Computational and Structural Biotechnology Journal 19 (2021) 3964–3977

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

# Insight into membraneless organelles and their associated proteins: Drivers, Clients and Regulators

Fernando Orti [a], Alvaro M. Navarro [a], Andres Rabinovich [a], Shoshana J. Wodak [b,c,1], Cristina Marino-Buslje [a,1,*]

[a] Bioinformatics Unit, Fundación Instituto Leloir. Avda. Patricias Argentinas 435, Buenos Aires B1405WE, Argentina
[b] VIB-VUB Center for Structural Biology, Flemish Institute for Biotechnology, Brussels, Belgium
[c] Structural Biology Brussels, Vrije Universiteit Brussel, Brussels, Belgium

ARTICLE INFO

*Article history:*
Received 14 March 2021
Received in revised form 26 June 2021
Accepted 27 June 2021
Available online 29 June 2021

*Keywords:*
Membraneless organelles
Liquid-liquid phase separation
Scaffold
Clients
Regulators
Liquid condensates
Protein disorder
Low-complexity regions

ABSTRACT

In recent years, attention has been devoted to proteins forming immiscible liquid phases within the liquid intracellular medium, commonly referred to as membraneless organelles (MLO). These organelles enable the spatiotemporal associations of cellular components that exchange dynamically with the cellular milieu.

The dysregulation of these liquid–liquid phase separation processes (LLPS) may cause various diseases including neurodegenerative pathologies and cancer, among others.

Until very recently, databases containing information on proteins forming MLOs, as well as tools and resources facilitating their analysis, were missing. This has recently changed with the publication of 4 databases that focus on different types of experiments, sets of proteins, inclusion criteria, and levels of annotation or curation.

In this study we integrate and analyze the information across these databases, complement their records, and produce a consolidated set of proteins that enables the investigation of the LLPS phenomenon. To gain insight into the features that characterize different types of MLOs and the roles of their associated proteins, they were grouped into categories: High Confidence MLO associated (including Drivers and reviewed proteins), Potential Clients and Regulators, according to their annotated functions. We show that none of the databases taken alone covers the data sufficiently to enable meaningful analysis, validating our integration effort as essential for gaining better understanding of phase separation and laying the foundations for the discovery of new proteins potentially involved in this important cellular process.

Lastly, we developed a server, enabling customized selections of different sets of proteins based on MLO location, database, disorder content, among other attributes (https://forti.shinyapps.io/mlos/).



## 1. Introduction

Eukaryotic cells organize their biological processes through numerous compartments or organelles which are often surrounded by a membrane. More recently, attention has been given to other types of supramolecular assemblies forming membraneless organelles [17,3,4,28]. Membraneless organelles (MLO) are condensates formed through liquid–liquid phase separation (LLPS), involving cellular components that form multivalent interactions and exchange dynamically with the intracellular medium in response to environmental signals [24,11,70]. MLOs carry out a wide range of functions enabled by heterogeneous mixtures of proteins and nucleic acids, such as metabolic processes and signaling pathways in the cytoplasm and in the nucleus [24,11,70].

Generally, MLOs contain tens to hundreds of macromolecules [16,31], but only a small subset of these components appears to be essential for the formation, structural integrity, and function of the condensate [92,36,45,22]. We refer to proteins that are essential for the formation of an MLO, and therefore directly or indirectly important for the function of the condensate, as High-Confidence (HC) MLO associated proteins. This category of proteins includes the small subset of proteins that have been assigned the

* Corresponding author.
*E-mail address:* cmb@leloir.org.ar (C. Marino-Buslje).
[1] Contributed equally

https://doi.org/10.1016/j.csbj.2021.06.042

role of Scaffolds/Drivers or Co-drivers of the LLPS process, on the basis of evidence that they undergo LLPS on their own (Scaffold/Drivers), or in conjunction with another protein or RNA (co-drivers) [69,27,38]. The remaining majority of the components are dispensable for MLO formation and often reside in the MLO only under certain conditions [23,34], although they may be important for the biological function of the MLO, and hence not be altogether dispensable for this reason. These proteins are referred to as Clients and Regulators. Clients are recruited to the condensate [10,25], whereas Regulators are often not part of the MLO proper [14,66].

The stability and dynamic properties of MLOs are the result of multivalent weak interactions between folded domains, intrinsically disordered proteins (IDP) or regions (IDRs), and interacting motifs [37].

LLPS is a tightly regulated process, which when perturbed, can undergo a transition from a physiological liquid condensate to pathological solid-like protein aggregates, leading to aging-associated diseases [6,82] and various neurodegenerative pathologies, among which Alzheimer, Parkinson, Huntington, ALS and FTD diseases stand out [2,18,68,100]. In this work we focus on the nine most representative MLOs comprising at least 15 associated proteins. Many other cellular condensates have been recently described in the literature but with fewer associated proteins [87]. For these representative MLOs, a very short description is given. The **Nucleolus** is the largest and better-studied condensate, it is implicated in rRNA transcription, processing and ribosomal subunit assembly, and its local protein content is twice as crowded as the nucleoplasm [78]. The **Nuclear speckles** are numerous irregular droplets, located adjacent to the interchromatin regions. They are enriched in long non-coding RNA in addition to specific proteins [44,21]. **Paraspeckles** are also located near the interchromatin regions and have an active role in gene expression regulation. The proteins NONO, PCPC1, and PSF are essential for the formation of these condensates, as is the long non-coding RNA NEAT1 [32,91]. **Cajal-Bodies** play an important role in the regulation of short non-coding RNAs, among which snRNAs of the spliceosome and snoRNA stand out. Coilin is one of the characteristic proteins of these bodies [59,57].

**PML-Bodies** contain various factors associated with a wide spectrum of functions, such as protein degradation, telomerase maintenance, and antiviral defense [53,20,61]. **The Nuclear Pore Complex (NPC)** is a large macromolecular assembly with a complex composition and diverse functions. Its main function is to regulate the macromolecule traffic to and from the nucleus and the cytoplasm, but it is also an active player in the gene expression pathway [85], chromatin organization and DNA repair [33].

The NPC is formed by a shell of well structured densely packed proteins , [95,13], but the natively unfolded phenylalanine – glycine rich (FG) segments of Nup proteins (FG-Nup), is what creates a liquid phase separated milieu at the center of the transport channel, a property that is deeply conserved [71,79,43]. The functional feature of the NPC – the permeability barrier – is made entirely of these FG repeats at the disordered region of the proteins, which readily phase separate. Importantly, macromolecules that are unable to interact with the disordered FG motifs of Nup are prevented from crossing the NPC [19,63]

Among the most relevant cytoplasmic MLOs are the **Stress Granules** and the processing bodies (**P-body**). **Stress Granules** are formed in response to various cellular stress signals, among their functions are inhibition of translation initiation and polysome disassembly. Stress Granules are formed by a group of RNA-binding proteins, binding to 80 rf-mRNA complexes that separate in a different liquid phase [48]. The **P-bodies** are a ubiquitous MLO conserved in eukaryotic cells, containing proteins involved in RNA turnover regulation. The accumulation of non-translating RNAs by translation inhibition is correlated with an increased number of p-bodies in the cell [58]. **Postsynaptic densities** (**PSD**) are membrane semi anchored, protein-enriched cellular compartments in close contact with the postsynaptic membranes. The SynGAP and PSD-95 proteins, in almost stoichiometric concentration form a multivalent complex that leads to the liquid–liquid phase separation in highly concentrated droplets in dynamic exchange with the free proteins in the cytoplasm [101].

Until very recently, databases with detailed and reliable information on LLPS proteins as well as tools and resources facilitating the study of their function and behaviour, were not available. Recently, four LLPS-dedicated databases partially addressing this need became available: **PhaSePro** [62], **PhasepDB** [98], **DrLLPS** [66] and **LLPSDB** [55]. Each of these databases archives a different set of proteins and focuses on different aspects such as the biophysical processes, LLPS assembly, MLO localization, protein distribution across MLOs and biological function. The experiments required to include a protein in the databases and the level of curation also differ considerably within and between databases.

The first goal of this work is to analyze the content of these databases despite their different structure, annotation and curation level. To this end we collated and complemented their records and produced a consolidated and curated set of proteins enabling a broader investigation of the LLPS phenomenon. The second goal is to identify features such as disorder content, low complexity regions, among others, in different groups of LLPS-associated proteins and correlate them with their function and distribution among MLOs. Lastly, we developed a web server (https://forti.shinyapps.io/mlos/) collating records from the 3 databases offering the scientific community easy means of retrieving sets of proteins filtered by specific criteria (localization to specific MLOs, source database, disorder content, etc.). We show that none of the currently available LLPS databases, taken alone, contains sufficient information to enable a general overview of the roles of LLPS associated proteins in the process of condensate formation, the various features of these proteins and the relations with their biological function and distribution among MLOs.

## 2. Material and methods

### 2.1. Databases

**PhaSepPro** [62]**:** is a manually curated database of LLPs driver proteins in various organisms, with emphasis on the biophysical properties that govern phase separation. It contains information on 121 proteins, and on 144 LLPS driving regions mapped onto these proteins and refers to 315 publications to support their protein classifications.

**PhaSepDB** [98]**:** classifies proteins into three groups depending on the origin of the data: **(1) Reviewed,** where each protein has proven LLPS association by at least one of the following assays: i) reconstitution of the liquid condensate with purified components *in vitro*, ii) *in vitro* or *in vivo* fluorescence recovery after photobleaching (FRAP), iii) droplets formation *in vivo* revealed by immunofluorescence with a fluorescent marker. **(2) Uniprot Reviewed,** proteins annotated in Uniprot as associated with membraneless organelles, subsequently verified using the same criteria as for the reviewed set and **(3) High-Throughput,** comprising proteins identified by High-throughput (HT) techniques applied to MLOs, such as Organelle purification, proximity labeling, Immunofluorescence image-based screen and affinity purification. As of October 2019, PhaSepDB includes 2957 proteins.

**DrLLPS** [66]**:** proteins in this database are classified as Drivers, Regulators and potential Clients. Drivers are defined as the drivers of LLPS, essential for the formation of the MLOs and the major com-

ponents which, alone or with other protein or nucleic acid, undergo LLPS. **Potential Clients** are defined as proteins identified to be in complex with known Scaffolds by conventional biochemical assays or mass spectrometry, or localized to an MLO by immunofluorescence, while **Regulators** are proteins that modulate the LLPS of Scaffold proteins and/or the stability and dynamic properties of the MLO. DrLLPS (latest update on 10 June 2019) contains information on 150 Scaffold proteins (LLPS drivers), 987 Regulators, and 8148 potential Client proteins. These proteins are assigned to 40 biomolecular condensates.

The fourth DB, **LLPSDB** [55], is entirely dedicated to *in vitro* LLPS experiments, most commonly using purified proteins and tightly controlling the biophysical conditions of droplet formation, such as different buffer solutions, crowding agents, protein concentration, temperature [15,29]. Although the database has valuable information about the proteins, these LLPS processes may not be representative of the phase separation process and MLO composition under physiological conditions. Also, the database does not provide information about MLO localization of the proteins. We nevertheless evaluated the overlap between the proteins stored in LLPSDB and those in the other three databases. This showed 75 of the 91 (82%) human proteins of LLPSDB to be redundant with those in the other 3 databases, of which 65 are LLPS Drivers. The 16 proteins unique to the LLPSDB lack annotations to specific MLOs, including in Uniprot, justifying their exclusion from our analysis. (for details, see **Fig. S1** and Table S1 of the Supplementary Material). For these reasons our analysis is limited to proteins in the **PhaSePro**, **PhaSepDB** and **DrLLPS** databases.

The contents of the 3 databases and their annotations were downloaded from the corresponding servers and their information was merged, completed and stored in a local database. For each protein we added information from Uniprot [84], domain information from PFAM [30], disorder content from MobiDB 3.0 [74] and post-translational modification information retrieved from PhosphoSitePlus [41]. Prion-Like domains (PLD) of human proteins were predicted by PLAAC using defaults parameters and a relative weighting of background frequencies ($\alpha$) of 50% from Homo Sapiens [5].

Data was downloaded, processed and plotted, using Python 3.7 scripts; graphs were prepared with Matplotlib and Seaborn Python libraries.

The consolidated dataset is freely available to the scientific community (https://forti.shinyapps.io/mlos/) allowing users to select proteins based on criteria such as disorder content (DC), Low-Complexity (LC) regions, database source among others. For example querying for proteins featuring LC regions, retrieves a list of 1200 proteins for further analysis. Proteins can be retrieved using a combination of criteria such as, being from humans, containing > 50% of disordered residues, having at least one RNA binding domain and a LC region of at least 100 residues. Applying these filters retrieves a group of 8 proteins (FUS, SFPQ, SRRM1, SRSF2, TAF15, RBM25, SLTM and SRSF8) from which the first 5 are of the Driver category.

### 2.2. Proteins are grouped into 3 categories based on the available annotations

For the purpose of the analysis the human proteins collated from the different databases are grouped into 3 categories on the basis of the annotations provided in the corresponding databases about their roles in the formation and integrity of the MLOs. Defining these roles relies heavily on the interpretation of LLPS experiments reported in the supporting literature. This is a challenging task for which agreed-upon practices are still lacking. Our division should therefore be considered more as work in progress, reflecting the current status of the field.

In the following we describe the criteria used to define the 3 protein groups. The so-called **High Confidence (HC) MLO associated** proteins include: 1) 101 proteins with direct supporting evidence of their driver and/or Co-driver role in the LLPS process, consolidated from 59 proteins contributed by PhaSePro [69,27], and 76 proteins by the DrLLPS database [66], 2) 490 reviewed proteins from PhaSepDB [98]. The latter proteins are not directly annotated as having direct experimental evidence supporting their role in driving LLPS, but are annotated in PhaSepDB as proteins whose role in MLO formation has been 'reviewed' or 'Uniport reviewed' by mining NCBI PubMed with appropriate keywords and by examining the full text of the publications for supporting evidence (detailed in section 2.1).

The **Potential Clients** group proteins with weaker supporting evidence on their participation in MLOs, since the information on their role is derived from high-throughput (HT) experiments. PhaSepDB contributes 1303 proteins detected in such experiments, and DrLLPS contributes 3176 proteins to this group. Both sets add up to 3209 non-redundant proteins, which for the sake of simplicity also be referred to as **Clients** in this study.

The **Potential Regulators** comprise proteins annotated as regulating the formation of the MLO or their stability, with no evidence for their incorporation into the MLO proper, while capable of changing the MLO morphology or function under certain circumstances. PhaSepDB and DrLLPS add up to 930 proteins to this category (also referred to as **Regulators** for simplicity sake). An overview of the consolidated datasets is presented in Fig. 1.

A protein may be annotated as a Driver in one MLO in one database, on the basis of its annotated scaffolding role, and as Client or Regulator in another, because of its HT recruitment. In such a case it is considered as a Scaffold for the purpose of the analysis as HT recruitment is a weaker evidence to assign a role. Also, as several experiments are often reported for each MLO type, a given protein can be assigned to two or three categories depending on the MLO. For example CPEB3 is reported to act as a Scaffold in P-bodies and a Client in Stress Granule. This is why there is an overlap between datasets shown in the Results section.

All the measured parameters are compared with a **control dataset,** which is the Human Proteome from Swiss-Prot (20.349 proteins).

## 3. Results

### 3.1. Databases overview

An integrative analysis of LLPS proteins and MLOs is challenging due to several reasons: the different sets and number of proteins and different focus and scope of individual databases, data structures, associated metadata, different experiments used to obtain the data and different curation level, among others. For example, PhaSePro contains a small number of proteins at a high curation level; DrLLPS, has different datasets depending on the role of the protein within the MLO, the supporting evidence and the curation level. PhaSepDB has less metadata and offers several data sets depending on the experimental method that are relied on to generate the annotation, this database emphasizes proteins localization more than their function in the LLPS process.

The organism coverage among the databases is also different: 49% of PhasePro, 83% of the low-throughput (LT) and 75% of the HT of PhaSepDB proteins are from humans, while in DrLLPS human Clients and Regulators proteins represent<40% of the proteins. Other organisms (non-human) are less well represented with no

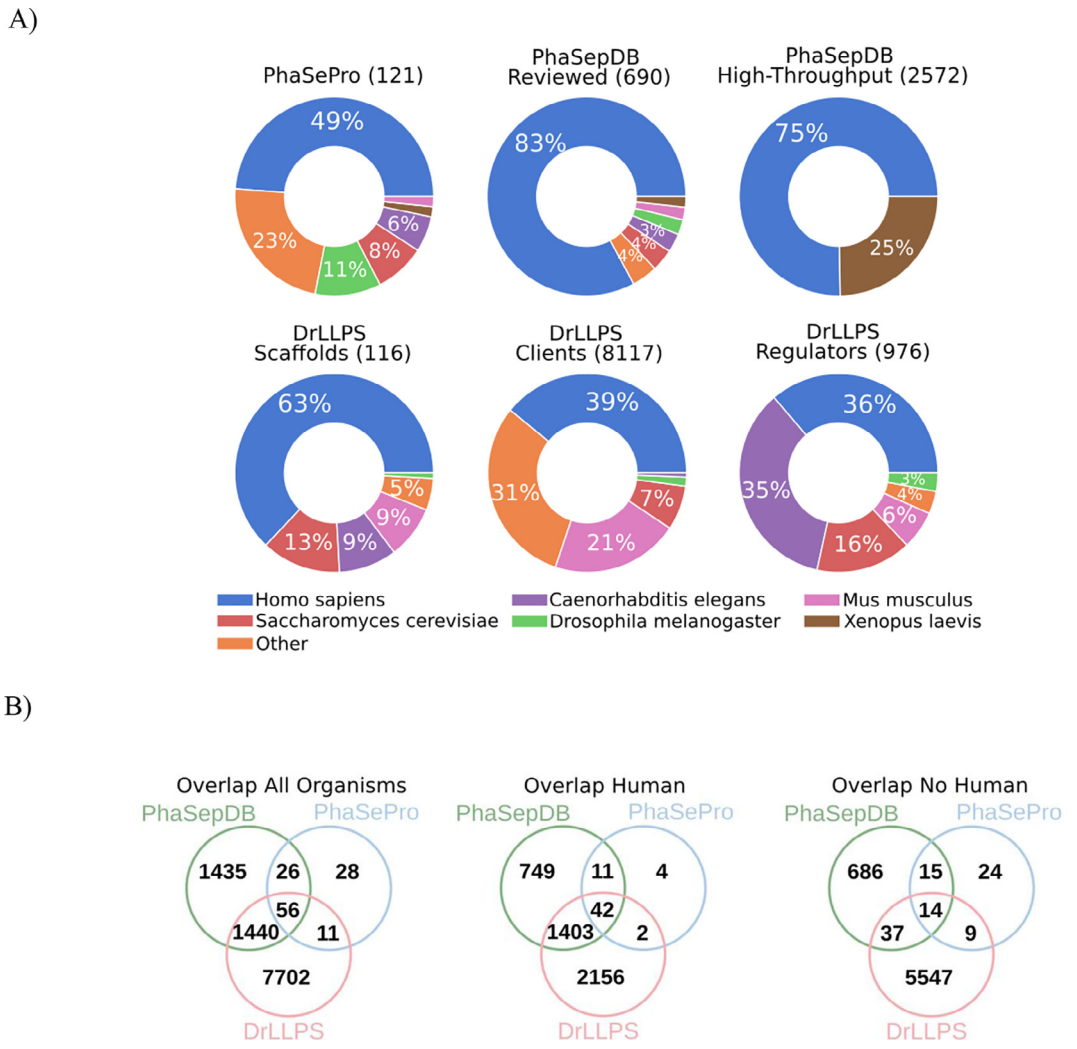


**Fig. 1.** General databases overview. A) distribution of entries by organisms in the databases. B) Venn diagram illustrating the overlap of the proteins stored in the different databases. Overlapped regions indicate the number of shared proteins; othewise, they are unique to a database.

single organism being represented in the datasets from all the databases **(**Fig. 1A **and** Suppl. Table 2**).**

Due to these differences, only 56 out of a total of 10,698 proteins are shared between all three databases: 42 human proteins out of 4367 and 14 non-human proteins out of 6331 (Fig. 1 **B**). This highlights the need to consolidate the data from the different databases in order to obtain a general overview and gain knowledge on the proteins engaged in LLPS and their functional role in MLOs.

### 3.2. Human proteins

Taking into account that human proteins make up the majority of the data, and that there is an incomplete and heterogeneous representation of the organisms in the databases, compounded by the intrinsic differences between organisms, the study of the MLOs across different organisms is not feasible at this point. We therefore focus our analysis on human proteins.

To overcome the inconveniences due to the different characteristics of the databases, we merge, unify and complete the information (whenever possible) and group the proteins into new augmented biologically relevant data sets of MLO associated proteins suitable for further analysis. These proteins are subdivided into three groups: ***High Confidence MLO associated***, **Clients** and **Regulators**, as described in Material and Methods section. Fig. 2 shows the composition of the datasets in these 3 groups and protein distribution within MLOs; the number of proteins that each database contributes to individual MLOs is listed in Suppl. Table 3**.**

Considering the large number of organelles described in the literature, of which many comprise only a few associated proteins, we focus on the 9 most representative MLOs, defined as those having at least 15 highly curated proteins in *HC MLO associated* proteins dataset. The remaining MLOs are clustered in this work under the label of "Others".

Our analysis first investigates the *HC MLO associated* set of proteins, as these are the best annotated and considered as essential constituents of the MLOs. This set of proteins is then compared with the Clients and Regulators sets. The following results are for the *HC MLO associated* set of proteins unless specified otherwise.

Few *HC MLO associated* proteins are present in all 3 databases, as expected, these proteins are well characterized as LLPS drivers or co-driver within the corresponding MLOs [69,27]. For example, this is the case of proteins like MAPT, HNRNPA1, TARDBP and FUS in paraspeckles [78], with FUS being also present in Stress Granules (SG) [24]. **Suppl.** Fig. 2 **and Suppl.** Fig. 3 displays the distribution of human proteins from the 3 different groups among the different MLOs, and the source databases.

To shed light on the composition and function of individual MLOs, in the following we analyze the shared proteins between

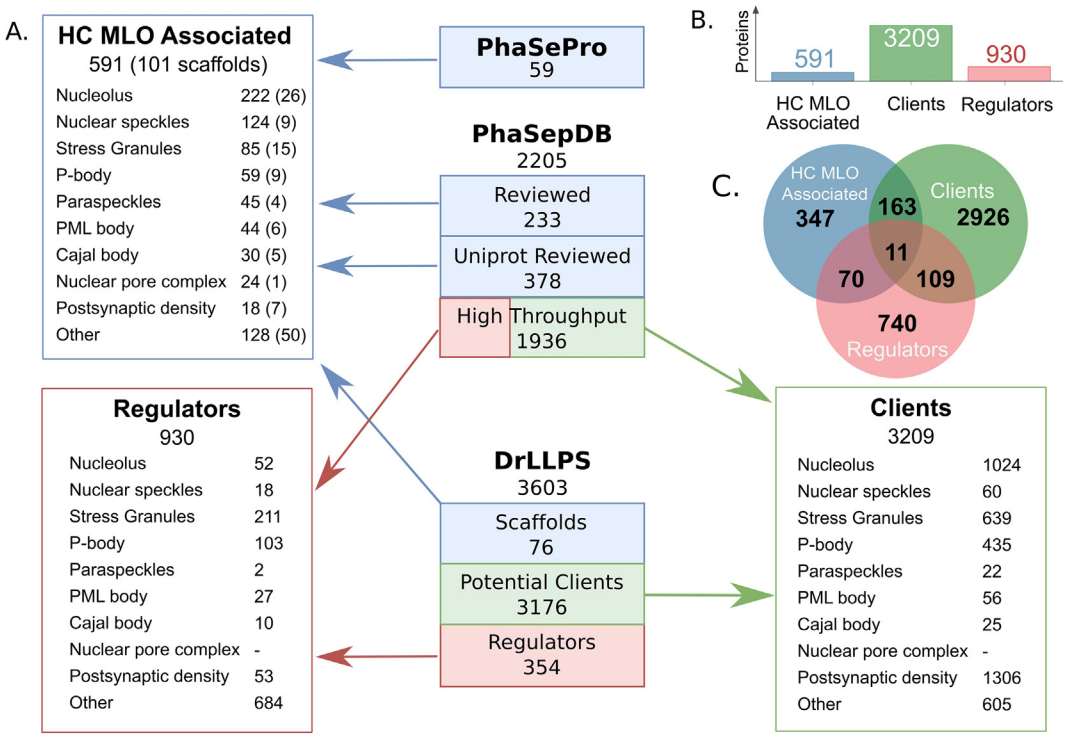


**Fig. 2.** A) Dataset compositions for the 3 groups of human proteins and the corresponding protein distribution among MLOs. The total number of proteins in each group and number of proteins within each MLO are listed. Numbers in brackets denote confirmed non-redundant Scaffold/Driver proteins. B) Total number of *HC MLO associated*, Potential Client and Regulator proteins respectively, in our consolidated dataset. C) Overlap of the proteins from the three groups. Note that some proteins play a diving/scaffolding role in one MLO while acting as Client in another MLO.

MLOs, domain enrichment of proteins in MLOs, their disorder content, content of low complexity regions and post-translational modifications (PTM).

### 3.3. Distribution of HC MLO associated, Clients and Regulator proteins in individual MLOs

Proteins common to two MLOs range from 20 to 60% except for the NPC that has only one protein (NUP98) in common with SG (1 out of 24 proteins). Half of the proteins present in P-bodies are also in SG, whereas SG shares 30% of its proteins with P-body. SG proteins are also present in many other MLOs as: Nucleolus, Nuclear Speckles and Paraspeckles (**Suppl.** Fig. 4).

Examples of proteins that are present in several MLOs, include FUS, EWSR1, HNRNPA1 and XPO1 present in almost 6 MLOs and DDX1, DDX3X, HABP4, HNRNPA2B1, KPNB1, TAF1 and TARDBP in almost 5 (Fig. 3 **A**). Not surprisingly, most of these proteins are drivers or co-drivers and are also present in all three databases.

The studied MLOs do not have the same proportion of proteins in each of the 3 protein groups, nor the same level of protein annotations. >60% of the proteins in Paraspeckles and Nuclear Speckles are *HC MLO associated* (of which several are drivers or co-drivers), whereas proteins of this set represent <10% of the proteins in P-bodies and SG. An extreme case is PSD with 95% of its proteins being Clients. In fact, all the proteins of PSD were detected in one experiment [12] and are annotated as Clients in DrLLPS (Fig. 3 **B**). On the other hand, NPC has exclusively *HC MLO associated* proteins, one of them (NUP98) is an LLPS Driver (see Discussion section).

### 3.4. Disorder content

The role of disordered proteins and regions in driving LLPS has been amply documented for driver proteins in the nucleolus [15], SG [93], Paraspeckles [39] and P-bodies [26], among others [78].

The amino acid composition of the disordered regions of our set of *HC MLO associated* proteins agrees with previous reports [76]. Cysteine, Isoleucine, Phenylalanine, Leucine and Valine are under-represented, while Glutamine, Arginine, Serine, Glycine and Proline are over-represented. However, discriminating by MLO allows us to observe sizable differences that may have a biological meaning and deserve further investigation (**Suppl.** Fig. 5 **A**).

It was shown that LLPS proteins have on average a higher disorder content (DC) than other proteins [78]. In agreement with this observation, our *HC MLO associated* set of proteins has a significantly higher disorder content (DC) than the Human proteome from Swiss-Prot, taken as reference, yet *HC MLO associated* proteins still exhibit a wide range of disorder content across MLOs (Fig. 4 **A**).

Proteins in the Nucleolus, Nuclear Speckles, SG, Paraspeckles and P-bodies are significantly more disordered than human proteins in Swiss-Prot. The median DC of PML-bodies, Cajal-bodies and PSD proteins is higher than those of the Swiss-Prot reference, albeit not significantly.

On the other hand, the median DC of NPC proteins is below that of the human proteome on the basis of currently available data in the LLPS protein databases: 1 driver protein from PhaSePro - Nup98- and 23 proteins from PhaSepDB. This result reflects the fact that the NPC, taken as a whole, has well structured densely

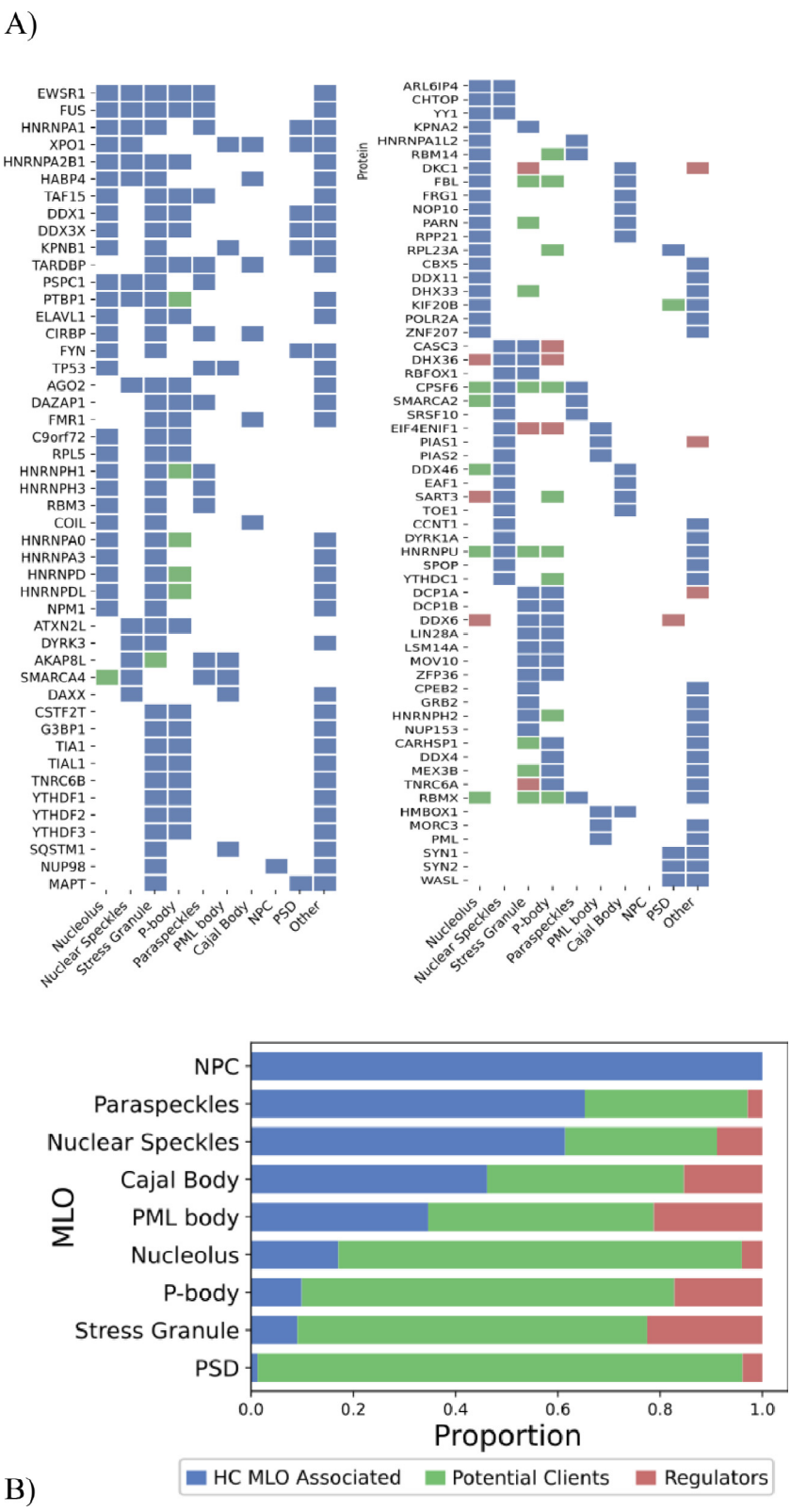


**Fig. 3.** Landscape of Protein distribution among MLOs. A) *HC MLO associated* proteins present in more than one MLO are colored blue. When in addition a protein is also annotated as a Client or Regulator in another MLO, it is colored green and red respectively (Y axis: proteins; X axis: MLOs. B) Proportion of proteins of a given category in different MLOs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

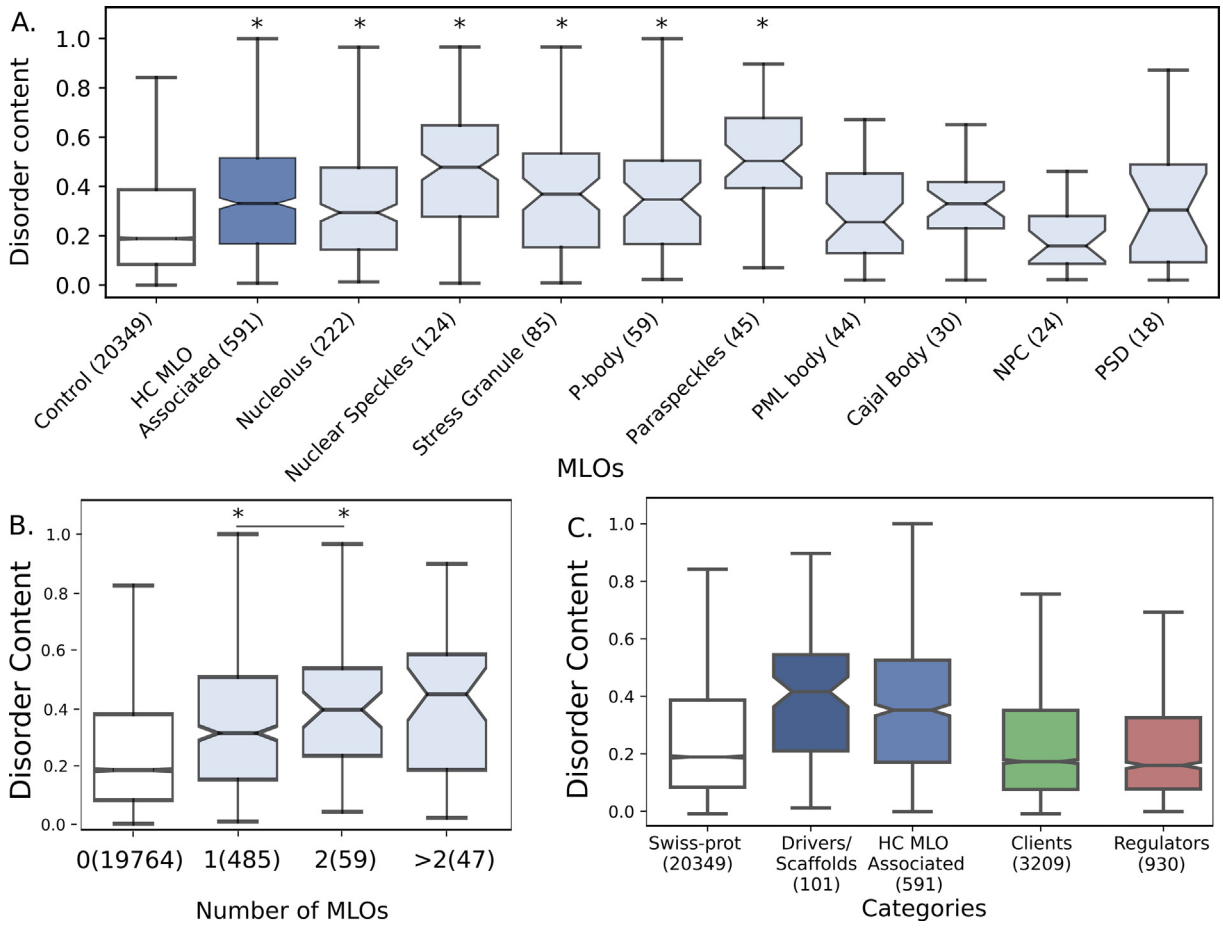


**Fig. 4.** Disorder content. A) Disorder content (DC) distribution by MLO. (*) indicates that there is a significant difference (Kruskal-Wallis test, p-value < 0.05) with respect to the control set of proteins (swiss-prot); B) DC distribution vs the number of MLO where the protein is present. The DC in the 3 ranges (1, 2 and > 2 MLOs) are significantly different from the control. The DC being in 1 and 2 MLOs is also significantly different (Kruskal-Wallis test, p-value < 0.05, denoted by a *); C) Disorder content by protein sets (drivers are individualized for comparison purposes).

packed nucleoporins organized in layers of coat and adaptor proteins [51,40].

The DC of Clients and Regulators in individual MLOs is lower than the DC of the *HC MLO associated* proteins except for PML bodies and cajal-Bodies where Regulators have the highest DC (shown in **Suppl.** Fig. 6**).**

When classifying our set of proteins by High (DC >70%), Middle (30> DC <70%) and Low (DC <30%) disorder content, the proportions are different between MLOs (**Suppl.** Fig. 7**)**. Except for PML-body and NPC, all the organelles have >50% of their proteins with middle and high DC. Extreme DC values are found for Paraspeckles and Nuclear Speckles with almost 80% of their proteins having high DC and for NPC, which exhibits the lowest number of proteins with medium or high DC (around 20%).

An interesting result is that the more different MLOs a protein takes part in, the higher the DC content of that protein, i.e more disordered proteins are found in more MLOs. This might suggest a predominant and similar role of these proteins in the different MLOs (Fig. 4 **B**) (See also suppl. Table 4**,** on proteins with >70% of disordered residues and their location).

Lastly, the DC is also different across the *HC MLO associated*, Clients and Regulators sets of proteins. The DC of *HC MLO associated* proteins differs significantly from that of the control human proteome, while those of Clients and Regulators do not (Fig. 4 **C.** The DC of the Driver set of proteins alone is also shown for comparison).

### 3.5. Low complexity regions

A special case of disordered regions are low complexity regions (LC), where the amino acid composition is biased toward a handful of amino acids, sometimes to a single one or a few. There are 579 LC regions belonging to 261 proteins in the *HC MLO associated* group, 83% of them overlap with the disordered regions. (Suppl. Table 5 contains all the LC regions of the entire dataset).

A striking example of a composition bias in LC regions, is the protein FMN2_HUMAN (formin 2), which has 339P in 29 identical repetitions of a $(P)_4$LPGAGI motif (in a 511 amino acid long P rich region). Formin 2 controls the cyclin-dependent kinase inhibitor p21 protein levels by forming a complex and preventing its degradation [96]. Formin 2 acts as an actin nucleation factor and promotes assembly of actin filaments together with SPIRE1 and SPIRE2. It also plays an active role in responses to DNA damage, cellular stress and hypoxia and in the vesicle transport along actin fibers [89,72]. Formin 2 is annotated as “reviewed” in PhaSepDB, and hence is a *HC MLO associated* protein in the nucleolus, and is classified as a potential Client in Postsynaptic Density in the DrLLPS.

Glycine-rich regions are the most over-represented in our set of *HC MLO associated* proteins, followed by Proline rich regions (Fig. 5 **A**). Arginine, Serine and Aspartic acid are also enriched in LC regions. Fig. 5 **A** shows that different MLO have specific aminoacid-rich regions. For example Glutamine (Q) and Glutamic

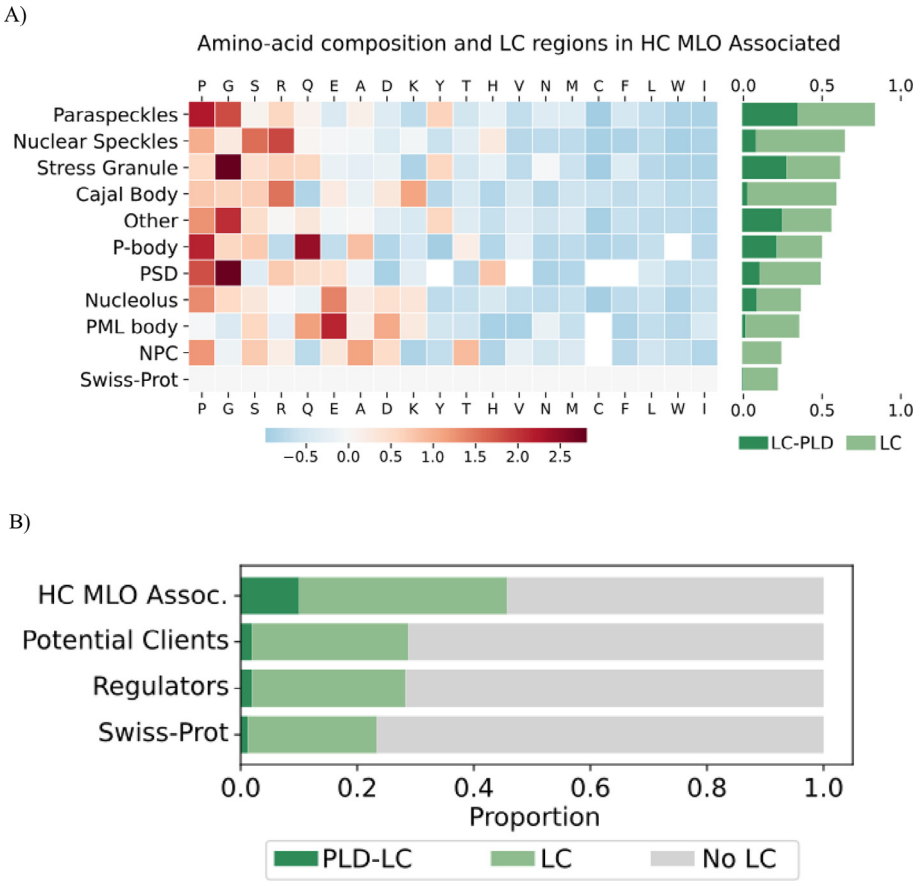


**Fig. 5.** A) Amino acid enrichment in LC regions by MLO in the High Confidence MLO associated dataset. From blue (depleted compared with Swiss-Prot), to red (enriched compared with Swiss-Prot). Enrichment was calculated as: DeltaFreq = (FreqAA in LC – FreqAA control)/FreqAA control. Green histogram: proportion of proteins having LC regions (from 0 to 1) by MLO. Dark green denotes the PLD domain proportion within the LC regions while light green shows other LC regions. B) Proportion of protein containing LC regions in the three categories of proteins (*HC MLO Associated*, Clients and Regulators). Dark green are proteins containing PLD, RGG or SR motifs (featured LC); light green, other LC regions; gray indicates non-LC regions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

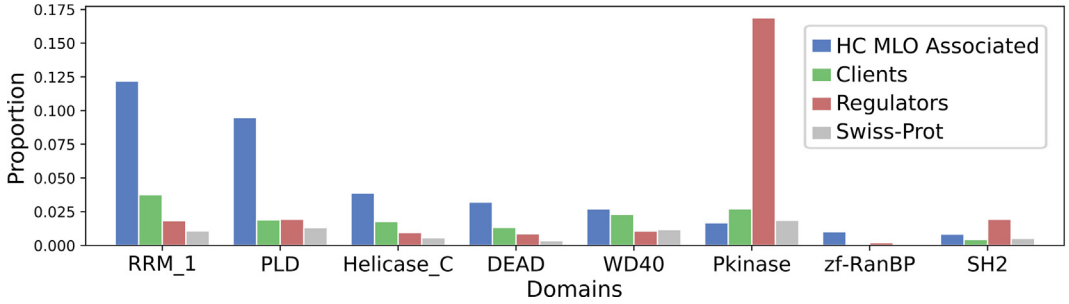


**Fig. 6.** Comparison of the most highly enriched domains in the different set of proteins. The proportion of proteins having RMM_1: RNA binding domain; PLD: Prion-Like Domain; Helicase_C; DEAD domain, WD40 domain, P-Kinase, zf-RanBP and SH2 domains in the three groups of proteins compared with the control (Swiss-Prot).

(E) are not enriched in LC regions when considering the full set of *HC MLO Associated* proteins (**Suppl. Fig. 8**), but are remarkably enriched in P-body and PML-body respectively. Also, the proportion of proteins with LC region varies in the different MLOs (Fig. 5 **A**). LC regions are higher in *HC MLO associated* than in Potential Clients and Regulators proteins (Fig. 5 **B**).

Although Fig. 5 shows the enrichment for individual amino acids, important LC regions feature frequent combinations of residues: glycine and Arginine (GR), Arginine and Serine (RS) (remarkably high in Nuclear Speckles), and Glycine Serine (GS). Serine in combinations with Aspartic (DS) and Glutamic (ES) are also very frequent, mainly in Nucleolus. Within the LC regions, PLD, RGG and RS motifs deserve a separate analysis as they make up the majority of our annotations and are strongly associated with LLPS and MLOs (Table 1 shows the proteins with PLD, RS-rich and P-rich regions and the MLO where they are present). They are mostly RNA binding and protein-protein interactions motifs [86,83,90].

Glycine rich regions are frequently associated with LLPS, because they are part of the so-called Prion-Like domains (PLD). **PLD domains** in the FET family of proteins (FUS, EWSR1 and TAF15), responsible for driving phase separation *in vivo* and *in vitro* [86,83,90], are enriched in G, Y, Q and N residues.

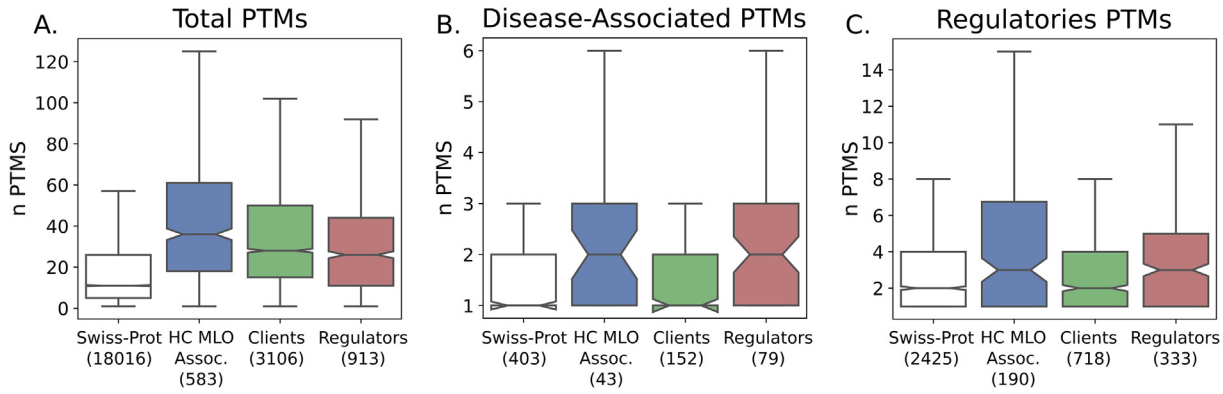


**Fig. 7.** A) Boxplot of the PTMs annotated to each protein by groups (between brackets are the number of proteins that have at least 1 PTM in the set); the PTMs annotated to proteins in Swiss-Prot are used as reference. B) Disease-associated PTMs, by groups. C) PMTs associated with regulatory functions, by groups.

**Table 1**
High-confidence MLO associated proteins with LC regions and their localization across the MLOs.

| MLO | Prion-Like Domain | TriRGG regions | RS-rich regions | Proline-Rich regions (PRM) |
|---|---|---|---|---|
| Nucleolus | HNRNPD, RBM14, TAF15, PSPC1, HNRNPA2B1, HNRNPH1, HNRNPH3, FUS, EWSR1, POLR2A, HNRNPA1L2, HNRNPA1, ILF3, DDX17, HNRNPA3, DROSHA, HNRNPDL, CHD7, ZNF207, HNRNPA0 | TAF15, FUS, EWSR1, FBL, HNRNPA1L2, HNRNPA1, HABP4, DDX17, INO80B, HNRNPA3, CHTOP, HNRNPA0 | ARL6IP4 | CEBPA, CHD7, DDX17, DROSHA, EWSR1, FMN2, KDM6A, NOL3, PELP1, PHF2, PHLDA1, POLR2A, SIN3A, TP53, ZNF207 |
| Nuclear Speckles | PSPC1, HNRNPA2B1, FUS, EWSR1, HNRNPA1, MAML2, KMT2E, HNRNPU, ATXN2L, MAML1, MAML3 | FUS, EWSR1, HNRNPA1, HABP4, ALYREF, CHTOP | NCBP1, SRSF10, SRSF2, SON, U2AF2, SRRM1, AKAP17A, ARL6IP4, SRRM2, SRSF1, SRSF6, SRSF3, PNISR, PPIG, CDK13, DDX46, ZNF638, SRSF4, CCNL2, BCLAF1, THRAP3, CCNL1 | AKAP8L, ATOH8, ATXN2L, CASC3, CBLL1, CBX4, CCNT1, CDK13, CEBPB, CPSF6, EAF1, EWSR1, HNRNPU, KMT2E, MAML1, MAML2, PABPN1, PNISR, RBM15, RBM15B, RING1, RREB1, SETD1A, SETD1B, SF3B2, SMARCA2, SMARCA4, SON, SRRM1, SRRM2, WT1 |
| Stress Granule | HNRNPD, TAF15, PSPC1, HNRNPA2B1, HNRNPH1, HNRNPH3, DAZAP1, FUS, EWSR1, ATXN2, TARDBP, YTHDF2, HNRNPA1, UBQLN2, YTHDF1, ATXN2L, CSTF2T, NUP153, HNRNPA3, HNRNPDL, HNRNPH2, HNRNPA0 | TAF15, FUS, EWSR1, HNRNPA1, LSM14A, HABP4, HNRNPA3, HNRNPA0 | – | ATXN2, ATXN2L, CASC3, CPEB2, CSTF2T, DAZAP1, DVL2, EIF4G1, EWSR1, G3BP1, PUM1, RC3H2, TNRC6B, YTHDF1, YTHDF2, YTHDF3, ZFP36 |
| P-Body | TAF15, HNRNPA2B1, DAZAP1, FUS, EWSR1, TARDBP, YTHDF2, YTHDF1, ATXN2L, TNRC6A, CSTF2T | TAF15, FUS, EWSR1, LSM14A | – | ATXN2L, CPEB3, CSTF2T, DAZAP1, EIF4E, EWSR1, G3BP1, MEX3A, PAN3, TNRC6A, TNRC6B, YTHDF1, YTHDF2, YTHDF3, ZFP36 |
| Paraspeckles | RBM14, TAF15, HNRNPUL1, PSPC1, SFPQ, HNRNPH1, HNRNPH3, DAZAP1, FUS, FAM98A, EWSR1, SS18L1, SOX9, HNRNPA1L2, TARDBP, HNRNPA1 | TAF15, HNRNPUL1, RBMX, SFPQ, FUS, HNRNPR, FAM98A, EWSR1, HNRNPA1L2, UBAP2L, HNRNPA1 | CPSF7, SRSF10, RBMX | AHDC1, AKAP8L, CPSF6, CPSF7, DAZAP1, EWSR1, FIGN, HNRNPK, HNRNPR, HNRNPUL1, MEX3C, NONO, PCED1A, RBM12, RBMX, SFPQ, SMARCA2, SMARCA4, SOX9, TP53, UBAP2L, ZNF335 |
| PML-Body | ATRX | – | – | AKAP8L, RPA1, SATB1, SMARCA4, THAP1, TP53 |
| Caja- Body | TARDBP | FBL, HABP4 | DDX46 | EAF1, ICE1, SMN1 |
| PSD | SYN1, HNRNPA1 | HNRNPA1 | – | SHANK3, SYN1, SYN2, SYNGAP1, WAS |
| Other | BRD4, HNRNPD, SYN1, TAF15, HNRNPA2B1, DAZAP1, FUS, EWSR1, POLR2A, SP1, TARDBP, YTHDF2, HNRNPA1, YTHDF1, , HNRNPU, TNRC6A, CSTF2T, NUP153, HNRNPAB, HNRNPA3, HNRNPDL, HNRNPH2, ZNF207, HNRNPA0 | TAF15, RBMX, KHDRBS1, FUS, EWSR1, HNRNPA1, HABP4, HNRNPA3, HNRNPA0 | RBMX | BRD4, CBX2, CCNT1, CPEB2, CSTF2T, DAZAP1, EWSR1, G3BP1, HNRNPU, HTT, KHDRBS1, POLR2A, RBMX, SOS1, SYN1, SYN2, TNRC6A, TNRC6B, TP53, WASL, YTHDF1, YTHDF2, YTHDF3, ZNF207 |

Important types of LC regions in Drivers proteins: PLD, RGG and RS. Gene name of the proteins and MLO where they are present. barrier

The dataset of *HC MLO associated* proteins has 56 members with PLDs distributed among different MLOs (**Suppl. Fig. 9**). Twenty two out of the 56 proteins with PLD domains are located in more than one MLO, with examples such as FUS and EWSR1, which are present in 5 MLOs.

A particularly important type of G-rich region to drive transition events is the one characterized by having two or three neighboring **RGG motifs** separated by 0–4 residues (RGG(X0–4)RGG(X0–4)RGG) [29,83]. In our *HC MLO associated* set of proteins, 21 have an RGG motif (two or three RGG repetitions). Proteins with RGG repetitive motifs are associated with a large number of functions related to RNA processes, such as transcription modulation, splicing, export and translation.

Another important motif in our dataset is the **proline-rich motif** (PRM) that upon interaction with SH3 domains promotes droplet formation. We found 97 proteins with P-rich regions

within the *HC MLO associated*, 287 in the Potential Clients and 91 in the Regulators sets of proteins.

Nuclear speckles have an enrichment in Serine and Arginine proteins (SR-proteins) that also play various roles in RNA processing, such as splicing activation and repression, mRNA export and translation. The SR family of proteins have SR rich regions and RNA binding domains that allows it to interact with RNA [86]. In humans, the SR protein family is encoded by nine genes, named splicing factors arginine serine-rich (SFRS) [99]. The *HC MLO associated* set of proteins has the SFRS 1,2,3,4,6 and 10 in nuclear speckles, SFRS10 is also present in paraspeckles. (Table 1).

### 3.6. Mlos domain enrichment

In addition to disordered and LC regions, we also mapped the Pfam domains to every protein in our dataset to analyze domain composition and distribution by MLO. Prion-like domains (PLD), as such, are not annotated as Pfam domains, so these domains were predicted using PLAAC software [5].

First, we analyzed domain enrichment in Driver proteins vs Swiss-Prot (exact fisher test, p-value < 0.05). Eighty two domains showed enrichment from which we selected the following 14 domains present in at least 5 proteins: RNA binding domain (RMM_1, PF00076), PLD, Helicase_C (PF00271), DEAD domain (PF00270), WD40 (PF00400), LSM Domain (LSM, PF01423), Zn-finger (zf-RanBP, PF00641), HMG_box (PF00505), SAP (PF02037), FHA (PF00498), MMR_HSR1 (PF01926), zf-CCCH (PF00642) and zf-CCHC (PF00098).

RNA binding and PLD domains remarkably predominate in *HC MLO associated* proteins, followed by helicase-3, dead and W-40 domains. RRM_1 is the largest group of eukaryotic RNA-binding proteins and is associated with a variety of RNA related functions. Several studies point out that in the FET family of proteins the interaction of RRM_1 and PLD domains regulates liquid-liquid separation [90]. Such is the case of FUS, where multivalent interactions among tyrosines from prion-like domains and Arginines from RNA-binding domains drives phase separation [90]. FUS phase behavior is also regulated in part by the phosphorylation of serine residues in the PLD domain that control physiological assembly preventing pathological aggregation [64]. Twenty seven proteins of the *HC MLO associated* set contain both domains (PLD plus RRM_1): CELF4, CSTF2, CSTF2T, DAZ1, DAZ3, DAZ4, DAZAP1, EWSR1, FUS, HNRNPA0, HNRNPA1, HNRNPA1L2, HNRNPA2B1, HNRNPA3, HNRNPAB, HNRNPD, HNRNPDL, HNRNPH1, HNRNPH2, HNRNPH3, PSPC1, RBM14, SFPQ, TAF15, TARDBP, TIA1, TIAL1.

RRM_1 and PLD are the more enriched domains in the Nucleolus, Nuclear Speckles, SG, P-Body and Paraspeckles. On the other hand, PML-Body and NPC are the only MLOs lacking RRM_1 domains, whereas NPC is the only MLO with no PLD nor RRM_1 domains, although these types of domains are present in NPC's of other species [1,35].

Helicase_C and DEAD domains are present in the DEAD Box family of proteins. DEAD box helicases are involved in various aspects of RNA metabolism, including nuclear transcription, pre mRNA splicing, ribosome biogenesis, nucleocytoplasmic transport, translation, RNA decay and organellar gene expression [24].

Helicase_C and DEAD domains are present in: ASCC3, BLM, DDX1, DDX17, DDX18, DDX21, DDX27, DDX31, DDX3X, DDX4, DDX46, DDX47, DDX5, DDX6, DHX33, DHX36, EIF4A3, MTREX, WRN. These proteins are associated with several functions including ribosomal biogenesis (ej DDX27), mRNA processing (ej DDX3X, DDX46) and translation repression (ej DDX6) [47].

WD40 domains are formed by a variable number of WD40 repeats. WD-repeats proteins are a large family found in all eukaryotes, implicated in signal transduction and transcription regulation, cell cycle control and apoptosis. Proteins containing WD40 repeats serve as platforms that coordinate multi-protein complex assemblies, where the repeating units serve as a rigid Scaffold for protein interactions and are mediators of transient interplay among other proteins [46,80]. An example are the WD40-repeat WD40 containing proteins in the Drivers dataset: are AAAS, DCAF13, FBXW7, GEMIN5, PLRG1, PWP1, RACK1, RBBP4, SEC13, UTP4, WDR12, WDR18, WDR36, WDR43, WDR74, WRAP53.

zf-RanBP (RanBP2-type Zinc Fingers) are zinc finger RNA-Binding motifs found in the E3 SUMO-protein ligase RanBP2 (RANBP2, also called Nucleoporin Nup358) and other proteins. RANBP2, located at the tip of the cytoplasmic fibrils of the NPC has a repetition of 8 zinc finger domains [97]. This nucleoporin binds single-stranded RNA and DNA [94], is a component of the nuclear export pathway [88] and in the sumoylation pathway [73] among other functions. RanBP2-type zinc fingers recognize single-stranded RNA (ssRNA) in a sequence-specific manner which is an important aspect of gene regulation that modulates mRNA processing [65]. Other proteins having zf-RanBP domains in our set of *HC MLO associated* proteins are FUS, EWSR1, NUP153 and TAF15 in Stress Granule, YAF15 in Nucleolus and MDM2 in PML-Body. Among the Potential Client proteins, RBM10 in P-Bodies, and YAF2 in the Nucleolus both contain zf-RanBP domains, and so does the Regulator protein RANDBP in P-Bodies.

PML-Body, Cajal-Body, NPC and PSD proteins have a more diverse set of domains. But since fewer proteins are currently associated with these MLOs, this observation needs to be verified as more data become available.

Taken together, the above observations show that the *HC MLO associated* set of proteins are strongly enriched in RNA binding proteins that are involved in the many RNA related processes across all the analyzed nuclear or cytoplasmic MLOs.

Next, we compare the proportion of the top 6 enriched domains in *HC MLO associated* proteins to those in the other 2 protein groups (Fig. 6). We find that RRM1, PLD, Helicase_C, DEAD, DW40 and zf-RanBP domains are enriched, in decreasing proportion, in the *HC MLO associated* set of proteins, WD40 and RMM_1 domains are also enriched in Potential Client proteins, although their enrichment is lower than in *HC MLO associated*. We studied the proportion of two other domains due to their enrichment in Potential Client and Regulator proteins: P_Kinase and SH2. The Kinase domain is remarkably enriched in Regulator proteins, agreeing with the regulatory function of the protein Kinases, also highlighting the importance of the PTMs (in this case phosphorylation) in the MLO regulation. The proportion of Kinase domains is likewise higher in Potential Clients than in *HC MLO associated* (Fig. 6).

### 3.7. Post-translational modifications

Previous studies have shown that certain Post-translational modifications (PTMs) regulate the structure and function of the MLOs [52,64,75,77,81].

For example, phosphorylation in low complexity regions of FUS decreases LLPS, preventing aggregation and lowering its toxicity [64]. On the other hand, TAU phosphorylation increases condensate formation [7], while acetylation reduces it [77]. Furthermore, some Client proteins are recruited or excluded from certain MLOs by means of PTMs [25].

We mapped the most frequent post-translational modifications (PTMs) to the proteins in our dataset: phosphorylation, methylation, acetylation, ubiquitination sumoylation and glycosylation. PTMs information was taken from the PhosphoSitePlus database [42], which stores PTM sites supported by low- and high-throughput data sources, manually curated by experts from the literature. **Suppl. Fig. 10** shows the number (A) and the most represented PTMs among MLOs (B) and the function altered by their Regulatory activity (C). We found that the *HC MLO associated*,

Potential Clients and Regulators set of LLPS-associated proteins taken together have more PTMs per protein than the control set (Swiss-Prot). *HC MLO associated* have more PTMs than Potential Clients and Regulators (Fig. 7 **A**), and certain MLO's such as PML-Bodies, Paraspeckles, SG and NPC have on average more modification sites than the *HC MLO associated's* median (**Suppl. Fig. 10 A**).

Several diseases are linked to PTMs in LLPS proteins. Suppl. Table 6 shows the proteins and the PTMs associated with a disease pathology. For example MAPT PTMs are linked with Alzheimer's disease (AD), chronic traumatic encephalopathy (CTE), Pick's disease (PiD), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), globular glial tauopathy (GGT), and argyrophilic grain disease (AGD), diabetes among others [56,50].

MYC PTMs are associated with acute lymphocytic leukemia, chronic myelogenous leukemia, leukemia, Burkitt's lymphoma, AIDS-related lymphoma, B cell lymphoma, ovarian cancer, breast cancer, hepatocellular carcinoma [60,49,9]. On the other hand, FUS phosphorylation in the LC domain decreases phase separation, preventing the aggregation and reducing its toxicity [64].

*HC MLO associated* and Regulator proteins have more PTMs related to diseases than the human proteome and the Client proteins (Fig. 7 **B**). Regulatory PTMs show a similar trend (Fig. 7 **C**). These two observations agree with the driving role and the regulatory function of proteins in these groups.

## 4. Discussion

Current LLPS-dedicated databases differ in their focus, annotation and curation levels among other characteristics, making analyses difficult and the extraction of new knowledge challenging.

The goal of our work has been to derive valuable insights into the landscape of membraneless organelles, their associated proteins, and their functional roles in cell biology. An important contribution of this work was to integrate the available data across databases, organize the information on LLPS proteins, and enrich it with complementary functional annotations. The integration was necessary, because the MLO and protein coverage of individual databases is different and incomplete, resulting in a limited amount of shared data across databases. The three databases analyzed here share only 42 of the 591 *HC MLO associated* proteins, the Potential Client and Regulator proteins also have an uneven distribution in the different databases. In addition to conducting the analysis, we considered it useful to make this integrated and enriched database available to the scientific community. To this end, we developed a web server enabling easy retrieval of sets of proteins necessary for a particular analysis without need to look for the information in each database. By applying filters on disorder content, LC regions, domain content, databases source, among others, the necessary protein sets can be assembled and all the associated information retrieved.

The proteins consolidated from the 3 databases were grouped into three sets: *HC MLO associated*, Client and Regulator proteins, on the basis of their level of annotation and their role in the LLP, to help obtain a general picture of the characteristics of the MLOs and improve our understanding of these fundamental ultrastructures and their roles in cellular function.

It is well established that IDR-containing proteins contribute significantly to the liquid-like character of MLO's by forming transient, low affinity multivalent interactions, enabling protein nucleation, assembly and disassembly of the corresponding organelles [67,78]. Our analysis confirms these observations. We find that on average, the disorder content (DC) of human proteins in our *HC MLO associated* set, is larger than in the human proteome of Swiss-Prot, but differs greatly between MLOs. Also, several authors observe specific biases in domain and sequence composition of proteins that undergo phase separation [54]. Our results confirm these biases, but also show that domain composition depends on the MLO and the functional role of the protein within the MLO. *The HC MLO associated* proteins tend to feature multiple functional modules, including folded domains as well as disordered and LC regions, thereby creating multivalency, a hall mark of property for LLPS formation. As confirmed here, this category of proteins is enriched in RNA- and to a lesser extent in DNA- binding domains and motifs such as, RRM_1, DEAD, Helicase_C. This highlights the fundamental role of RNA in MLO formation, stabilization and function, but very little is currently known about the RNA molecular composition, structures, and the mechanistic underpinning of its functional roles. Elucidating the roles of nucleic acids in LLPS processes, MLOs dynamics and regulation, is clearly an important direction for future research.

Post-translational modifications are also prevalent in both *HC MLO associated* and Regulator proteins in agreement with their important role in MLOs fate and the fine tuning of the underlying regulatory processes, whose malfunction may cause disease.

We find that the disorder content, low complexity regions, as well as motifs are not predominant across all MLOs, suggesting that other mechanisms, not involving these features may underlie the formation, function and stability of certain types of MLOs. As already mentioned, the NPC is a good example of such MLOs.

Another example is the PSD, whose assembly is driven by the formation of the SynGAP/PSD-95 complex, which does not involve protein-interactions mediated by IDRs [101].

The more comprehensive dataset used in our analysis allowed us to uncover commonalities and differences across MLOs that would not be brought to light in individual studies, usually performed on reduced and often different sets of associated proteins. Taken together, the generated information should be a good starting point for future research on a particular protein, protein groups or specific MLOs.

It is important to note that the categorization of proteins according to their role in the LLPS process derived from the different databases is far from perfect. As such, it reflects the currently limited information available in the literature about many MLO's and their precise cellular components, and how the latter are modified in different cellular states. This in turn is mirrored by the sparse and incomplete annotations in the source databases, a situation that is however bound to change. With the flurry of publications reporting studies on well known or new experimentally verified LLPS systems, new information on proteins playing key roles in these systems is accumulating fast. A good example is the XPO1 protein, only recently assigned the role of Client in the NPC, mediating the transport of >200 proteins [8], and is therefore not annotated as such in the databases.

In addition, much of what we know about MLOs and have been able to analyze so far, is derived from organelles in human cells, leaving a gap in our knowledge of equivalent systems in other organisms. We expect however, that many of the trends observed in the present study should hold in organisms across the Eukaryotic kingdoms, and conversely, that the information obtained on MLOs from other organisms should be valuable in interpreting the data on human MLOs. But extracting meaningful biological knowledge from analyzing these systems will require deriving community-wide guidelines for defining MLOs in terms of the molecular players involved, their roles in the LLPS process and their biological functions. These guidelines should form the basis for a controlled vocabulary that would help standardize information retrieval and database annotations in this fast developing field.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We thank Dr. Javier Iserte and Dr. Franco L. Simonetti for the discussions and for proofreading the manuscript. FO, AR are fellows of the National Research Council (CONICET), CMB is researcher of the CONICET.

## Funding

FO and AMN are PhD fellows and CMB is a researcher of CONICET. This work was partially fund by PICT-2018-01015.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.06.042.

## References

[1] Ader, C., S. Frey, W. Maas, H. B. Schmidt, D. Görlich, and M. Baldus. 2010. "Amyloid-like Interactions within Nucleoporin FG Hydrogels." Proceedings of the National Academy of Sciences of the United States of America 107 (14). https://doi.org/10.1073/pnas.0910163107.

[2] Aguzzi A, Altmeyer M. Phase Separation: Linking Cellular Compartmentalization to Disease. Trends Cell Biol 2016;26(7):547–58. https://doi.org/10.1016/j.tcb.2016.03.004.

[3] Alberti S. The Wisdom of Crowds: Regulating Cell Function through Condensed States of Living Matter. J Cell Sci 2017;130(17):2789–96.

[4] Alberti S, Gladfelter A, Mittag T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell 2019;176 (3):419–34.

[5] Alberti S, Halfmann R, King O, Kapila A, Lindquist S. A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins. Cell 2009;137(1):146–58.

[6] Alberti S, Hyman AA. Biomolecular Condensates at the Nexus of Cellular Stress, Protein Aggregation Disease and Ageing. Nat Rev Mol Cell Biol 2021;22 (3):196–213.

[7] Ambadipudi S, Biernat J, Riedel D, Mandelkow E, Zweckstetter M. Liquid-Liquid Phase Separation of the Microtubule-Binding Repeats of the Alzheimer-Related Protein Tau. Nat Commun 2017;8(1):275.

[8] Azmi AS, Uddin MH, Mohammad RM. "The Nuclear Export Protein XPO1 – from Biology to Targeted Therapy". *Nature Reviews*. Clinical Oncology 2021;18 (3):152–69. https://doi.org/10.1038/s41571-020-00442-4.

[9] Bahram F, von der Lehr N, Cetinkaya C, Larsson LG. C-Myc Hot Spot Mutations in Lymphomas Result in Inefficient Ubiquitination and Decreased Proteasome-Mediated Turnover. Blood 2000;95(6):2104–10.

[10] Banani SF, Rice AM, Peeples WB, Lin Y, Jain S, Parker R, et al. Compositional Control of Phase-Separated Cellular Bodies. Cell 2016;166(3):651–63.

[11] Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular Condensates: Organizers of Cellular Biochemistry. Nat Rev Mol Cell Biol 2017;18 (5):285–98. https://doi.org/10.1038/nrm.2017.7.

[12] Bayés À, Collins MO, Croning MDR, van de Lagemaat LN, Choudhary JS, Grant SGN, et al. Comparative Study of Human and Mouse Postsynaptic Proteomes Finds High Compositional Conservation and Abundance Differences for Key Synaptic Proteins. PLoS ONE 2012;7(10):e46683. https://doi.org/10.1371/journal.pone.0046683.

[13] Beck M, Hurt Ed. The Nuclear Pore Complex: Understanding Its Function through Structural Insight. Nat Rev Mol Cell Biol 2017;18(2):73–89. https://doi.org/10.1038/nrm.2016.147.

[14] Berchtold D, Battich N, Pelkmans L. A Systems-Level Study Reveals Regulators of Membrane-Less Organelles in Human Cells. Mol Cell 2018;72 (6):1035–1049.e5.

[15] Berry, Joel, Stephanie C. Weber, Nilesh Vaidya, Mikko Haataja, and Clifford P. Brangwynne. 2015. "RNA Transcription Modulates Phase Transition-Driven Nuclear Body Assembly." Proceedings of the National Academy of Sciences of the United States of America 112 (38): E5237–45.

[16] Buchan JR, Parker R. Eukaryotic Stress Granules: The Ins and Outs of Translation. Mol Cell 2009;36(6):932–41.

[17] Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, et al. Protein Phase Separation: A New Phase in Cell Biology. Trends Cell Biol 2018;28(6):420–35.

[18] Candelise N, Scaricamazza S, Salvatori I, Ferri A, Valle C, Manganelli V, et al. Protein Aggregation Landscape in Neurodegenerative Diseases: Clinical Relevance and Future Applications. Int J Mol Sci 2021;22(11):6016. https://doi.org/10.3390/ijms22116016.

[19] Celetti G, Paci G, Caria J, VanDelinder V, Bachand G, Lemke EA. The Liquid State of FG-Nucleoporins Mimics Permeability Barrier Properties of Nuclear Pore Complexes. The Journal of Cell Biology 2020;219(1). https://doi.org/10.1083/jcb.201907157.

[20] F.T.M. Chang J.D. McGhie F.L. Chan M.C. Tang M.A. Anderson J.R. Mann et al. PML Bodies Provide an Important Platform for the Maintenance of Telomeric Chromatin Integrity in Embryonic Stem Cells 41 8 2013 2013 4447 4458.

[21] Chen, Weizhong, Zhangming Yan, Simin Li, Norman Huang, Xuerui Huang, Jin Zhang, and Sheng Zhong. 2018. "RNAs as Proximity-Labeling Media for Identifying Nuclear Speckle Positions Relative to the Genome." iScience 4 (June): 204–15.

[22] Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, et al. John N Hutchinson, Sergio A Sara, Alexander W Ensminger, Archa H Fox, Andrew Chess, Jeanne B Lawrence An Architectural Role for a Nuclear Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles". Mol Cell 2009;33(6):717–26.

[23] Dellaire, G., C. H. Eskiw, H. Dehghani, R. W. Ching, and D. P. Bazett-Jones. 2006. "Mitotic Accumulations of PML Protein Contribute to the Re-Establishment of PML Nuclear Bodies in G1." Journal of Cell Science 119 (Pt 6). https://doi.org/10.1242/jcs.02817.

[24] Mitrea DM, Kriwacki RW. Phase Separation in Biology; Functional Organization of a Higher Order. Cell Communication and Signaling: CCS 2016;14(1). https://doi.org/10.1186/s12964-015-0125-7.

[25] Ditlev JA, Case LB, Rosen MK. Who's In and Who's Out-Compositional Control of Biomolecular Condensates. J Mol Biol 2018;430(23):4666–84.

[26] Elbaum-Garfinkle, Shana, Younghoon Kim, Krzysztof Szczepaniak, Carlos Chih-Hsiung Chen, Christian R. Eckmann, Sua Myong, and Clifford P. Brangwynne. 2015. "The Disordered P Granule Protein LAF-1 Drives Phase Separation into Droplets with Tunable Viscosity and Dynamics." Proceedings of the National Academy of Sciences of the United States of America 112 (23): 7189–94.

[27] Farahi N, Lazar T, Wodak SJ, Tompa P, Pancsa R. Concentration and Dosage Sensitivity of Proteins Driving Liquid-Liquid Phase Separation. Cold Spring Harbor Laboratory 2021. https://doi.org/10.1101/2021.02.19.430946.

[28] Fare CM, Villani A, Drake LE, Shorter J. Higher-Order Organization of Biomolecular Condensates. Open Biology 2021;11(6):210137. https://doi.org/10.1098/rsob.210137.

[29] Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, et al. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell 2016;165(7):1686–97.

[30] Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: The Protein Families Database. Nucleic Acids Res 2014;42(D1):D222–30.

[31] K.-W. Fong Y. Li W. Wang W. Ma K. Li R.Z. Qi et al. Whole-Genome Screening Identifies Proteins Localized to Distinct Nuclear Bodies 203 1 2013 149 164.

[32] Fox, A. H., and A. I. Lamond. 2010. "Paraspeckles." Cold Spring Harbor Perspectives in Biology 2 (7). https://doi.org/10.1101/cshperspect.a000687.

[33] Géli V, Lisby M. Recombinational DNA Repair Is Regulated by Compartmentalization of DNA Lesions at the Nuclear Pore Complex. BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology 2015;37(12):1287–92. https://doi.org/10.1002/bies.201500084.

[34] Grousl, T., P. Ivanov, I. Frýdlová, P. Vasicová, F. Janda, J. Vojtová, K. Malínská, et al. 2009. "Robust Heat Shock Induces eIF2alpha-Phosphorylation-Independent Assembly of Stress Granules Containing eIF3 and 40S Ribosomal Subunits in Budding Yeast, Saccharomyces Cerevisiae." Journal of Cell Science 122 (Pt 12). https://doi.org/10.1242/jcs.045104.

[35] Halfmann R, Jarosz DF, Jones SK, Chang A, Lancaster AK, Lindquist S. Prions Are a Common Mechanism for Phenotypic Inheritance in Wild Yeasts. Nature 2012;482(7385):363–8. https://doi.org/10.1038/nature10875.

[36] M. Hanazawa M. Yonetani A. Sugimoto PGL Proteins Self Associate and Bind RNPs to Mediate Germ Granule Assembly in C. Elegans 192 6 2011 929 937 10.1083/jcb.201010106.

[37] Harmon TS, Holehouse AS, Pappu RV. Differential Solvation of Intrinsically Disordered Linkers Drives the Formation of Spatially Organized Droplets in Ternary Systems of Linear Multivalent Proteins. New J Phys 2018;20 (4):045002. https://doi.org/10.1088/1367-2630/aab8d9.

[38] Harmon, Tyler S., Alex S. Holehouse, Michael K. Rosen, and Rohit V. Pappu. 2017. "Intrinsically Disordered Linkers Determine the Interplay between Phase Separation and Gelation in Multivalent Proteins." eLife 6 (November). https://doi.org/10.7554/eLife.30294.

[39] S. Hennig G. Kong T. Mannen A. Sadowska S. Kobelke A. Blythe et al. Prion-like Domains in RNA Binding Proteins Are Essential for Building Subnuclear Paraspeckles 210 4 2015 529 539.

[40] Hoelz A, Debler EW, Blobel G. The Structure of the Nuclear Pore Complex. Annu Rev Biochem 2011;80(1):613–43. https://doi.org/10.1146/annurev-biochem-060109-151030.

[41] Hornbeck, Peter V., Jon M. Kornhauser, Vaughan Latham, Beth Murray, Vidhisha Nandhikonda, Alex Nord, Elżbieta Skrzypek, Travis Wheeler, Bin Zhang, and Florian Gnad. 2019. "15 Years of PhosphoSitePlus®: Integrating Post-Translationally Modified Sites, Disease Variants and Isoforms." Nucleic Acids Research 47 (Database issue): D433.

[42] Hornbeck, Peter V., Bin Zhang, Beth Murray, Jon M. Kornhauser, Vaughan Latham, and Elzbieta Skrzypek. 2015. "PhosphoSitePlus, 2014: Mutations,

PTMs and Recalibrations." Nucleic Acids Research 43 (Database issue): D512–20.
[43] Hülsmann B, Labokha A, Görlich D. The Permeability of Reconstituted Nuclear Pores Provides Direct Evidence for the Selective Phase Model. Cell 2012;150 (4):738–51. https://doi.org/10.1016/j.cell.2012.07.019.
[44] Hutchinson, J. N., A. W. Ensminger, C. M. Clemson, C. R. Lynch, J. B. Lawrence, and A. Chess. 2007. "A Screen for Nuclear Transcripts Identifies Two Linked Noncoding RNAs Associated with SC35 Splicing Domains." BMC Genomics 8 (February). https://doi.org/10.1186/1471-2164-8-39.
[45] A.M. Ishov A.G. Sotnikov D. Negorev O.V. Vladimirova N. Neff T. Kamitani et al. PML Is Critical for ND10 Formation and Recruits the PML-Interacting Protein Daxx to This Nuclear Structure When Modified by SUMO-1 147 2 1999 221 234 10.1083/jcb.147.2.221.
[46] Jain BP, Pandey S. WD40 Repeat Proteins: Signalling Scaffold with Diverse Functions. Protein J 2018;37(5):391–406.
[47] Jarmoskaite I, Russell R. DEAD-Box Proteins as RNA Helicases and Chaperones. Wiley Interdisciplinary Reviews. RNA 2011;2(1):135–52.
[48] Ji S, Luo Y, Cai Q, Cao Z, Zhao Y, Mei J, et al. LC Domain-Mediated Coalescence Is Essential for Otu Enzymatic Activity to Extend Drosophila Lifespan. Mol Cell 2019;74(2):363–377.e5.
[49] Kalkat, Manpreet, Jason De Melo, Katherine Ashley Hickman, Corey Lourenco, Cornelia Redel, Diana Resetca, Aaliya Tamachi, William B. Tu, and Linda Z. Penn. 2017. "MYC Deregulation in Primary Human Cancers." Genes 8 (6). https://doi.org/10.3390/genes8060151.
[50] Kametani F, Yoshida M, Matsubara T, Murayama S, Saito Y, Kawakami I, et al. Comparison of Common and Disease-Specific Post-Translational Modifications of Pathological Tau Associated With a Wide Range of Tauopathies. Front Neurosci 2020;14. https://doi.org/10.3389/fnins.2020.581936.
[51] Kim SJ, Fernandez-Martinez J, Nudelman I, Shi Yi, Zhang W, Raveh B, et al. Integrative Structure and Functional Anatomy of a Nuclear Pore Complex. Nature 2018;555(7697):475–82.
[52] Kim TH, Tsang B, Vernon RM, Sonenberg N, Kay LE, Forman-Kay JD. Phospho-Dependent Phase Separation of FMRP and CAPRIN1 Recapitulates Regulation of Translation and Deadenylation. Science 2019;365(6455):825–9. https://doi.org/10.1126/science:aax4240.
[53] V. Lallemand-Breitenbach H. de The PML Nuclear Bodies Cold Spring Harbor Perspectives in Biology 2 5 2010 a000661 a661.
[54] Lin Y, Currie SL, Rosen MK. Intrinsically Disordered Sequences Enable Modulation of Protein Phase Separation through Distributed Tyrosine Motifs. The Journal of Biological Chemistry 2017;292(46):19110–20.
[55] Q. Li X. Peng Y. Li W. Tang J. Zhu J. Huang et al. LLPSDB: A Database of Proteins Undergoing Liquid–liquid Phase Separation in Vitro 48 D1 2020 2020 D320 D327.
[56] Liu Y, Liu F, Grundke-Iqbal I, Iqbal K, Gong C-X. Brain Glucose Transporters, O-GlcNAcylation and Phosphorylation of Tau in Diabetes and Alzheimer's Disease. J Neurochem 2009;111(1):242–9.
[57] Logan MK, McLaurin DM, Hebert MD, Misteli T. Synergistic Interactions between Cajal Bodies and the miRNA Processing Machinery. Mol Biol Cell 2020;31(15):1561–9.
[58] Luo Y, Na Z, Slavoff SA. P-Bodies: Composition, Properties, and Functions. Biochemistry 2018;57(17):2424–31.
[59] Machyna M, Heyn P, Neugebauer KM. "Cajal Bodies: Where Form Meets Function". *Wiley Interdisciplinary Reviews*. RNA 2013;4(1):17–34. https://doi.org/10.1002/wrna.1139.
[60] Malempati S, Tibbitts D, Cunningham M, Akkari Y, Olson S, Fan G, et al. Aberrant Stabilization of c-Myc Protein in Some Lymphoblastic Leukemias. Leukemia 2006;20(9):1572–81.
[61] Mediani L, Guillén-Boixet J, Alberti S, Carra S. Nucleoli and Promyelocytic Leukemia Protein (PML) Bodies Are Phase Separated Nuclear Protein Quality Control Compartments for Misfolded Proteins. Molecular & Cellular Oncology 2019;6(6):e1415624. https://doi.org/10.1080/23723556.2019.1652519.
[62] B. Mészáros G. Erdős B. Szabó É. Schád Á. Tantos R. Abukhairan et al. PhaSePro: The Database of Proteins Driving Liquid–liquid Phase Separation 2020 10.1093/nar/gkz848.
[63] Mobbs GW, Hoelz A. Nucleoporin Condensates Drive Nuclear Pore Complex Assembly in Oocytes. Trends Biochem Sci 2020;45(4):278–80.
[64] Monahan Z, Ryan VH, Janke AM, Burke KA, Rhoads SN, Zerze GH, et al. Phosphorylation of the FUS Low-Complexity Domain Disrupts Phase Separation, Aggregation, and Toxicity. The EMBO Journal 2017;36 (20):2951–67.
[65] Nguyen CD, Mansfield RE, Leung W, Vaz PM, Loughlin FE, Grant RP, et al. Characterization of a Family of RanBP2-Type Zinc Fingers That Can Recognize Single-Stranded RNA. J Mol Biol 2011;407(2):273–83. https://doi.org/10.1016/j.jmb.2010.12.041.
[66] W. Ning Y. Guo S. Lin B. Mei Y.u. Wu P. Jiang et al. DrLLPS: A Data Resource of Liquid-Liquid Phase Separation in Eukaryotes 48 D1 2020 2020 D288 D295 10.1093/nar/gkz1027.
[67] Nott T, Petsalaki E, Farber P, Jervis D, Fussner E, Plochowietz A, et al. Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. Mol Cell 2015;57(5):936–47.
[68] Oliveira, Guilherme A. P. de, Yraima Cordeiro, Jerson L. Silva, and Tuane C. R. G. Vieira. 2019. "Liquid-Liquid Phase Transitions and Amyloid Aggregation in Proteins Related to Cancer and Neurodegenerative Diseases." Advances in Protein Chemistry and Structural Biology 118 (September): 289–331.
[69] Pancsa Rita, Vranken Wim, Mészáros Bálint. Computational Resources for Identifying and Describing Proteins Driving Liquid-Liquid Phase Separation. *Briefings Bioinf*, February. 2021. https://doi.org/10.1093/bib/bbaa408.
[70] Pancsa Rita, Schad Eva, Tantos Agnes, Tompa Peter. Emergent Functions of Proteins in Non-Stoichiometric Supramolecular Assemblies. Biochim Biophys Acta, Proteins Proteomics 2019;1867(10):970–9. https://doi.org/10.1016/j.bbapap.2019.02.007.
[71] Patel Samir S, Belmont Brian J, Sante Joshua M, Rexach Michael F. Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex. Cell 2007;129(1):83–96.
[72] Peng Kang-Wei, Liou Ying-Ming. Differential Role of Actin-Binding Proteins in Controlling the Adipogenic Differentiation of Human CD105-Positive Wharton's Jelly Cells. BBA 2012;1820(4):469–81.
[73] Pichler Andrea, Gast Andreas, Seeler Jacob S, Dejean Anne, Melchior Frauke. The Nucleoporin RanBP2 Has SUMO1 E3 Ligase Activity. Cell 2002;108 (1):109–20.
[74] Damiano Piovesan Francesco Tabaro Lisanna Paladin Marco Necci Ivan Mičetić Carlo Camilloni et al. MobiDB 3.0: More Annotations for Intrinsic Disorder, Conformational Diversity and Interactions in Proteins 46 D1 2018 2018 D471 D476 10.1093/nar/gkx1071.
[75] Qamar Seema, Wang GuoZhen, Randle Suzanne J, Ruggeri Francesco Simone, Varela Juan A, Lin Julie Qiaojin, et al. FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. Cell 2018;173(3):720–734.e15. https://doi.org/10.1016/j.cell.2018.03.056.
[76] Radivojac Predrag, Iakoucheva Lilia M, Oldfield Christopher J, Obradovic Zoran, Uversky Vladimir N, Dunker A Keith. Intrinsic Disorder and Functional Proteomics. Biophys J 2007;92(5):1439–56.
[77] Saito Makoto, Hess Daniel, Eglinger Jan, Fritsch Anatol W, Kreysing Moritz, Weinert Brian T, et al. Acetylation of Intrinsically Disordered Regions Regulates Phase Separation. Nat Chem Biol 2019;15(1):51–61.
[78] Sawyer Iain A, Sturgill David, Dundr Miroslav. "Membraneless Nuclear Organelles and the Search for Phases within Phases". *Wiley Interdisciplinary Reviews*. RNA 2019;10(2):e1514. https://doi.org/10.1002/wrna.2019.10.issue-210.1002/wrna.1514.
[79] Schmidt H Broder, Görlich Dirk. Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. Trends Biochem Sci 2016;41 (1):46–61.
[80] Smith Temple F, Gaitatzes Chrysanthe, Saxena Kumkum, Neer Eva J. The WD Repeat: A Common Architecture for Diverse Functions. Trends Biochem Sci 1999;24(5):181–5.
[81] Spannl Stephanie, Tereshchenko Maria, Mastromarco Giovanni J, Ihn Sean J, Lee Hyun O. Biomolecular Condensates in Neurodegeneration and Cancer. Traffic 2019;20(12):890–911.
[82] Taniue Kenzui, Akimitsu Nobuyoshi. Aberrant Phase Separation and Cancer. *The FEBS Journal*, February. 2021. https://doi.org/10.1111/febs.15765.
[83] Thandapani Palaniraja, O'Connor Timothy R, Bailey Timothy L, Richard Stéphane. Defining the RGG/RG Motif. Mol Cell 2013;50(5):613–23.
[84] The UniProt Consortium. UniProt: The Universal Protein Knowledgebase. Nucleic Acids Res 2017;45(D1):D158–69.
[85] Tran Elizabeth J, Wente Susan R. Dynamic Nuclear Pore Complexes: Life on the Edge. Cell 2006;125(6):1041–53. https://doi.org/10.1016/j.cell.2006.05.027.
[86] Twyffels Laure, Gueydan Cyril, Kruys Véronique. Shuttling SR Proteins: More than Splicing Factors. The FEBS Journal 2011;278(18):3246–55.
[87] Uversky Vladimir N. Intrinsically Disordered Proteins in Overcrowded Milieu: Membrane-Less Organelles, Phase Separation, and Intrinsic Disorder. Curr Opin Struct Biol 2017;44(June):18–30.
[88] Vetter Ingrid R, Nowak Christine, Nishimoto Takeharu, Kuhlmann Jürgen, Wittinghofer Alfred. Structure of a Ran-Binding Domain Complexed with Ran Bound to a GTP Analogue: Implications for Nuclear Transport. Nature 1999;398(6722):39–46.
[89] Vizcarra, Christina L., Barry Kreutz, Avital A. Rodal, Angela V. Toms, Jun Lu, Wei Zheng, Margot E. Quinlan, and Michael J. Eck. 2011. "Structure and Function of the Interacting Domains of Spire and Fmn-Family Formins." Proceedings of the National Academy of Sciences of the United States of America 108 (29): 11884–89.
[90] Wang Jie, Choi Jeong-Mo, Holehouse Alex S, Lee Hyun O, Zhang Xiaojie, Jahnel Marcus, et al. A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. Cell 2018;174(3):688–699.e16.
[91] Jason A. West Mari Mito Satoshi Kurosaka Toru Takumi Chiharu Tanegashima Takeshi Chujo et al. Structural, Super-Resolution Microscopy Analysis of Paraspeckle Nuclear Body Organization 214 7 2016 817 830.
[92] Wheeler, R. J., and A. A. Hyman. 2018. "Controlling Compartmentalization by Non-Membrane-Bound Organelles." Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 373 (1747). https://doi.org/10.1098/rstb.2017.0193.
[93] Wippich Frank, Bodenmiller Bernd, Trajkovska Maria Gustafsson, Wanka Stefanie, Aebersold Ruedi, Pelkmans Lucas. Dual Specificity Kinase DYRK3 Couples Stress Granule Condensation/dissolution to mTORC1 Signaling. Cell 2013;152(4):791–805.
[94] Wu Jian, Matunis Michael J, Kraemer Doris, Blobel Günter, Coutavas Elias. Nup358, a Cytoplasmically Exposed Nucleoporin with Peptide Repeats, Ran-GTP Binding Sites, Zinc Fingers, a Cyclophilin A Homologous Domain, and a Leucine-Rich Region. The Journal of Biological Chemistry 1995;270 (23):14209–13.

[95] Xu Songli, Powers Maureen A, Weis Karsten. In Vivo Analysis of Human Nucleoporin Repeat Domain Interactions. Mol Biol Cell 2013;24(8):1222–31.
[96] Yamada Kayo, Ono Motoharu, Perkins Neil D, Rocha Sonia, Lamond Angus I. Identification and Functional Characterization of FMN2, a Regulator of the Cyclin-Dependent Kinase Inhibitor p21. Mol Cell 2013;49(5):922–33.
[97] Yaseen, N. R., and G. Blobel. 1999. "Two Distinct Classes of Ran-Binding Sites on the Nucleoporin Nup-358." Proceedings of the National Academy of Sciences of the United States of America 96 (10): 5516–21.
[98] Kaiqiang You Qi Huang Chunyu Yu Boyan Shen Cristoffer Sevilla Minglei Shi et al. PhaSepDB: A Database of Liquid-Liquid Phase Separation Related Proteins 48 D1 2020 2020 D354 D359 10.1093/nar/gkz847.
[99] Zahler A M, Lane W S, Stolk J A, Roth M B. SR Proteins: A Conserved Family of Pre-mRNA Splicing Factors. Genes Dev 1992;6(5):837–47.
[100] Zbinden Aurélie, Pérez-Berlanga Manuela, De Rossi Pierre, Polymenidou Magdalini. Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force. Dev Cell 2020;55(1):45–68.
[101] Zeng Menglong, Shang Yuan, Araki Yoichi, Guo Tingfeng, Huganir Richard L, Zhang Mingjie. Phase Transition in Postsynaptic Densities Underlies Formation of Synaptic Complexes and Synaptic Plasticity. Cell 2016;166(5):1163–1175.e12.