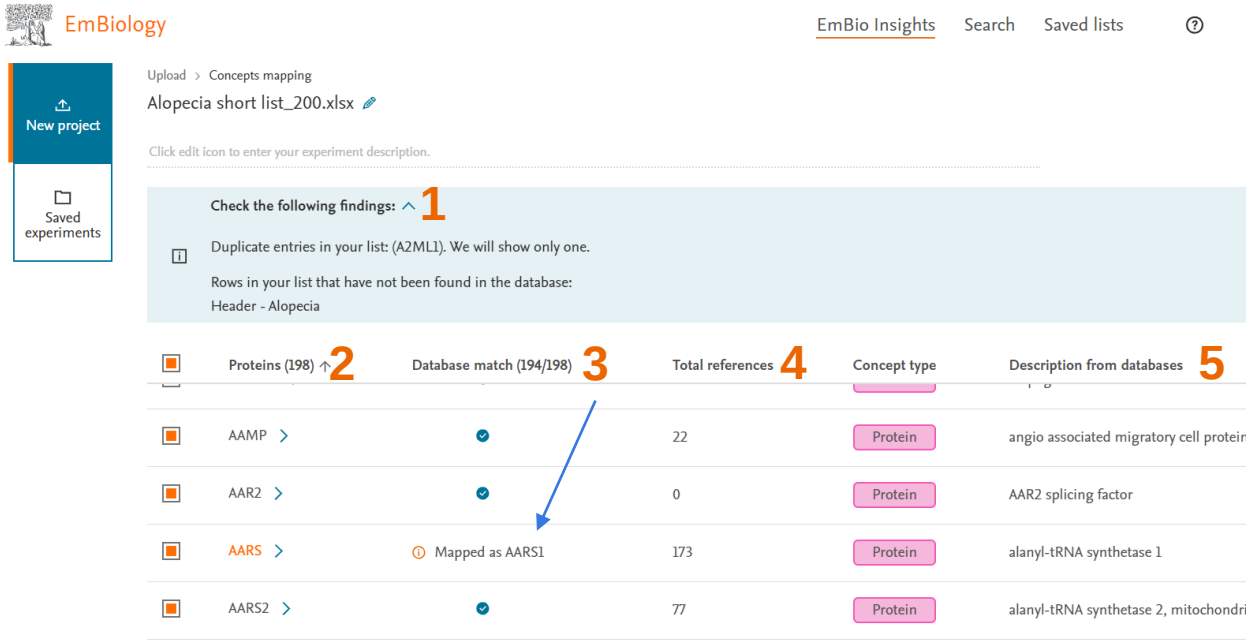
Edit the file name and provide a description

Click '**Open analysis**' to see information about each protein in your list



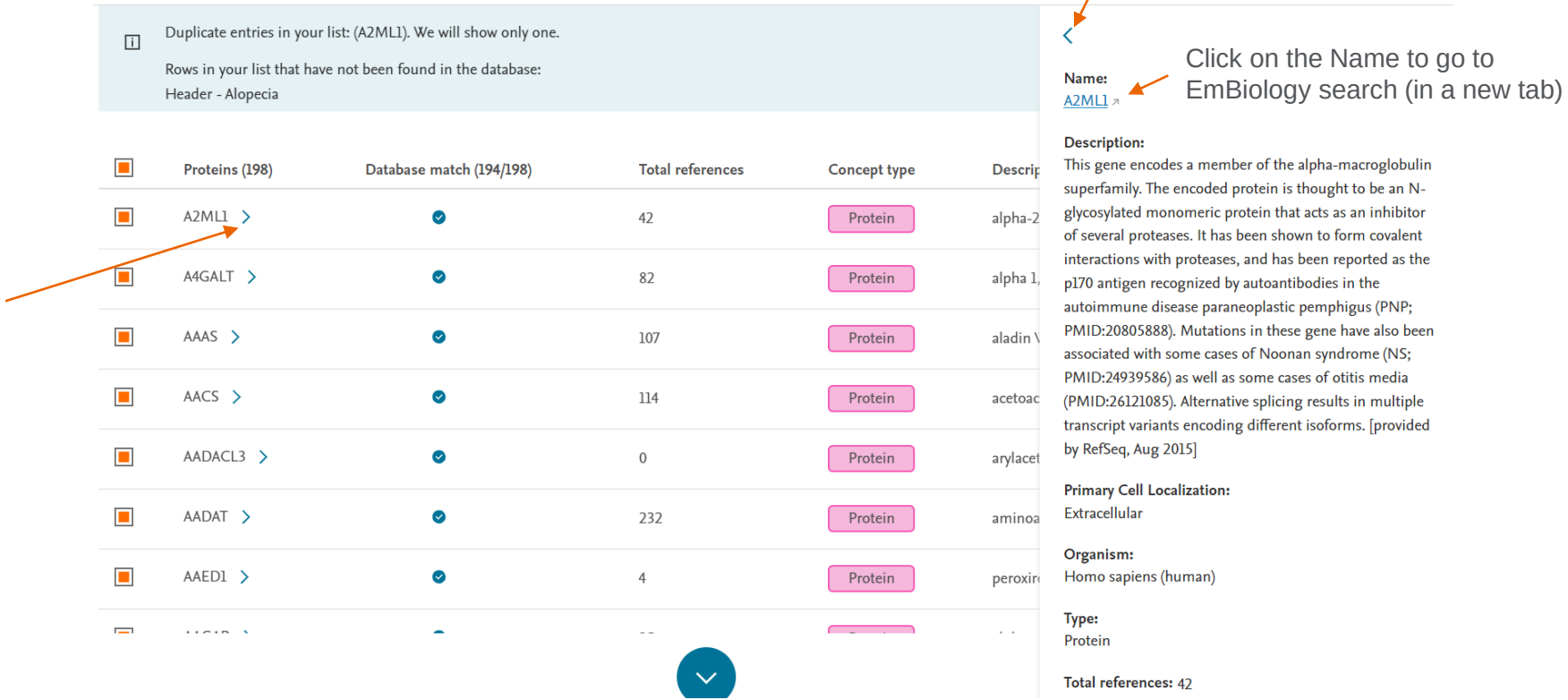


The concepts mapping list shows proteins from your list mapped to the EmBiology database and includes information on each protein

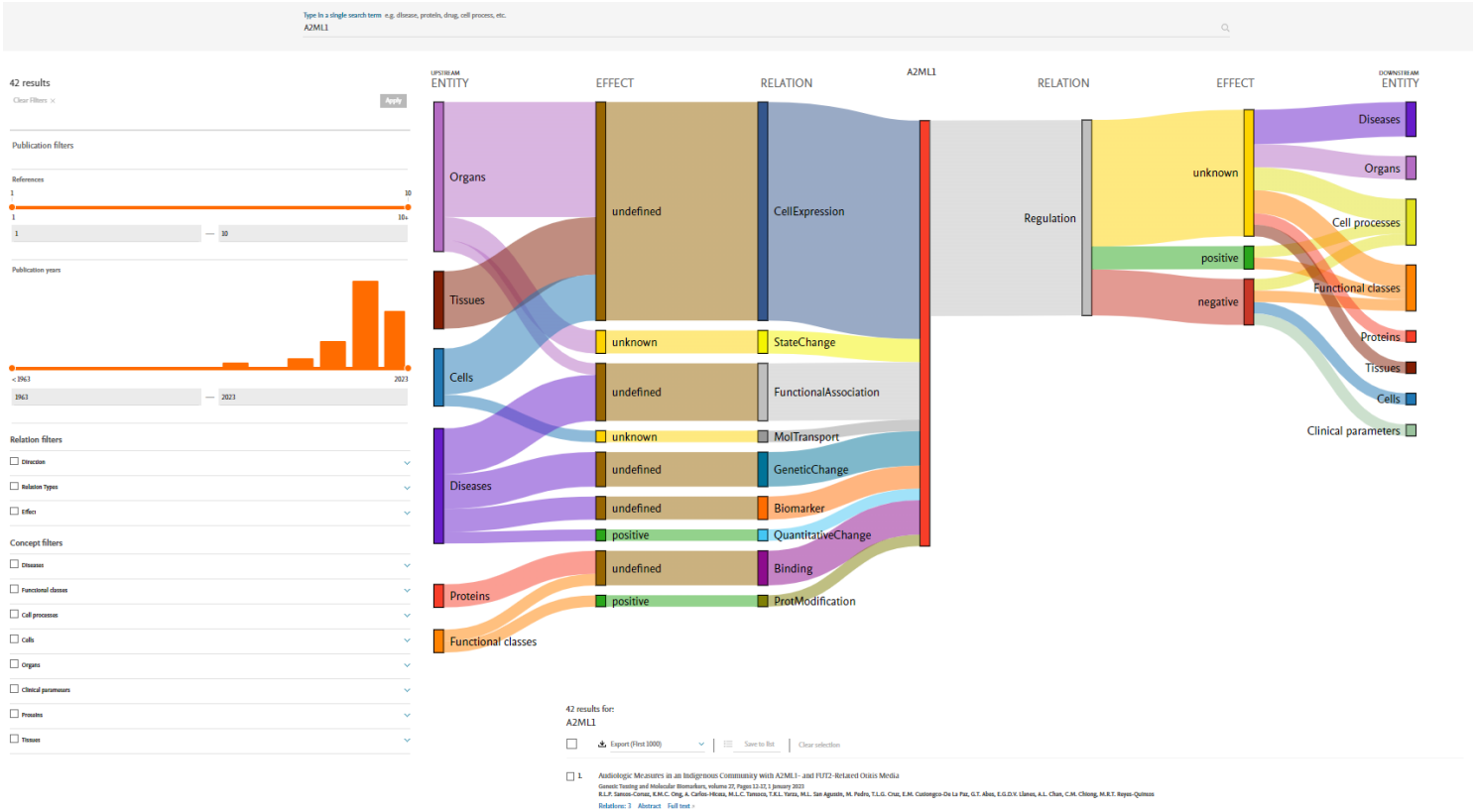


1. Duplicates and terms that can't be matched with proteins found in the EmBiology database are shown in the header
2. Number of proteins successfully mapped to the database
3. In some cases, the protein name in the uploaded list is a synonym of the preferred term – cases where the protein is mapped to a differently named term is clearly indicated
4. Total number of references refers to the number of articles/clinical trials where the protein is mentioned
5. A short description of the protein

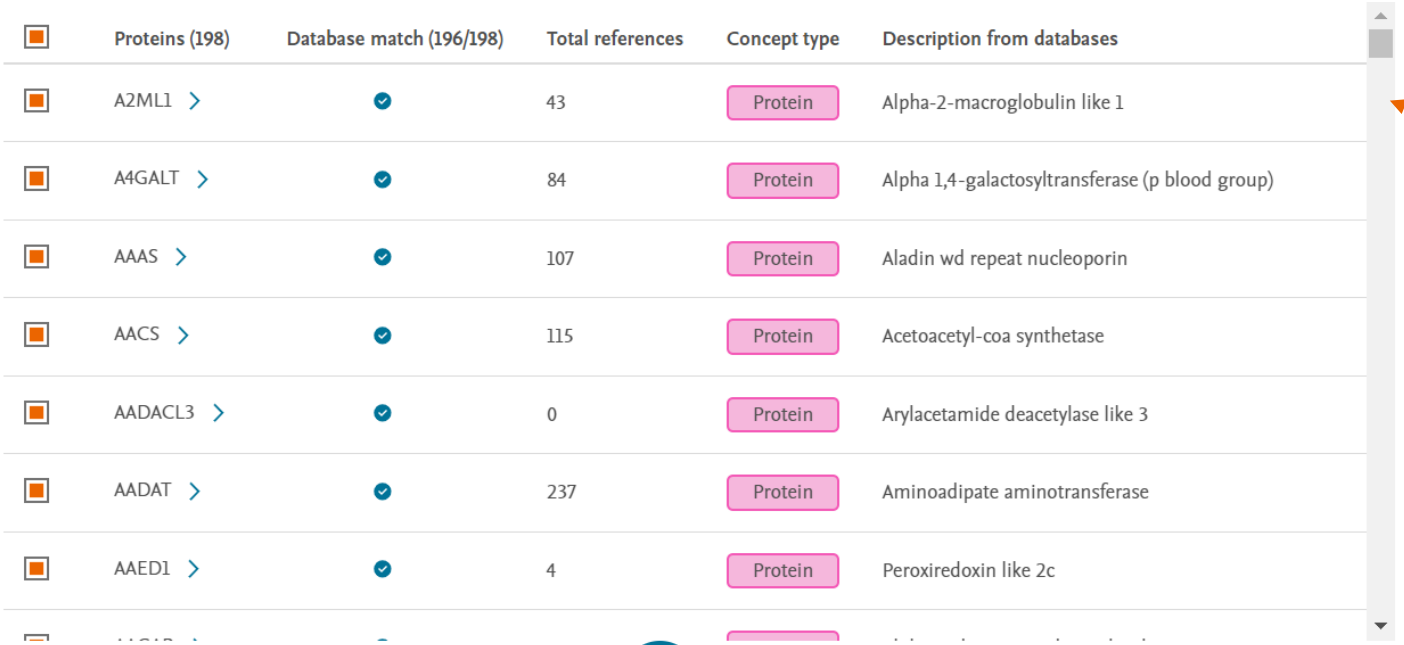
Click on the > beside the protein name to get a detailed description of function, primary cell location and organism the protein is found in



Clicking on the protein name in the information panel opens the EmBiology Search in a new tab



To see relationships for all proteins on your list, click find connections – or (***new***) apply additional filters



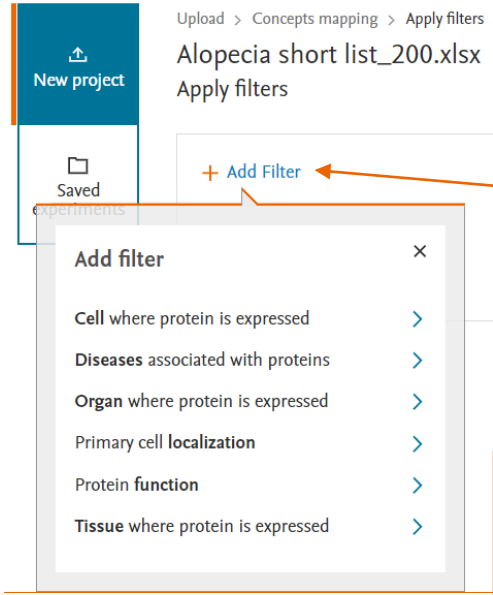
Proteins (198)	Database match (196/198)	Total references	Concept type	Description from databases
A2ML1	✓	43	Protein	Alpha-2-macroglobulin like 1
A4GALT	✓	84	Protein	Alpha 1,4-galactosyltransferase (p blood group)
AAAS	✓	107	Protein	Aladin wd repeat nucleoporin
AACS	✓	115	Protein	Acetoacetyl-coa synthetase
AADACL3	✓	0	Protein	Arylacetamide deacetylase like 3
AADAT	✓	237	Protein	Aminoadipate aminotransferase
AAED1	✓	4	Protein	Peroxiredoxin like 2c

Use the scrollbar to see all proteins in your list

Click ‘Find connections’ to skip filters and immediately see relationships for all proteins on your list – skip to slide 13 for this workflow

Apply filters to narrow down your list of proteins before identifying connections for the proteins in your list

Applying filters to your uploaded list of proteins



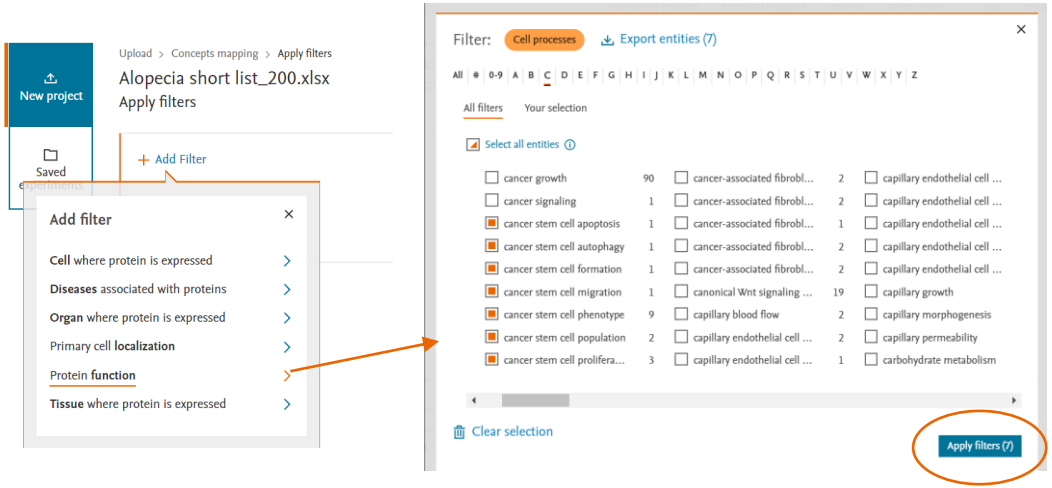
Click + Add Filter to add a filter that will narrow down the number of proteins that will be analysed, based on filter selections

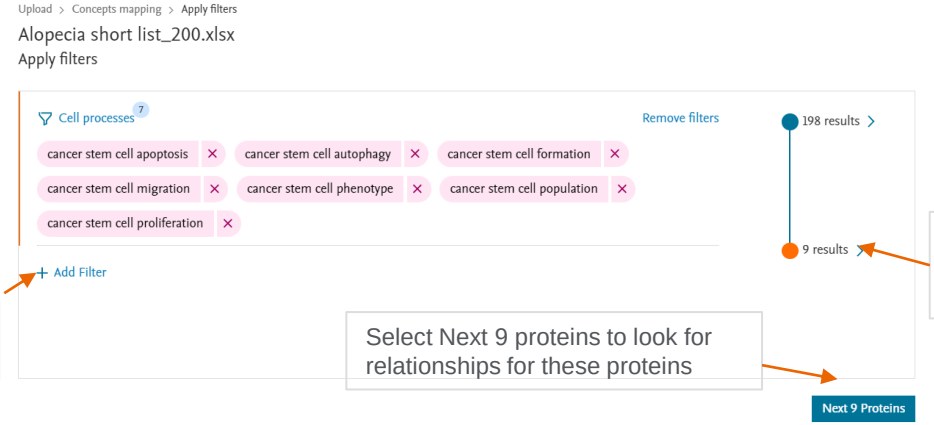
Number of proteins that will be analysed

Skip filters to directly find relationships for all proteins on your list

- **Cell where protein is expressed**: Shows cells that have a cell expression relationship with proteins from your uploaded list. Selecting items from this filter will reduce the proteins to be analysed by ones known to be expressed in the selected cell types.
- **Diseases associated with proteins**: Shows diseases with a regulation relationship with proteins from your uploaded list. Selecting items from this filter will reduce the proteins to be analysed by ones known to be associated with these diseases.
- **Organ where protein is expressed**: Shows organs that have a cell expression relationship with proteins from your uploaded list. Selecting items from this filter will reduce the proteins to be analysed by ones known to be expressed in the selected organs.
- **Primary cell location**: Shows the primary cell location for proteins from your uploaded list. Selecting items from this filter reduces the proteins to be analysed by ones expressed in the selected locations.
- **Protein function**: Shows cell processes that have a regulation relationships with proteins from your uploaded list. Selecting items from this filter will reduce the proteins to be analysed by ones known to be involved in the selected cell processes.
- **Tissue where protein is expressed**: Shows tissues that have a cell expression relationship with proteins from your uploaded list. Selecting items from this filter will reduce the proteins to be analysed by ones known to be expressed in the selected tissue types

For example – apply the Protein Function filter





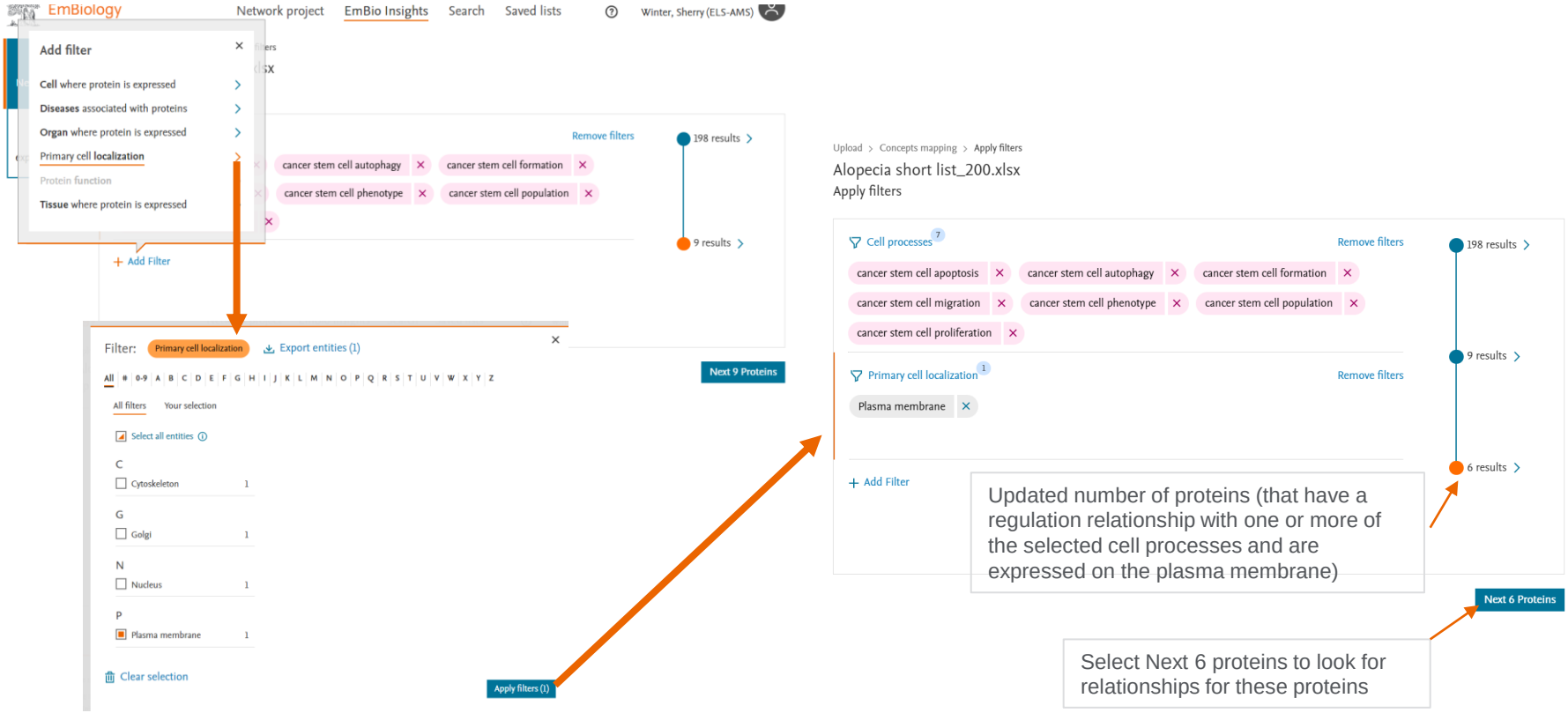
The selected cell processes (protein functions) are shown. Individual ones can be removed if required

Click + Add Filter to apply another filter to the remaining list of 9 proteins

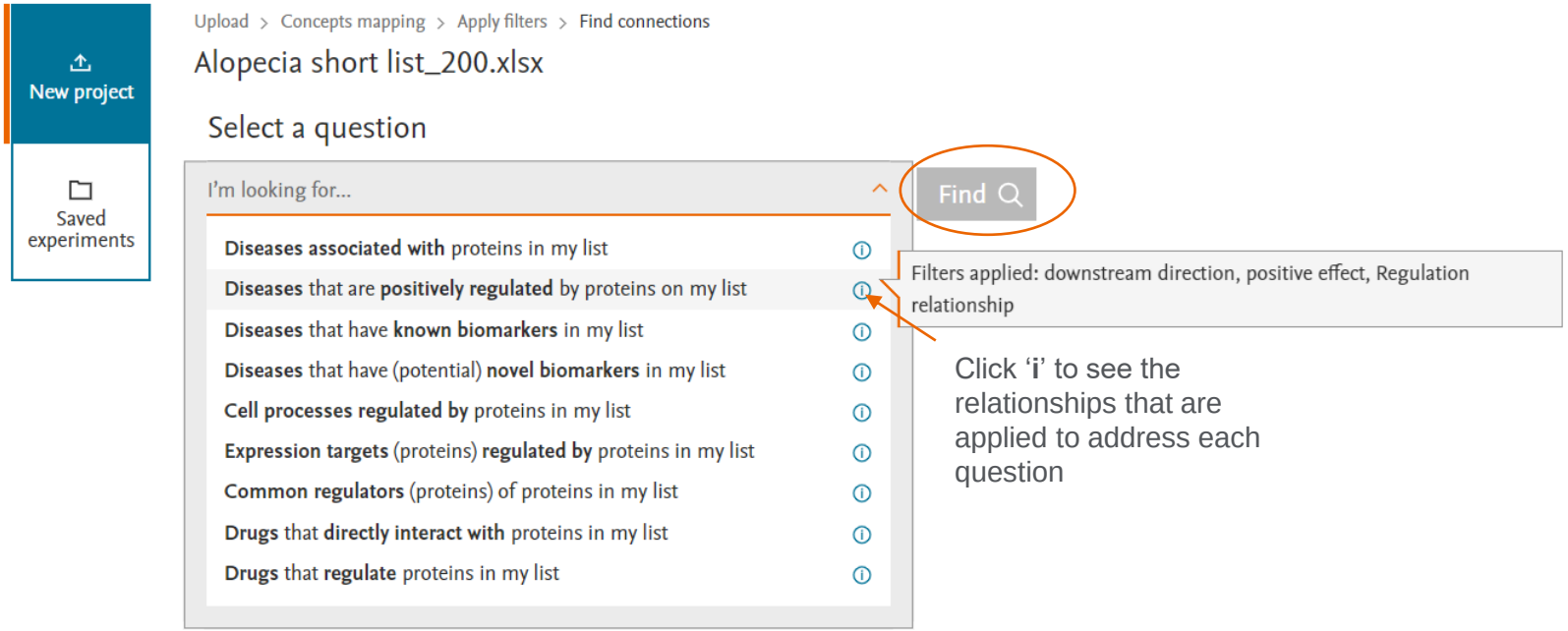
Updated number of proteins (that have a regulation relationship with one or more of the selected cell processes)

Select Next 9 proteins to look for relationships for these proteins

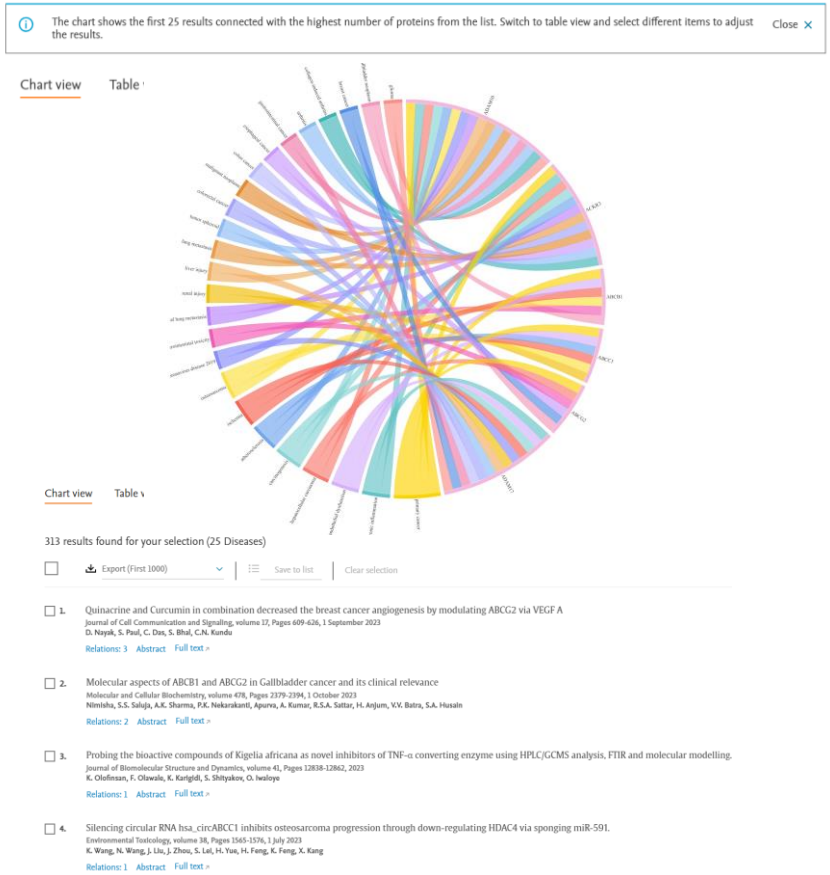
For example – apply a 2nd (protein localization) filter



Select the question to investigate, which will define the relationships searched



EmBiology Insights – discover relationships for multiple terms at one time



- The top 25 associations (based on number of protein connections) is shown by default, along with associated literature
- You can adjust the associations by clicking on table view, which will include the top 500 associations, and making a different selection
- The list of literature at the bottom of the visual includes all literature supporting the relationships shown in the visual
    - Results include information on relations, abstract and link to full text (information is the same as EmBiology search results)

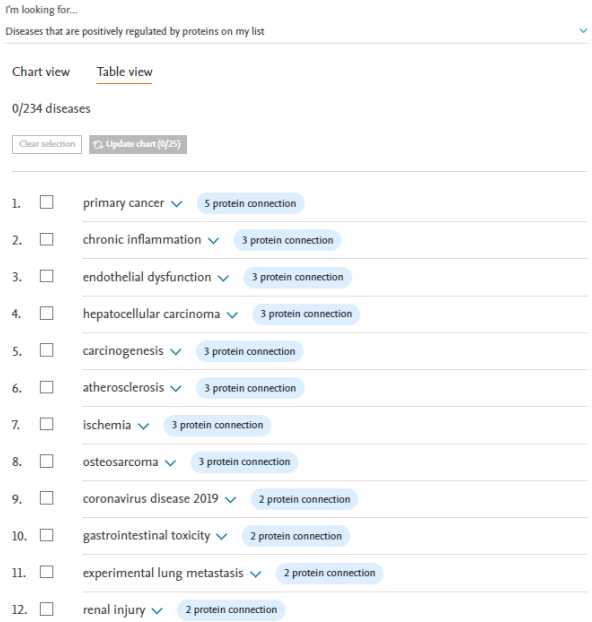
Applying filters to the visual →

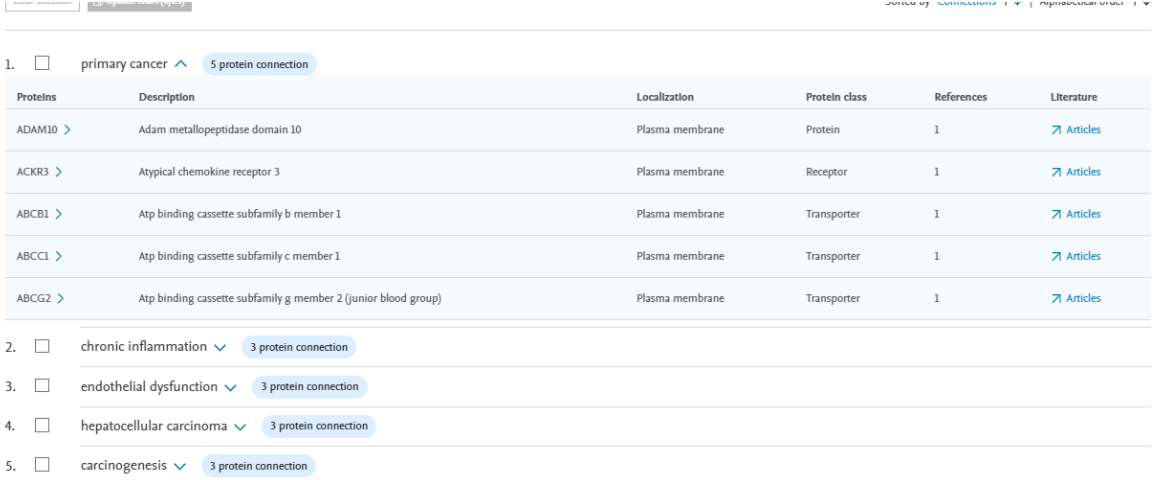
EmBiology Insights – discover relationships for multiple terms at one time

Clicking on 1 (or more) specific diseases applies a filter for the list of literature results

- Clicking on 1 (or more) specific diseases applies a filter for the list of literature results
  - For example, the literature supporting relationships between a single disease and proteins from your list that have a relationship with that disease

Table view: Detailed information for up to 500 results





By default, the top 25 results (based on connections with proteins from the list) are shown in the visual. Users can make other selections for the visual and see the complete list of results (up to 500) in the table

Within each table row, details on individual proteins are seen, including the description, localisation, protein class and number of references supporting the relationship of the protein with the result. A link takes the user to a page with the list of literature supporting the relationship.

Note: information in the table will be made exportable in Q1 2024