

Information on CD-CODE is presented in seven pages (Home, About, Proteins, Condensates, Statistics, Encyclopedia and Help). This guide is aimed to explain how to use those CD-CODE pages.

1. Select whether you are looking for a protein or a condensate

2. Choose organism here

Species Protein Condensate

Species: All

Search: Enter Uniprot ID, gene name, or protein name

Examples: [UNE6](#), [nucleolus__3702](#)

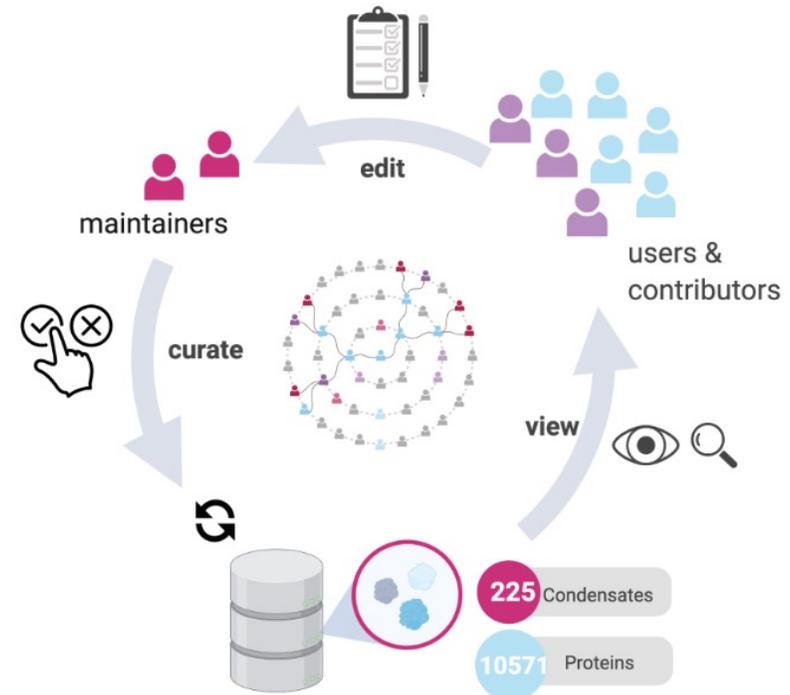
3. Insert protein or condensate name

4. Search here

[Search](#)

CD-CODE (CrowDsourcing COndensate Database and Encyclopedia) is a comprehensive, semi-manually curated crowdsourcing database of biomolecular condensates and their constituents as well as an encyclopedia for the scientific terms used to describe them. Biomolecular condensates are membrane-less organelles that are formed by liquid-liquid phase separation (LLPS). In cells, phase separation is driven by driver proteins and nucleic acid, which recruit client molecules into the condensate. CD-CODE allows for the dynamic and fast addition, and review of new condensates and proteins by contributing users. CD-CODE is a semi-manually curated and annotated database that aggregates information from primary literature and other protein and LLPS databases. (Dev)

[Read more](#)



Search result for searched protein (CRB)

Filter

Functional Type

Previous 1 Next

Uniprot	Gene Name	Name	Functional Type	Species
A0A1P8ATL2 (A0A1P8ATL2_ARATH)	CRB	Chloroplast RNA binding protein	client	Arabidopsis thaliana (3702)

Showing 1 to 1 of 1 entries

Click on this link to go to protein detail page

Previous 1 Next

Protein detail page

CRB

General Information

1. The upper part shows details of the protein searched and relevant links (e.g. Uniprot, Pubmed).



Name	Protein Chloroplast RNA binding protein
Species	Arabidopsis thaliana (3702)
Ensembl ID	
Ensembl Gene ID	
UniProt	A0A1P8ATL2 (A0A1P8ATL2_ARATH)
Antibodies	A0A1P8ATL2
Functional Type	client
Pubmed	29104584
Sequence	<code>MSKEFYELLRHNCESF</code>

Protein details can be edited by contributors.

2. The bottom part shows condensates to whom the searched protein is part of, the number of proteins in each condensate and the species.



Condensates

Name	No. of Proteins	Species
Nucleolus	1881	Arabidopsis thaliana

CRB

General Information

Name	<p>- Driver: A protein that undergoes phase separation or self assembles to form liquid droplets independent of other proteins, induces the formation of a condensate, or is essential for the integrity of a condensate. This protein is known to be a driver in at least one of the known condensates.</p> <p>- Client: A protein that is part of a condensate, but is driven to the condensate by a driver protein. This protein is known to be a client in at least one of the known condensates, but never known to be a driver.</p> <p>- Regulator: A protein that biochemically or enzymatically regulates the formation of a condensate, but is not a part of the condensate. This protein is never known to be a driver or client.</p>
Species	
Ensembl ID	
Ensembl Gene ID	
UniProt	
Antibodies	
Functional Type ⓘ	
Pubmed ⓘ	
Sequence	

This (i) button explains how protein functional types are defined in CD-CODE.

CRB

General Information

Name	<p>Publications providing evidence for protein's LLPS behavior in general</p>
Species	
Ensembl ID	
Ensembl Gene ID	
UniProt	
Antibodies	
Functional Type ⓘ	
Pubmed ⓘ	
Sequence	

This (i) button explains the PMIDs.

Condensate search

Protein Condensate

Species

Search

Nucleolus: Mus musculus
Stress Granule: Homo sapiens
Rbfox1 Rnp Granule: Drosophila melanogaster
Search for **Nucleolus**

1. Type condensate name here

Search

3. Search here

2. Condensate can be selected from the suggestions

4. Search results

Filter

Previous 1

Name	No. Proteins	Species	Confidence Score
Nucleolus	21	Mus musculus (10090)	★★★★★
Nucleolus	1881	Arabidopsis thaliana (3702)	★★★★★
Nucleolus	1415	Homo sapiens (9606)	★★★★★
Nucleolus	248	Bos taurus (9913)	★★★★★
Nucleolus	116	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (559292)	★★★★★
Nucleolus	1	Xenopus laevis (8355)	★★★★★
Nucleolus	1	Felis catus (9685)	★★★★★
Nucleolus	13	Rattus norvegicus (10116)	★★★★★
Nucleolus	1	Oreochromis niloticus (8128)	★★★★★
Nucleolus	3	Oryctolagus cuniculus (9986)	★★★★★

5. Click the condensate you are searching for to go to the condensate detail page

Nucleolus

General Information

Species: Mus musculus (10090)

Description: Nucleolus is the largest and the first membrane-less organelle which was identified. It is composed of a fibrillar centre (FC), dense fibrillar component (DFC) and granular component (GC). The nucleolus is a dynamic structure that disassembles during mitosis and reassembles after cell division. The nucleolus is responsible for ribosome biogenesis in eucaryotic cells, which are essential macromolecular machines responsible for synthesizing all proteins required by the cell. The nucleoli are now known to act as sensors and regulators of cellular responses, such as apoptosis, DNA damage response, epigenetic modification, genome stability, proliferation, senescence and signal transduction. 

Markers 

No. of Proteins: 21

Evidence star:     

[Q80Z11](#) 

1. The upper part shows details if the condensate searched for, such as description and the marker of the condensate.

New protein can be added to condensates.

Proteins

[Add a protein to this condensate](#)

2. The lower part shows information on the protein members of the condensate searched for.

Data can be edited by registered users as contributor.

Show 10 entries

Filter

Previous 1 2 3 Next

[Download CSV](#)

UniProt	Gene Name	Name	Pubmed	Functional Type	Experimental Evidence	Driver Criterion	Confidence Score	Action
P60843	Elf4a1	Eukaryotic initiation factor 4A-I	15635413				★★★★★	Edit Remove
P61161	Actr2	Actin-related protein 2	12429849 15635413				★★★★★	Edit Remove
P61965	Wdr5	WD repeat-containing protein 5	15635413				★★★★★	Edit Remove
P62242	Rps8	40S ribosomal protein S8	15635413				★★★★★	Edit Remove
P62245	Rps15a	40S ribosomal protein S15a	12429849 15635413				★★★★★	Edit Remove
P62264	Rps14	40S ribosomal protein S14	12429849 15635413				★★★★★	Edit Remove
P62309	Snrpg	Small nuclear ribonucleoprotein G	15635413				★★★★★	Edit Remove
P62320	Snrpd3	Small nuclear ribonucleoprotein Sm D3	15635413				★★★★★	Edit Remove
P62849	Rps24	40S ribosomal protein S24	15635413				★★★★★	Edit Remove
P62852	Rps25	40S ribosomal protein S25	15635413				★★★★★	Edit Remove

Confidence score of each protein being part of the condensate.

CD-CODE

CD-CODE (CrowD sourcing COndensate Database and Encyclopedia) is a comprehensive, semi-manually curated crowdsourcing database of **biomolecular condensates** and their constituents as well as an **encyclopedia** for the scientific terms used to describe them. Biomolecular condensates are membrane-less organelles that are formed by **liquid-liquid phase separation (LLPS)**. In cells, phase separation is driven by **driver** proteins and nucleic acid, which recruit **client** molecules into the condensate.

CD-CODE allows for the dynamic and fast addition, and review of new condensates and proteins by contributing users. CD-CODE is a semi-manually curated and annotated database that aggregates information from primary literature and other protein and LLPS databases.

1. Uniprot: <https://uniprot.org>
2. PhaSePro: <https://phasepro.elte.hu>
3. LLPSDB: <http://bio-comp.org.cn/lpsdb/home.html>
4. DrLLPS: <http://llps.biocuckoo.cn>
5. PhaSepDB: <http://db.phasep.pro>

If you are interested in becoming a contributor, please join us [here](#).

Protein functional types are explained.

Definitions at CD-CODE

Biomolecular condensates: Membrane-less organelles formed by LLPS observed *in vivo*.

Synthetic condensates: *In vitro* experiments of LLPS.

Markers: Proteins that are used to define and label/mark a specific biomolecular condensate. They can help in unique identification of membrane-less organelles. Markers are often drivers.

Proteins in CD-CODE are divided into four groups:

1. **Drivers:** Proteins which fulfill at least one of the following criteria (confidence score max. 3 stars).
 - Undergo phase separation or self-assembly into liquid droplets independent of other proteins (1/3 stars).
 - Induce the formation of a condensate (2/3 stars).
 - Are essential for the integrity of a condensate (3/3 stars).
2. **Clients:** Proteins which are part of a condensate, but they are recruited into the condensate by driver proteins. They are not essential to the integrity of a condensate.
3. **Regulators:** Proteins which biochemically or enzymatically regulate the formation of a condensate, but are structurally not part of it.
4. **Unknown:** Proteins that have been associated with condensates but for which the role is unknown.

It is important to note that a particular protein may be a driver in one condensate and a client in another one. Moreover, the formation of condensates might be regulated by regulator proteins through specific biochemical activities, but they do not necessarily need to be a part of the condensate.

Confidence score for a protein being in a condensate (max. 5 stars):

Number of stars	Description
★☆☆☆☆	PubMed reference annotated
★★☆☆☆	High throughput experiment (eg. Mass spect.)
★★★☆☆	<i>In vitro</i>
★★★★☆	<i>In cellulo</i>
★★★★★	<i>In vivo</i>

CD-CODE scoring system is explained.

Confidence score for a condensate (max. 5 stars): mean of the confidence scores of its member proteins.

Protein page



Proteins

Proteins can be filtered and sorted according to functional types.

List of proteins can be downloaded here.

Filter

Functional Type

[Download CSV](#)

Uniprot	Gene Name	Name	Functional Type	Species
P62750 (RL23A_HUMAN)	RPL23A	60S ribosomal protein L23a	client	Homo sapiens (9606)
Q9GZQ8 (MLP3B_HUMAN)	MAP1LC3B	Microtubule-associated proteins 1A/1B light chain 3B	client	Homo sapiens (9606)
Q32P51 (RA1L2_HUMAN)	HNRNPA1L2	Heterogeneous nuclear ribonucleoprotein A1-like 2	driver	Homo sapiens (9606)
O43561 (LAT_HUMAN)	LAT	Linker for activation of T-cells family member 1	driver	Homo sapiens (9606)
P06748 (NPM_HUMAN)	NPM1	Nucleophosmin	driver	Homo sapiens (9606)
Q14011 (CIRBP_HUMAN)	CIRBP	Cold-inducible RNA-binding protein	driver	Homo sapiens (9606)
P22626 (ROA2_HUMAN)	HNRNPA2B1	Heterogeneous nuclear ribonucleoproteins A2/B1	driver	Homo sapiens (9606)
Q13151 (ROA0_HUMAN)	HNRNPA0	Heterogeneous nuclear ribonucleoprotein A0	client	Homo sapiens (9606)
P09651 (ROA1_HUMAN)	HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens OX=9606 GN=HNRNPA1 PE=1 SV=5	driver	Homo sapiens (9606)
O43670 (ZNF207_HUMAN)	ZNF207	BUB3-interacting and GLEBS motif-containing protein ZNF207	driver	Homo sapiens (9606)

Condensate page

List of condensates can be sorted and new condensates will be added.



- Biomolecular Condensates
- Synthetic Condensates
- Add New Condensate

Condensates

Filter

Download CSV

Name	No. Proteins	Species	Confidence Score
Nucleolus	1881	Arabidopsis thaliana (3702)	★★★★★
Postsynaptic Density	1487	Mus musculus (10090)	★★★★★
Nucleolus	1415	Homo sapiens (9606)	★★★★★
Synaptic vesicle pool condensate	1382	Homo sapiens (9606)	★★★★★
Stress granule	1006	Homo sapiens (9606)	★★★★★
Mitochondrial cloud	637	Xenopus laevis (8355)	★★★★★
P-body	597	Homo sapiens (9606)	★★★★★
Centrosome	549	Homo sapiens (9606)	★★★★★
Stress Granule	454	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (559292)	★★★★★

Biomolecular Condensates

Synthetic Condensates

Add New Condensate

List of biomolecular condensates is shown here and can be downloaded.



Download CSV

Biomolecular Condensates (Membrane-less Organelles)

Filter

Name	No. Proteins	Species	Confidence Score
Nucleolus	1881	Arabidopsis thaliana (3702)	★★★★★
Postsynaptic Density	1487	Mus musculus (10090)	★★★★★
Nucleolus	1415	Homo sapiens (9606)	★★★★★
Synaptic vesicle pool condensate	1382	Homo sapiens (9606)	★★★★★
Stress granule	1006	Homo sapiens (9606)	★★★★★
Mitochondrial cloud	637	Xenopus laevis (8355)	★★★★★
P-body	597	Homo sapiens (9606)	★★★★★

Biomolecular Condensates

Synthetic Condensates

Add New Condensate

List of synthetic condensates is shown here and can be downloaded.



Download CSV

Synthetic Condensates (Experiments)

Filter

Name	No. Proteins	Species	Confidence Score
Synthetic Condensate 000283	10	Chimeras (0000)	★★★★★
Synthetic Condensate 000375	10	Homo sapiens (9606)	★★★★★
Synthetic Condensate 000331	8	Drosophila melanogaster (7227)	★★★★★
Synthetic Condensate 000340	6	Drosophila melanogaster (7227)	★★★★★
Synthetic Condensate 000373	6	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (559292)	★★★★★
Synthetic Condensate 000372	6	Homo sapiens (9606)	★★★★★

'Add New Condensate' function requires the contributor to fill a form and provide evidence for their novel condensate to be added to CD-CODE.

Biomolecular Condensates

Synthetic Condensates

Add New Condensate

New Condensate

Synthetic/Biomolecular Synthetic Biomolecular

Condensate Name

Name

Proteins

Search

UNIPROT

FUNCTIONAL TYPE

PUBMED

DRIVER CRITERION

EXPERIMENTAL EVIDENCES

Synonyms

Synonyms (comma-separated)

Description

Description

Comments

Please provide evidence/PubMed ID.

Add Condensate

Cancel

Details of proteins included need to be provided.

Proteins

Search

UNIPROT FUNCTIONAL TYPE PUBMED DRIVER CRITERION EXPERIMENTAL EVIDENCES

Aliases

Description

Comments

Please, specify **Functional Type, PubMed IDs, Driver Criterion** and **Experimental**

Uniprot ID * 

Functional Type

PubMed IDs *

Driver Criterion

- Self-PS
- Induces Assembly
- Essential For Integrity

Experimental Evidences

- In Vitro
- In Vivo
- In Cellulo
- Mass Spectrometry
- Colocalization
- FRAP
- Others

The UniProt ID (accession) is the 6 to the 10-character-long name displayed on the heading of the protein page from Uniprot. It is a unique identifier for a protein from the whole Uniprot database (https://www.uniprot.org/help/accession_numbers). Please note that different proteins from the same gene will also have different Uniprot IDs. For example, The Uniprot ID of "FUS Human" is P35637 (<https://www.uniprot.org/uniprot/P35637>). It is usually the last part of the URI of the HTTPS link of the protein page.

Uniprot ID and relevant PMIDs need to/ maybe provided.

Contributors can determine the functional type of their protein here.

Criteria for defining a protein as a driver need to be determined .

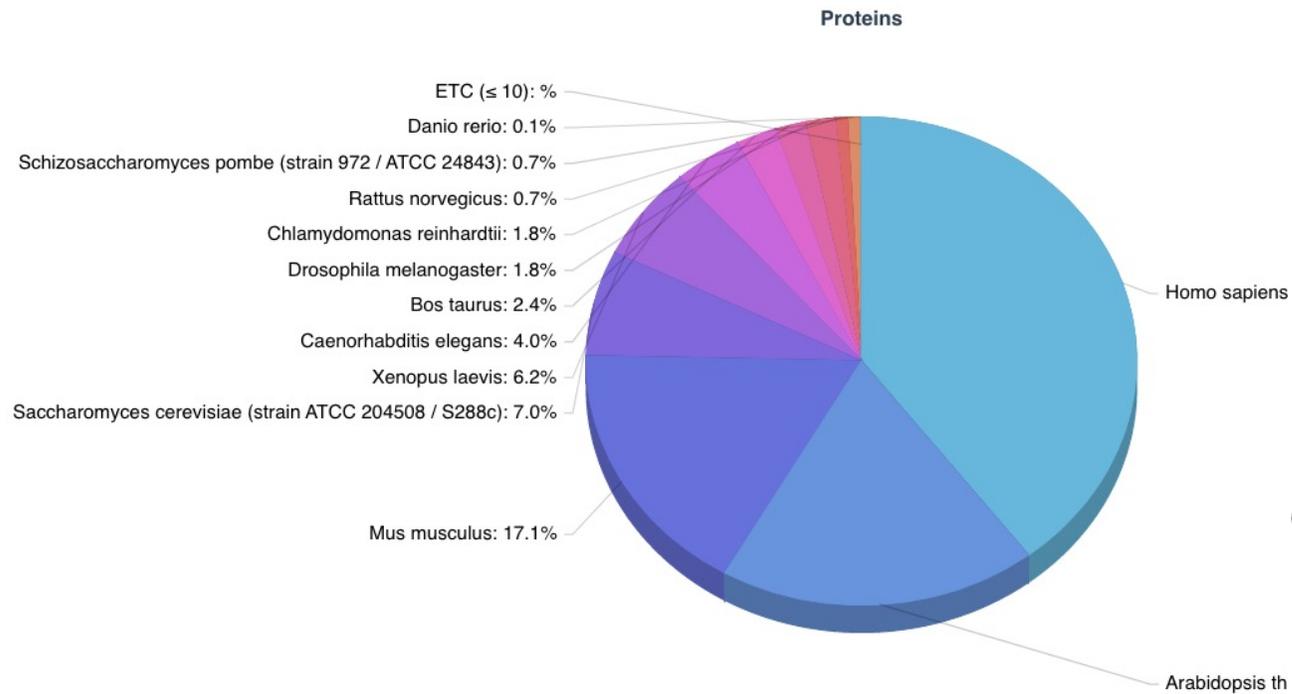
Experimental details must be provided.

Statistics page

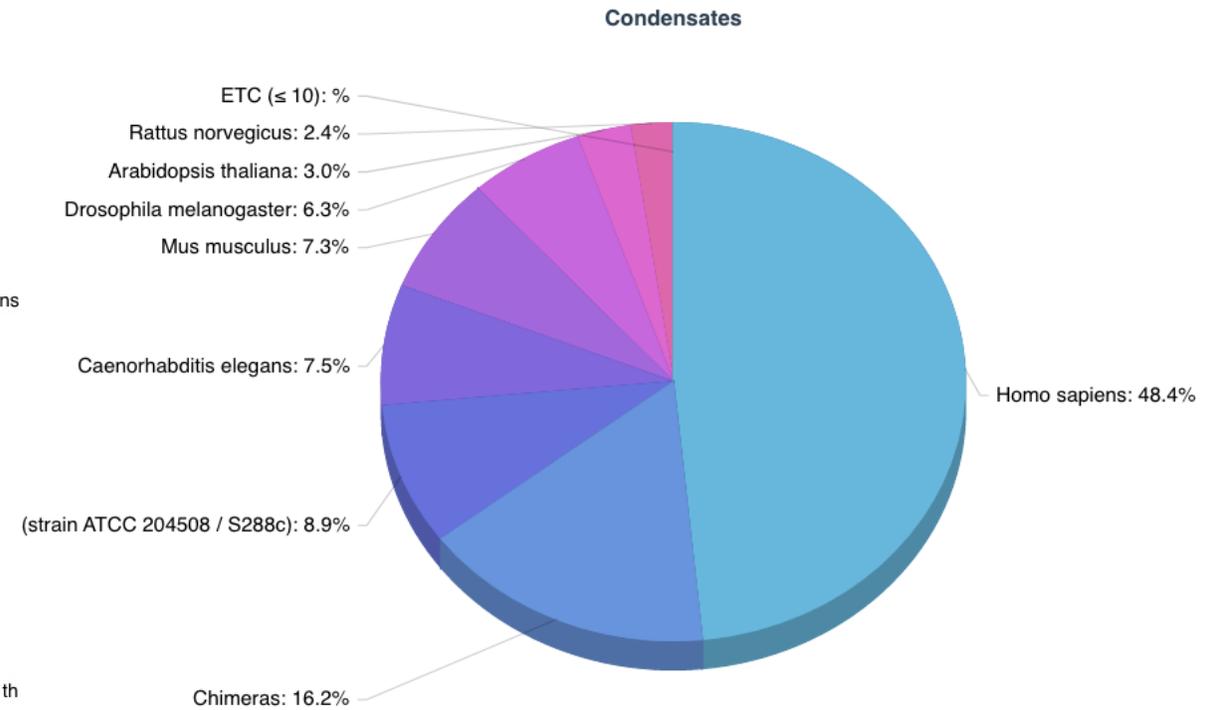
Distribution of data on CD-CODE is shown here.

[Home](#) [About](#) [Proteins](#) [Condensates](#) **[Statistics](#)** [Encyclopedia](#) [Help](#)

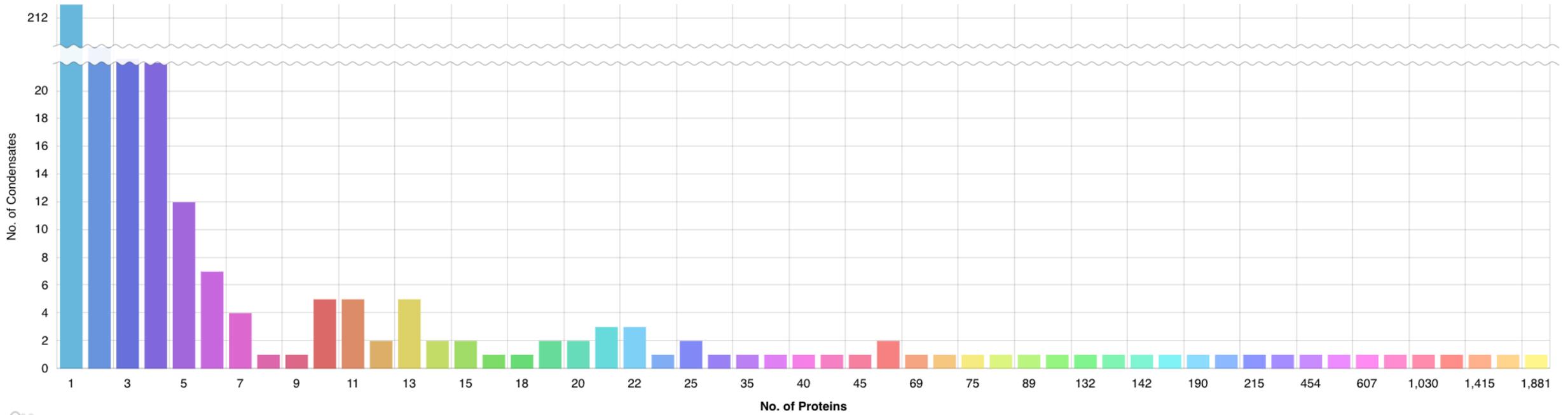
Protein distribution according to species.



Condensate distribution according to species.



Condensate proteome size distribution



Encyclopedia page

Details about biomolecular condensates are shown here.



Encyclopedia

Welcome to the CD-CODE Encyclopedia!

This is a glossary-like companion resource alongside the [CD-CODE](#)  database that aims to codify and standardise the concepts and nomenclature in condensate research; it can also be used as a standalone reference to anything related to the condensate research field.

The encyclopedia is a collaborative project so feel free to join!

Table of Contents

[Biomolecular condensates](#)

← A list of biomolecular condensates can be found here.

[Scaffolds \(Drivers\) and Clients](#)

[Mechanisms of Condensate Formation](#)

[Intrinsically Disordered Regions](#)

[Experimental Techniques](#)

[Valency](#)

[Synonyms](#)

[Phase Diagram](#)

[Condensate Markers](#)

You can edit this page in [CD-Code Wiki](#) ← Contributors can also edit the encyclopedia.

List of biomolecular condensates.

Biomolecular condensates

Biomolecular condensates are micron-scale compartments in eukaryotic cells that operate to concentrate proteins and nucleic acids but lack surrounding membranes (membraneless). RNA metabolism, ribosome biogenesis, the DNA damage response, and signal transmission are all affected by these condensates. Recent research has discovered that multivalent macromolecular interactions promote liquid-liquid phase separation, which is a fundamental organizing principle for biomolecular condensates ([28225081](#) , [31675499](#) ). The first membraneless compartment was observed within the nucleus of neuronal cells in the 1830s and was later termed the nucleolus ([21106648](#) ). Since then, many such compartments have been discovered in the nucleus, cytoplasm and on membranes of eukaryotic and prokaryotic cells.

[Nucleolus](#) ← Go to a biomolecular condensate of choice, eg. Nucleolus.

[Nuclear bodies](#)

[GW-body](#)

[miRISC condensate](#)

[Tis granule](#)

[RNP granules](#)

[T-cell signalosome](#)

[Sec-body](#)

[Paraspeckle](#)

[P-body](#)

[Stress granule](#)

[Centrosome](#)

[Cajal body](#)

[Nuclear speckle](#)

[Transcriptional condensate](#)

[Postsynaptic density](#)

[Synaptic vesicle pool condensate](#)

[Presynaptic active zone condensate](#)

[Neuronal RNP granule](#)

[Cytoplasmic protein granule](#)

[Inclusion body](#)

[P62 inclusion](#)

[Tau inclusion](#)

[Htt inclusion](#)

Help page

This page contains a guide on how to use CD-CODE and its functionalities.

[Home](#) [About](#) [Proteins](#) [Condensates](#) [Statistics](#) [Encyclopedia](#) [Help](#)

Users can Sign up to become contributors to be able to edit/add data on Cd-CODE.

[Home](#) [About](#) [Proteins](#) [Condensates](#) [Statistics](#) [Encyclopedia](#) [Help](#)

[Sign in](#) [Sign up](#)

Login to CD-Code

Email*

Insert your email here

Password*

Insert your password here

Sign in

[Forgot Password?](#)

Don't have an account yet? [Join CD-CODE as a contributor.](#)

Sign up here.

Sign in here.

Users can join to become contributor here.