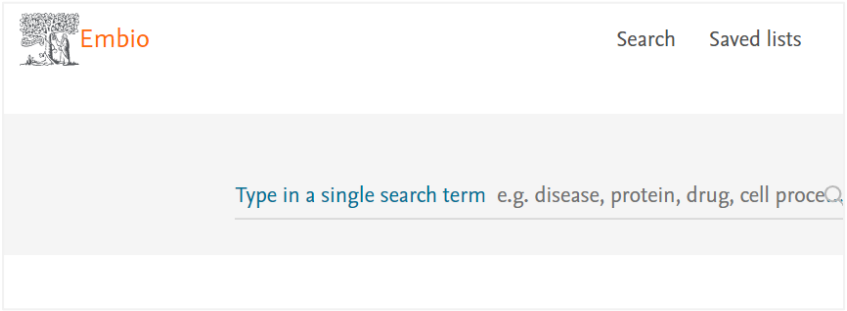
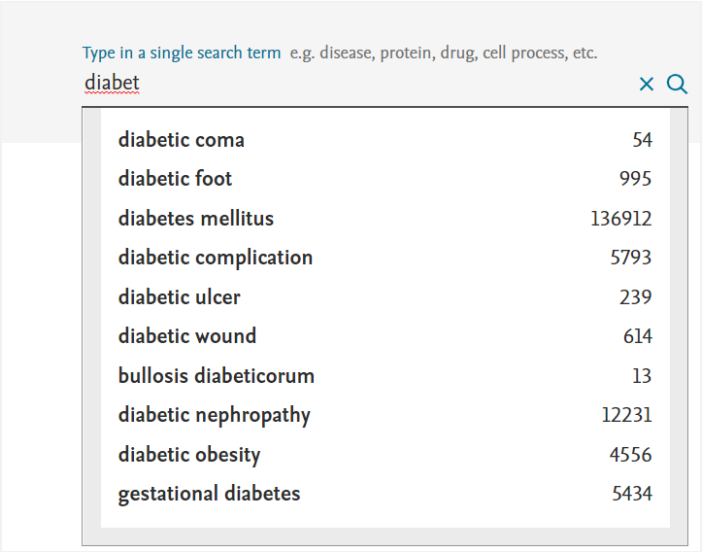
EmBiology Quick User Guide

February 2023

Contents

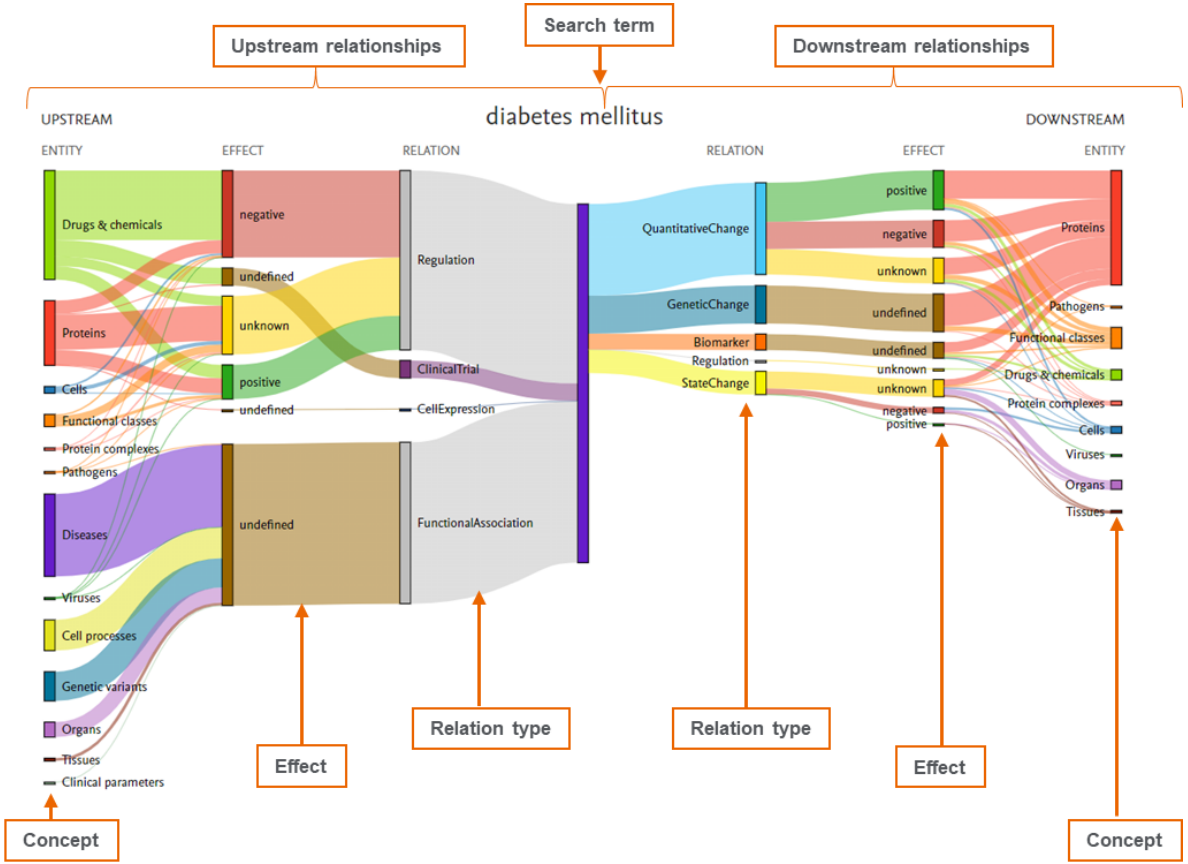
Starting a search





- Type a term into the search bar. After 3 letters are entered, suggestions will appear below the search bar. Either continue typing in the search term and press ‘enter’ or select from one of the suggested terms.
  - Search for any of the >1.2M concepts in EmBiology (drug, disease, protein, virus, cell process, etc.)
  - Note that several synonyms often refer to the same term and it is possible to type your preferred synonym for a concept. EmBiology will indicate what the preferred term for that concept is and results will include articles and clinical trials that mention both the preferred term and all synonyms.

Interpreting the visual overview of results



Upstream/downstream

- Upstream relationships represent biological concepts that **act o**n the search term (e.g., SIRT1 negatively regulates diabetes mellitus).
- Downstream relationships represent biological concepts that are **acted on by** the search term (IL6 has a positive quantitative change as a result of diabetes mellitus)

Concept

- Represent different biological entities (e.g., proteins, drugs, cell processes, etc) that have a relationship with the search term). For a description of all concept types, see page x.

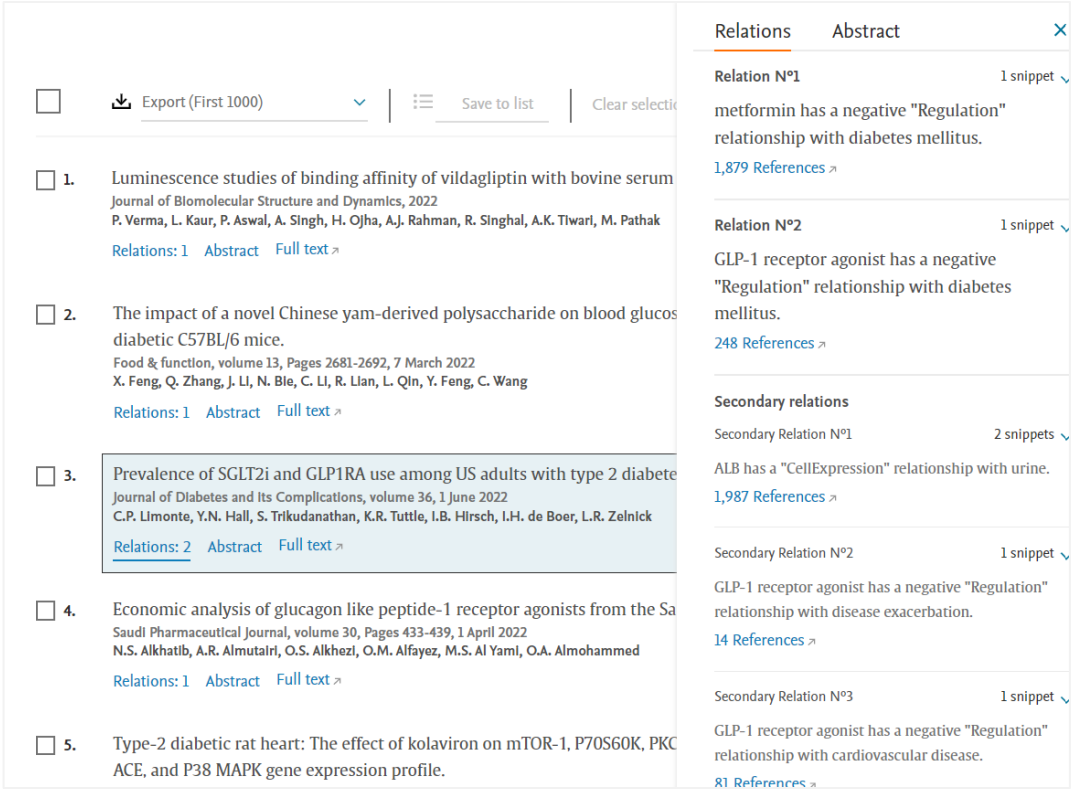
Effect

- Can be positive, negative or unknown (in cases where the effect is not mentioned) or undefined (for relationship types that do not act positively or negatively)

Relation type

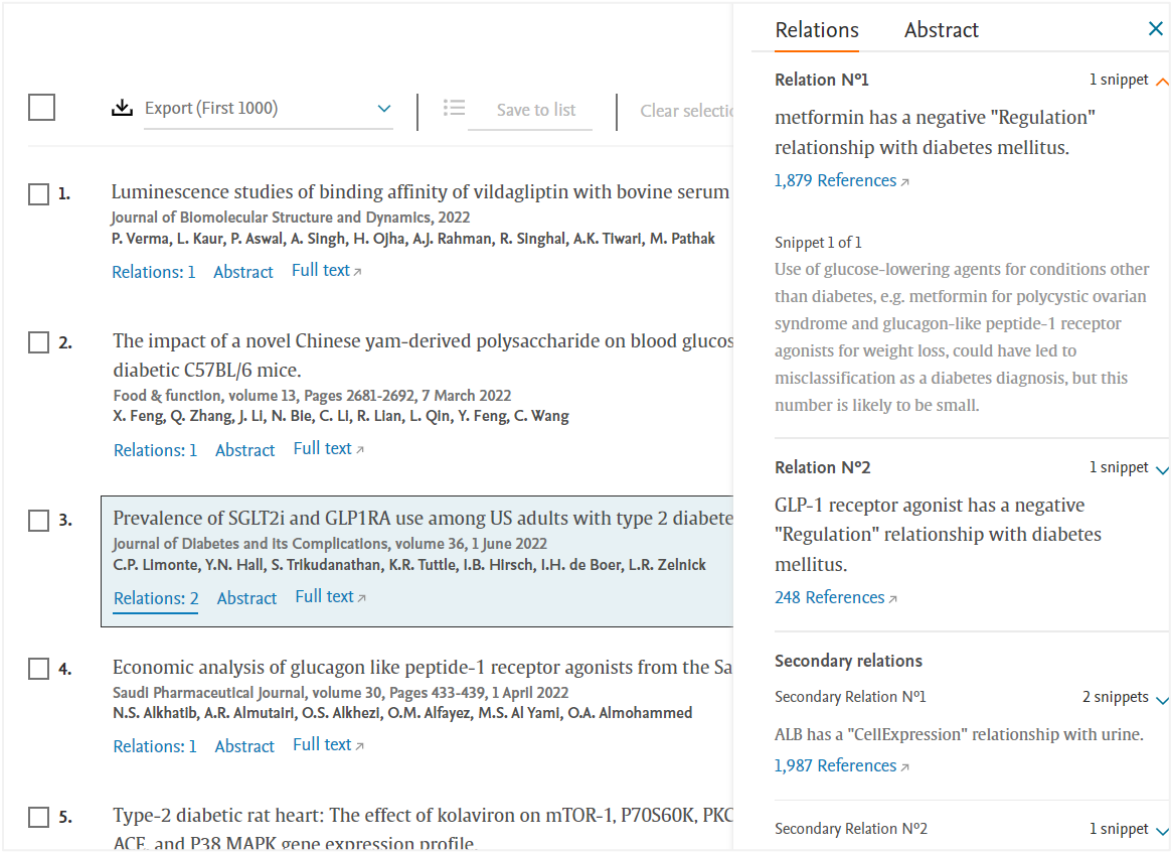
- Represent different types of relationships concepts have with the search term. For a description of relationships, see page x.

Understanding your results



- The results list includes all articles and clinical trials where a biological concept has a relationship with your search term
- Applying filters further narrows down the results list to those records with a specific type of relationship
- Each record includes information on the relations found in the record, a link to the abstract/clinical trial information and a link out to the full text/CT.gov

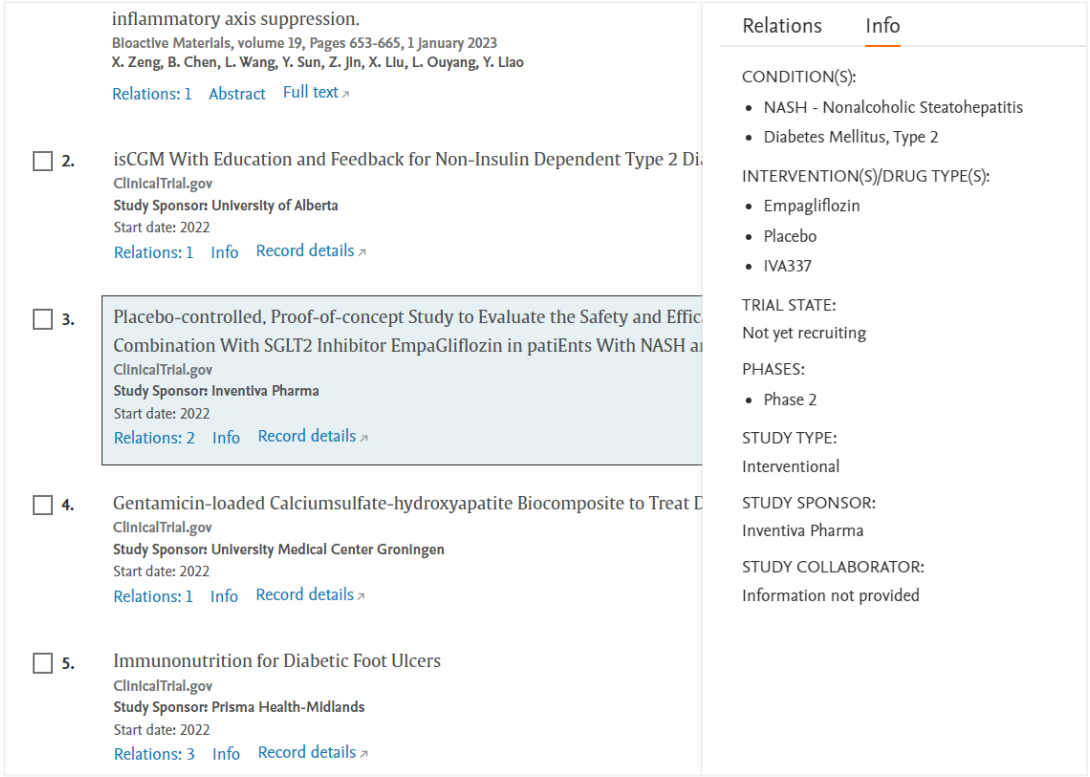
Understanding your results – Relations information



Relation information includes the following:

- **Sentence describing the relationship**: e.g., *metformin has a negative regulation relationship with diabetes mellitus*
- **One or more text snippets**: text from the article that was interpreted by automatic indexing to include the relationship. Snippets can be useful to see the context of the relationship and determine it's relevance to your work.
  (Clinical trials do not have text snippets)
- **Link to additional references mentioning the relationship**: opens up a new window of articles/trials that include the specific relationship and can be a useful indicator of how well studied/known is a relationship
- **Secondary relationships**: additional relationships found in the record that are not directly related to your search/filters

Understanding your results – Clinical Trial information

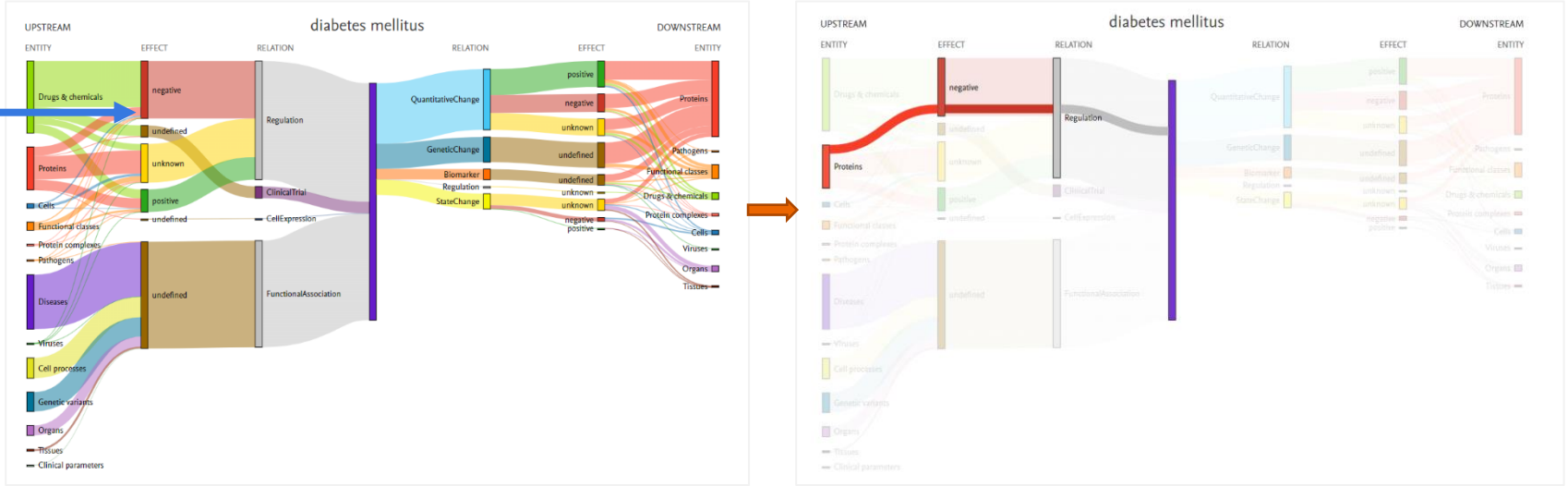


Clinical trials do not include Abstracts. Instead, they include:

- Condition(s)
- Interventions
- Trial State
- Phase(s)
- Study type
- Study sponsor
- Study collaborator

It is possible to filter for specific clinical trial information (e.g., trial phase) by exporting clinical trial results and applying filters in the exported file.

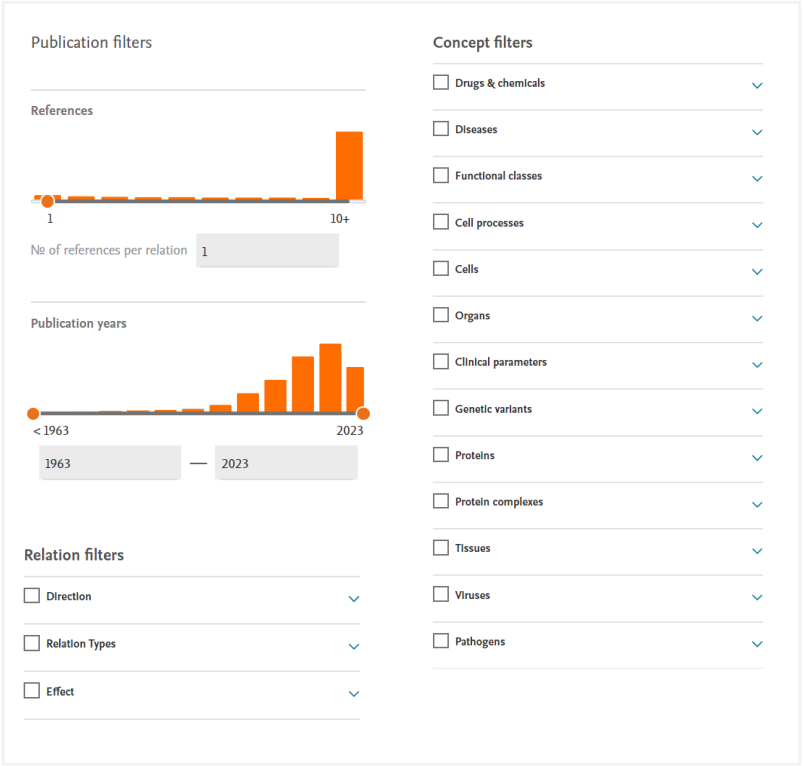
How to apply filters – single filter



It is possible to apply a single filter by clicking directly on the visual.

- For example, to filter by proteins that negatively regulate diabetes mellitus, click on the corresponding area on the graph. This will narrow down search results to those that include at least one **primary** relationship where a protein negatively regulates diabetes mellitus.

How to apply filters – multiple filters



Publication filters

- **References**: Allows filtering by a specific number of references per relation. For example, if the No. references per relation = 5, results will include articles/clinical trials where all primary relationships are supported by 5 or more references.
- **Publication years**: Allows filtering by specific years. For example, setting the lower limit to 1992 and upper limit to 2021 means results will include all articles/clinical trials published between the years 1992 and 2021.

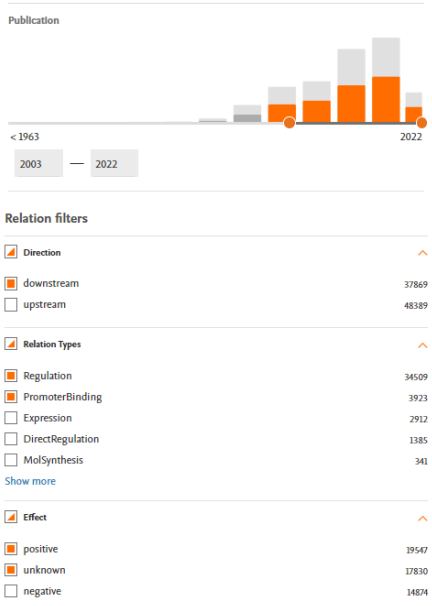
Relation filters

- **Direction**: results will include articles/clinical trials where all primary relationship are 1) upstream (meaning concepts **act on** the search term) or 2) downstream (concepts are **acted on by** the search term).
- **Relation**: results will include articles/clinical trials where all primary relationships are of the selected relation(s).
- **Effect**: results will include articles/clinical trials where all primary relationships are of the selected effect(s) – positive, negative, unknown (meaning the effect of the relationship has not been identified) and undefined (meaning the relationship type does not have a positive/negative effect)

Concept Filters: results will include articles/clinical trials where all primary relationships are of the selected concept(s). For example, selecting upstream, regulation, positive and protein means results will include articles/clinical trials where all primary relationships are proteins that have a positive regulation relationship with your search term.

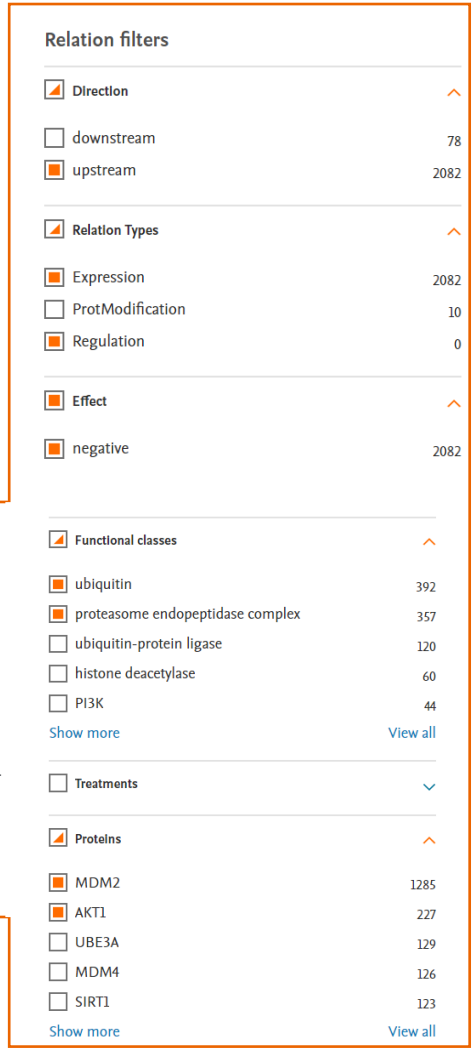
Filter behaviour

- Between different types of filters, selections are combined with AND
- Multiple selections within an individual filter category are combined with OR
  - **All selections under Concept filters combined with OR**



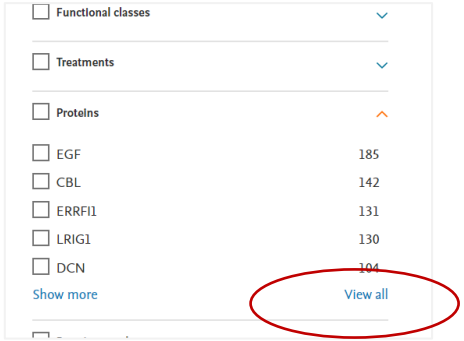
Results will include articles:
Between 2003-2022
AND
Downstream relations
AND
Regulation OR promoter binding relations
AND
Positive OR unknown effect

Results will include articles with:
Upstream relations
AND
Regulation OR Expression relations
AND
Negative effect
AND
Ubiquitin **OR** proteasome…. **OR** MDM2 **OR** AKT1

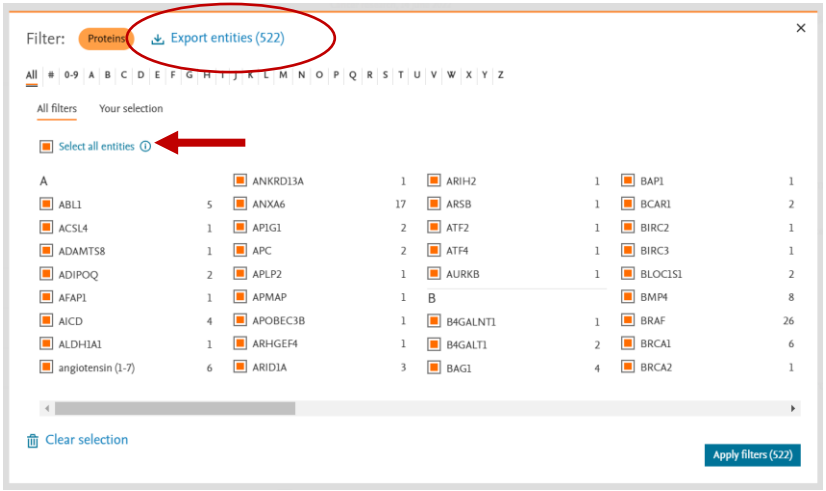


* **Apply Publication and Relation filters first, then Concept filters (the list of concept filters will be narrowed down to relevant options)**

Exporting relationships

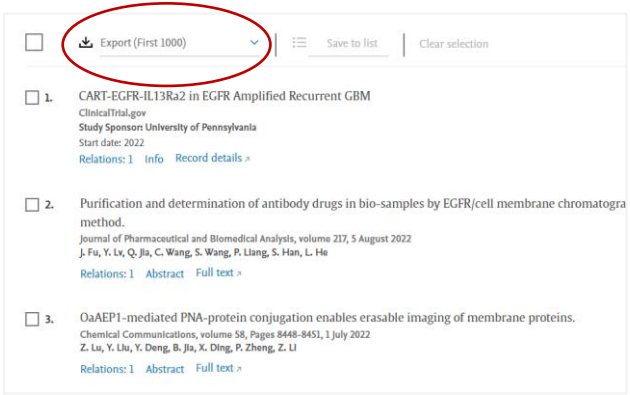


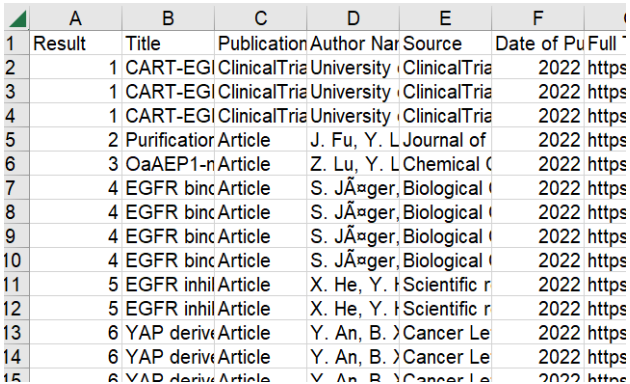
- Perform search and apply all relevant publication/relationship filters
- Using the filter pop-up panel, you can export information about individual concepts or the entire set of information for a concept. For example:
  - Search by EGFR
  - Apply upstream/positive relationship filters
  - Expand the Concept filter – Protein and select View all.
  - A popup shows all proteins that are upstream of EGFR with a negative relationship. Select individual proteins or Select all entities – you can export this list or apply filters from this popup.



# Date of export: 02.02.2023						
# Search term: EGFR						
# Filters:						
# Publication:						
# Relation: Direction: UPSTREAM; Effects: negative						
# Concept: Protein						
Search Te	Acts on / is acted	Concept	Relation Object Typ	Relation E	# of refere	Link
EGFR	is acted on by	EGF	Expression	negative	188	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	LRIG1	DirectRegulation	negative	121	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	LRIG1	Expression	negative	43	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	CBL	Expression	negative	147	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	ERRFI1	DirectRegulation	negative	129	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	ERRFI1	Expression	negative	14	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	DCN	DirectRegulation	negative	101	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	DCN	Expression	negative	10	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	MIR7-1	miRNAEffect	negative	99	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	SNX1	DirectRegulation	negative	33	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	SNX1	Expression	negative	33	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	MIR133A1	miRNAEffect	negative	54	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	HGS	DirectRegulation	negative	30	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	HGS	Expression	negative	22	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	MIR146A	miRNAEffect	negative	48	https://embio.elsevier.com/reference?query=EGFI
EGFR	is acted on by	MIR145	miRNAEffect	negative	35	https://embio.elsevier.com/reference?query=EGFI

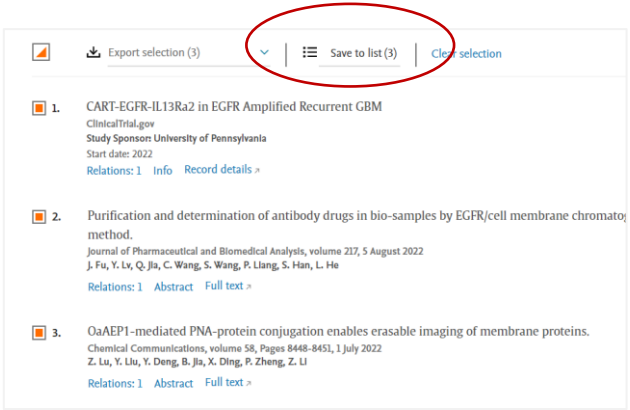
Saving and exporting results

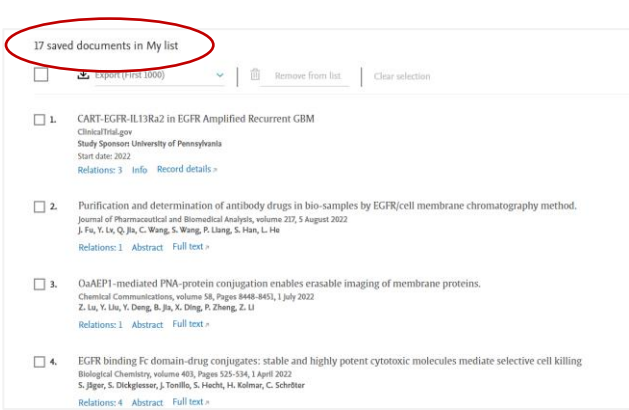




	A	B	C	D	E	F	
1	Result	Title	Publication	Author Nar	Source	Date of Pu	Full
2	1	CART-EGI	ClinicalTria	University (	ClinicalTria	2022	https
3	1	CART-EGI	ClinicalTria	University (	ClinicalTria	2022	https
4	1	CART-EGI	ClinicalTria	University (	ClinicalTria	2022	https
5	2	Purificatior	Article	J. Fu, Y. L	Journal of	2022	https
6	3	OaAEP1-n	Article	Z. Lu, Y. L	Chemical (	2022	https
7	4	EGFR binc	Article	S. JÃ¤ger,	Biological (	2022	https
8	4	EGFR binc	Article	S. JÃ¤ger,	Biological (	2022	https
9	4	EGFR binc	Article	S. JÃ¤ger,	Biological (	2022	https
10	4	EGFR binc	Article	S. JÃ¤ger,	Biological (	2022	https
11	5	EGFR inhil	Article	X. He, Y. H	Scientific r	2022	https
12	5	EGFR inhil	Article	X. He, Y. H	Scientific r	2022	https
13	6	YAP derive	Article	Y. An, B.)	Cancer Le	2022	https
14	6	YAP derive	Article	Y. An, B.)	Cancer Le	2022	https
15	6	YAP deriv	Article	Y. An, B.)	Cancer Le	2022	https

- Export up to 1000 results or select individual results
- Export files include all bibliographic information, relationship information, text snippets and clinical trail information





- Save records (articles or clinical trials) to a list. Access your saved list by clicking ‘Saved lists’ (top right-hand of the screen.

Glossary of terms – Types of filters

Filter	Description
Publication filters	
Reference	Results will only include articles with relations supported by a specific number of references. Setting the filter to ‘5’ means you will only see relations that have been reported in 5+ articles
Publication (year)	Results will include articles published on/before/between/after specific dates. Adjust the slide or fill in numbers for the publication years for which you would like to limit results
Direction filters	
Upstream or downstream	Results will include relations that are upstream or downstream of the search term. Upstream relations: search term is the ‘object’ (concepts act in some way on the search term). Downstream relations: search term is the ‘subject’ (it acts in some way) on downstream concepts
Effect filters	
Positive, negative, unknown or undefined	Most relations can have an effect (positive/negative). In cases where this information is not found in the snippet of text, the effect will be labelled ‘unknown’. In cases where there is no effect for the relationship type (Binding, Cell Expression, Functional Association, the effect is labelled ‘undefined’.
Relation filters	Articles in the results list will be selected based on the Search Term and relations selected in the Relations filter (see detailed information in slides 2-3)
Concept filters	Articles in the results list will be selected based the Search Term and concepts selected in the Concepts Filter. See detailed information in slide 4

Glossary of terms – Types of relations and examples

Relation	Description	Examples
Expression	Changes protein abundance by affecting levels of transcript or protein stability	MDM2 has a negative "Expression" relationship with TP53
miRNA Effect	Inhibitory effect of an miRNA on its mRNA target	miR-30 has a negative "miRNAEffect" relationship with TP53
Promoter Binding	Binds to the promoter of a gene	FOXC1 has a positive "PromoterBinding" relationship with MMP7
Regulation	Changes activity by an unknown mechanism (may be direct or indirect	SOCS3 has a positive "Regulation" relationship with diabetes mellitus
Direct Regulation	Influences activity by direct physical interaction	BRCA1 has a "DirectRegulation" relationship with BARD1
Binding	Direct physical interaction between two molecules	FANCD2 has a "Binding" relationship with BRCA1
Protein Modification	Changes the modification of the target molecule, usually by direct interaction	SRC has a "ProtModification" relationship with GRB2
Biomarker	Identification of proteins/complexes/functional classes/metabolites that are prognostic or diagnostic biomarkers for a disease	Lung cancer has a "Biomarker" relationship with IL6
Genetic Change	Genetic changes such as gene deletions, amplifications, mutations or epigenetic changes	Lung cancer has a "GeneticChange" relationship with ALK
Quantitative Change	Changes in abundance/activity/expression of a gene/protein/small molecule in a disease state	Breast cancer has a positive "QuantitativeChange" relationship with AGK
State Change	Changes in a protein's post-translational modification status or alternative splicing events	Breast cancer has a "StateChange" relationship with estrogen receptor

Types of relations and examples (continued)

Relation	Description	Examples
Functional Assoc.	Between a disease and a cellular process or another disease	Chronic pancreatitis has a "FunctionalAssociation" relationship with pancreatic cancer
Chemical Reaction	Either enzyme catalysed or spontaneous chemical reaction	CYP3A has a "ChemicalReaction" relationship with ticagrelor
Molecular Synthesis	Changes the concentration of the target	CYP3A has a "MolSynthesis" relationship with midazolam
Molecular Transport	Changes the localization of the target	Tamoxifen has a positive "MolTransport" relationship with MAPK3
Clinical Trial	Clinical trials conducted for a drug against a disease (from CT.gov)	Tamoxifen has a positive “Clinical Trial" relationship with breast cancer*
Cell Expression	Expression of proteins within or on the surface of a cell	Hepatocyte has a "CellExpression" relationship with EGFR

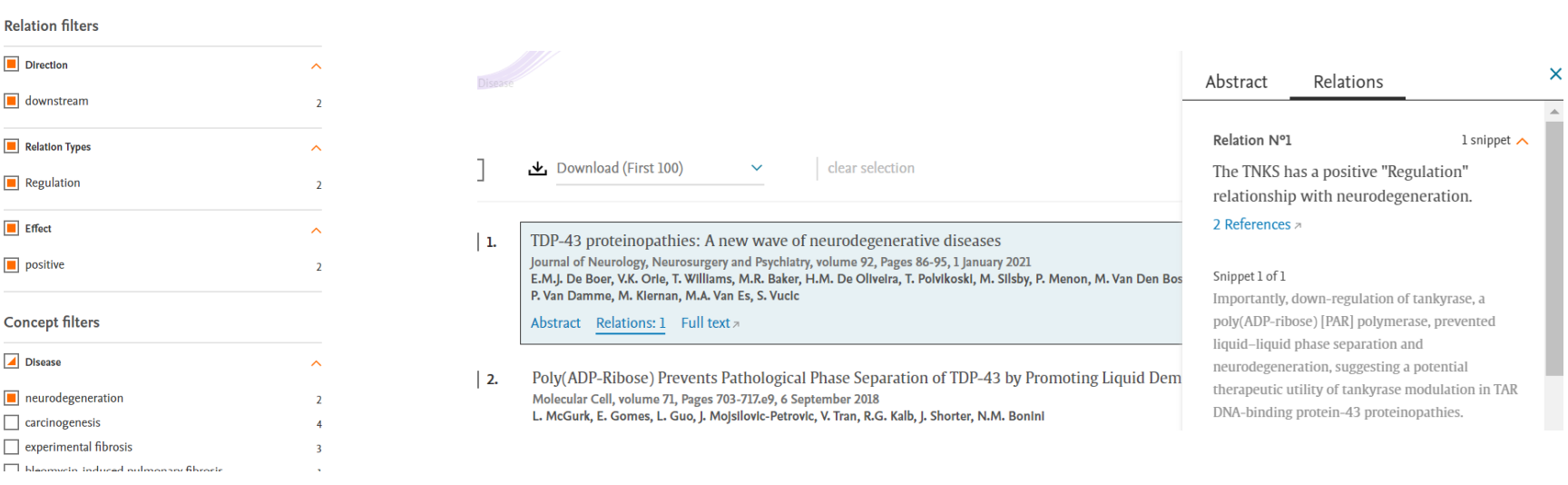
Glossary of terms – Types of concepts

Concept	Description	Examples
Drugs & chemicals	Naturally occurring metabolites, small molecules found in cell, drugs (incl. small molecules & biologics)	Cisplatin, glucose, ATP, trastuzumab
Diseases	Health conditions and disease terms	Neoplasm, acute lung injury
Proteins	Represents both genes and the gene products, including proteins and miRNAs	TP53, mir293
Functional classes	Proteins classes based on biological function	DNA helicase, CYP3A
Complexes	One or more polypeptides that form a complex via physical interactions	Ribosome, p53-MDM2
Genetic variant	Variants searchable by SNPID and text	rs11553421, mutation in TP53, 524G>T (p.Arg175Leu)
Cells	Mammalian cell types & cell lines	T-cell, fibroblast, 298T
Cell Process	Biological processes	DNA damage, cytokinesis
Organs	Mammalian organ types	Brain, heart, lung
Tissues	Mammalian tissue types	Bone marrow, endothelium
Viruses	An agent that causes infectious diseases	Influenza A, Sars-CoV-2, HIV-1
Clinical parameter	Measured parameters of the human body used in clinical practice	Overall survival, platelet count
Treatment	Non-chemical treatments and environmental conditions	Radiation, hypoxia, stress
Cell object	Organelles and other structural components of the cell	Mitochondria, chromatin

Example workflows

Starting with a potential protein target: what diseases can be affected by my protein?

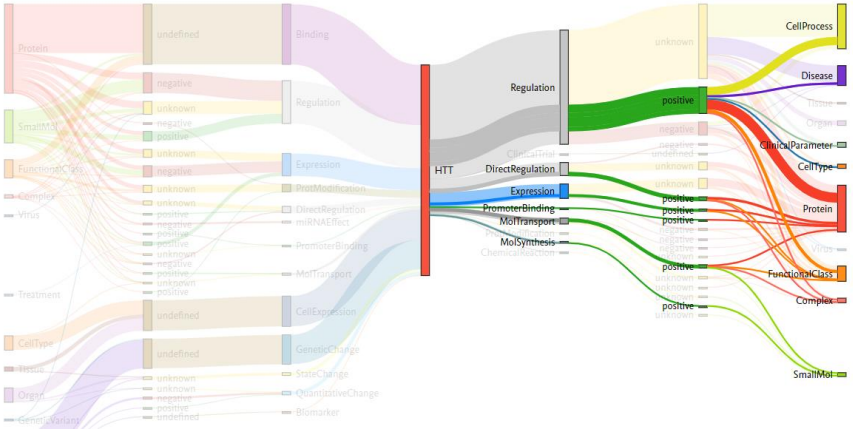
- Enter protein (e.g., TANK1) into the Search Field (dropdown will help you to find correct term)
  - If text indicates ‘TANK1: use TNKS’ – TNKS is the preferred term will appear in subsequent fields
- Apply filters: downstream, positive, disease
  - Articles in the results list include TANK1 positively regulating a disease (by an unknown mechanism)
- Look at all diseases positively regulate by TANK1 and/or further narrow down diseases using the Disease concept filter (de-select All diseases, select relevant ones and click ‘Apply’)



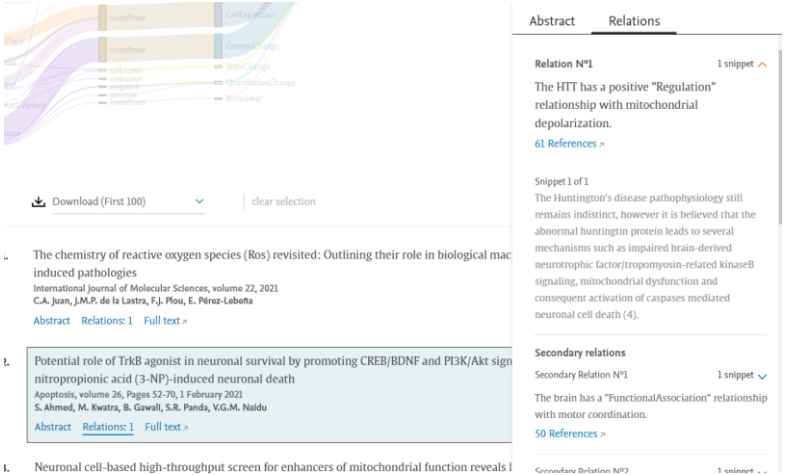
Starting with protein: what do increases in my protein cause?

- Enter protein (e.g., HTT) into the Search Field (dropdown will help you to find correct term)
- Apply filters: downstream, positive

TIP: It is not necessary to be an expert all categories – at first, select the fewest filters and leave the rest unchecked to see what information appears. Then narrow down as needed.



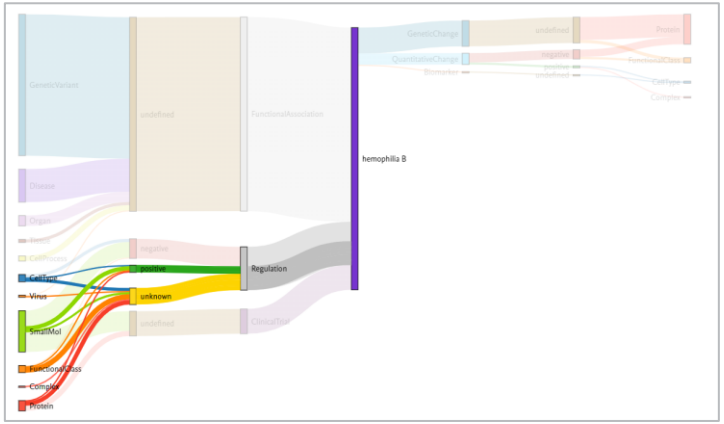
Articles in the results list include any concept that is positively affected by HTT



Starting with disease

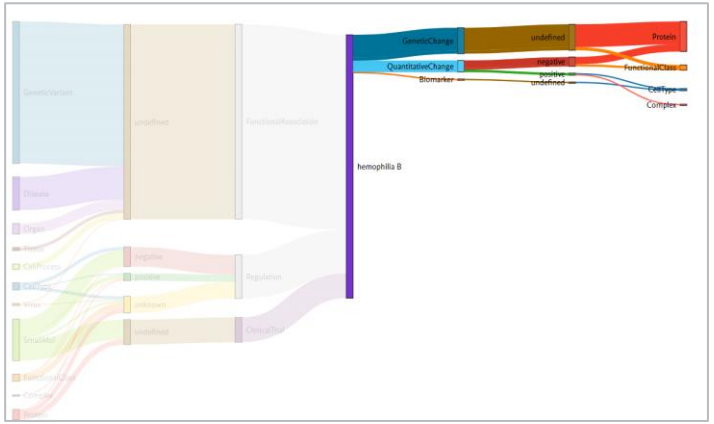
What are potential drug targets?

- Enter disease (Hemophilia B) into the Search Field (dropdown will help you to find correct term – you many also wish to look under the general term of Hemophilia)
- Apply filters: upstream, regulation, positive, unknown
  - Articles in the results list include concepts regulating a disease (either positively or with unknown effect) – it is important not to exclude ‘unknown’ effect as EmBio will only indicate an effect is ‘positive’ if it is in the snippet of text from where the relationship is extracted



What are potential biomarkers?

- Enter disease (Hemophilia B) into the Search Field (dropdown will help you to find correct term)
- Apply filters: downstream (Genetic change, Quantitative Change, State Change, Biomarker)
  - Articles in the results list include concepts that are changed as a result of the disease and could therefore be measured



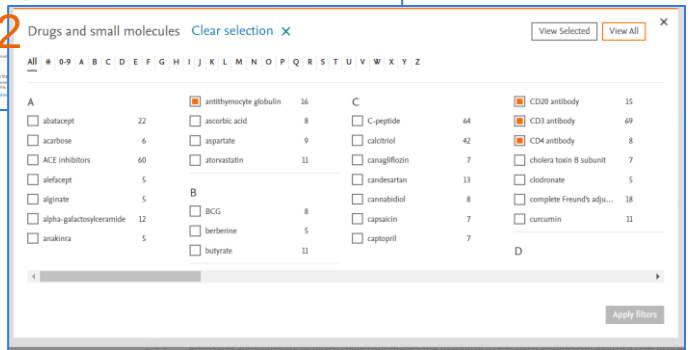
Starting with disease

What biologics negatively regulate the disease?

1. Enter disease (Type 1 diabetes) into the Search Field (there are other types of diabetes you may also want to search)
- Apply filters: Direction: upstream; Effect: negative, Relations: 5+ (this limits the search to relations that are reported in 5+ publications)
2. Then view the Drugs & Small molecules filter popup to select the individual names of relevant biologics.

1

2



3. Results include articles that have a relationship with Type 1 diabetes negatively regulated by one of the selected biologics
4. The results list can be exported for an overview of who is publishing in this field.

3

Tips for using EmBiology

- Apply the Publication filters and Relation filters before making selections in the concept filters – the available options in concept filters will depend on the years and relationships you've selected
- If you are applying multiple concept filters in different categories, apply your selection from one category before making selections in the other
- Do you only want to see well studied relationships? Use the 'References' filter to limit result. (e.g., setting the 'References' filter to 5 limits results to articles that include relations that have been reported 5 or more times
- It is not necessary to be an expert in the types of content that is in all categories – e.g., if you are interested in what concepts may positively affect your search term, select upstream and positive – leave the rest unchecked and see what categories appear
- Concept names that appear in left-hand filters and pop-ups are the preferred terms, which can be frustrating if you're searching for a filter by a different name.
    - To find out what the preferred term is, open up an new window of EmBio and type your term in the search box – the dropdown suggestions will include the preferred term (e.g., if you would like to filter by CD20 but it's not in the list of available filters, see in the dropdown that the preferred term is MS4A1 and search for this term. Note that synonym-based searching of filters is on our 2022 roadmap!)