



Article

# The Patterning and Proportion of Charged Residues in the Arginine-Rich Mixed-Charge Domain Determine the Membrane-Less Organelle Targeted by the Protein

Tamami Miyagi <sup>1</sup>, Rio Yamazaki <sup>1</sup>, Koji Ueda <sup>2</sup>, Satoshi Narumi <sup>3</sup> , Yuhei Hayamizu <sup>4</sup> , Hiroshi Uji-i <sup>5,6</sup>, Masahiko Kuroda <sup>1,\*</sup> and Kohsuke Kanekura <sup>1,\*</sup>

<sup>1</sup> Department of Molecular Pathology, Tokyo Medical University, 6-1-1 Shinjuku, Shinjuku-ku, Tokyo 160-8402, Japan; s121060@tokyo-med.ac.jp (T.M.); s122066@tokyo-med.ac.jp (R.Y.)

<sup>2</sup> Cancer Proteomics Group, Cancer Precision Medicine Center, Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto-ku, Tokyo 135-8550, Japan; koji.ueda@jfcf.or.jp

<sup>3</sup> Department of Molecular Endocrinology, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya-ku, Tokyo 157-8535, Japan; narumi-s@ncchd.go.jp

<sup>4</sup> Department of Materials Science and Engineering, School of Materials and Chemical Technology, Tokyo Institute of Technology, 2-12-1, Ookayama, Meguro-ku, Tokyo 152-8550, Japan; hayamizu.y.aa@m.titech.ac.jp

<sup>5</sup> Research Institute for Electronic Science (RIES), Hokkaido University, Kita 10 Nishi 20, Kita Ward, Sapporo 001-0020, Japan; hiroshi.ujii@es.hokudai.ac.jp

<sup>6</sup> Division of Information Science and Technology, Graduate School of Information Science and Technology, Hokkaido University, Kita 14 Nishi 9, Kita Ward, Sapporo 001-0020, Japan

\* Correspondence: kuroda@tokyo-med.ac.jp (M.K.); kanekura@tokyo-med.ac.jp (K.K.)



**Citation:** Miyagi, T.; Yamazaki, R.; Ueda, K.; Narumi, S.; Hayamizu, Y.; Uji-i, H.; Kuroda, M.; Kanekura, K. The Patterning and Proportion of Charged Residues in the Arginine-Rich Mixed-Charge Domain Determine the Membrane-Less Organelle Targeted by the Protein. *Int. J. Mol. Sci.* **2022**, *23*, 7658. <https://doi.org/10.3390/ijms23147658>

Academic Editor: Vladimir N. Uversky

Received: 13 June 2022

Accepted: 8 July 2022

Published: 11 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Membrane-less organelles (MLOs) are formed by biomolecular liquid–liquid phase separation (LLPS). Proteins with charged low-complexity domains (LCDs) are prone to phase separation and localize to MLOs, but the mechanism underlying the distributions of such proteins to specific MLOs remains poorly understood. Recently, proteins with Arg-enriched mixed-charge domains (R-MCDs), primarily composed of R and Asp (D), were found to accumulate in nuclear speckles via LLPS. However, the process by which R-MCDs selectively incorporate into nuclear speckles is unknown. Here, we demonstrate that the patterning of charged amino acids and net charge determines the targeting of specific MLOs, including nuclear speckles and the nucleolus, by proteins. The redistribution of R and D residues from an alternately sequenced pattern to uneven blocky sequences caused a shift in R-MCD distribution from nuclear speckles to the nucleolus. In addition, the incorporation of basic residues in the R-MCDs promoted their localization to the MLOs and their apparent accumulation in the nucleolus. The R-MCD peptide with alternating amino acids did not undergo LLPS, whereas the blocky R-MCD peptide underwent LLPS with affinity to RNA, acidic poly-Glu, and the acidic nucleolar protein nucleophosmin, suggesting that the clustering of R residues helps avoid their neutralization by D residues and eventually induces R-MCD migration to the nucleolus. Therefore, the distribution of proteins to nuclear speckles requires the proximal positioning of D and R for the mutual neutralization of their charges.

**Keywords:** liquid–liquid phase separation; membrane-less organelle; nuclear speckle; nucleolus

## 1. Introduction

Eukaryotic cells carry out a wide variety of biochemical reactions simultaneously. Some of these reactions lead to toxic byproducts and are mutually incompatible. Hence, these biochemical reactions occur in parallel in isolated compartments in typical lipid-membrane-bound organelles. In contrast, cells form dynamic ribonucleoprotein compartments to respond rapidly to various external and internal stimuli [1,2]. These ribonucleoprotein bodies, formed through the liquid–liquid phase separation (LLPS) of proteins

and RNA, are known as membrane-less organelles (MLOs) [3]. MLOs such as nuclear speckles, stress granules, and the nucleolus are formed according to the requirements of the cell and are eliminated when unused. These structures can change their size, fluidity, and protein composition according to requirements, to mediate a dynamic cellular response. The most important factor in distinguishing between membrane-bound organelles and MLOs is the ability of material to move freely across the boundaries of the organelle. This is attributed to the presence or absence of a lipid membrane. In addition, due to this feature, disease-associated proteins such as amyotrophic lateral sclerosis (ALS)-causing *C9ORF72* dipeptides are often incorporated into MLOs and influence their dynamics [4]. MLOs have constituent proteins with low-complexity domains (LCDs) and/or intrinsically disordered regions (IDRs). The repetitive motifs of LCDs promote multivalent interaction, and the flexibility of IDRs makes them strong driving factors of phase separation [5,6]. Multivalent interactions occurring through LCD-IDRs promote the polymerization of macromolecules, leading to phase separation and MLO formation. However, the process of distribution of MLO-targeting proteins to specific MLOs remains elusive.

The localization of molecules to the nucleolus is the most widely studied process in MLO targeting. The consensus regarding the nucleolus localization signals (NoLS) is a sequence with a cluster of basic amino acids such as arginine (R) and lysine (K) [7,8]. Findings from the comparison of various NoLS and proteomic studies on the nucleolus suggest that NoLS do not contain any specific sequences that act as a recognition motif for binding to specific receptor molecules [9,10]; instead, they undergo high-affinity interactions with nucleolar molecules such as proteins and nucleic acids, thus facilitating nucleolar localization. However, the precise mechanism remains unknown. The targeting of nuclear speckles by proteins has also been investigated. A nuclear speckle is an MLO with an essential role in mRNA splicing, the splicing factor storage, and modification [11]. SON protein and serine/arginine repetitive matrix protein 2 (SRRM2), both natively unfolded proteins with large LCDs, have been identified as the scaffolding molecules for nuclear speckles [12]. Recently, the R-enriched mixed-charge domain (R-MCD) was shown to determine whether proteins will localize to nuclear speckles [13]. Interestingly, the combination of aspartate (D) and R in proteins triggers their incorporation into nuclear speckles, whereas the presence of glutamate (E) and K exerts a limited effect on protein localization. However, the role of the R-MCD structure–function relationship in protein distribution to nuclear speckles is yet to be fully characterized.

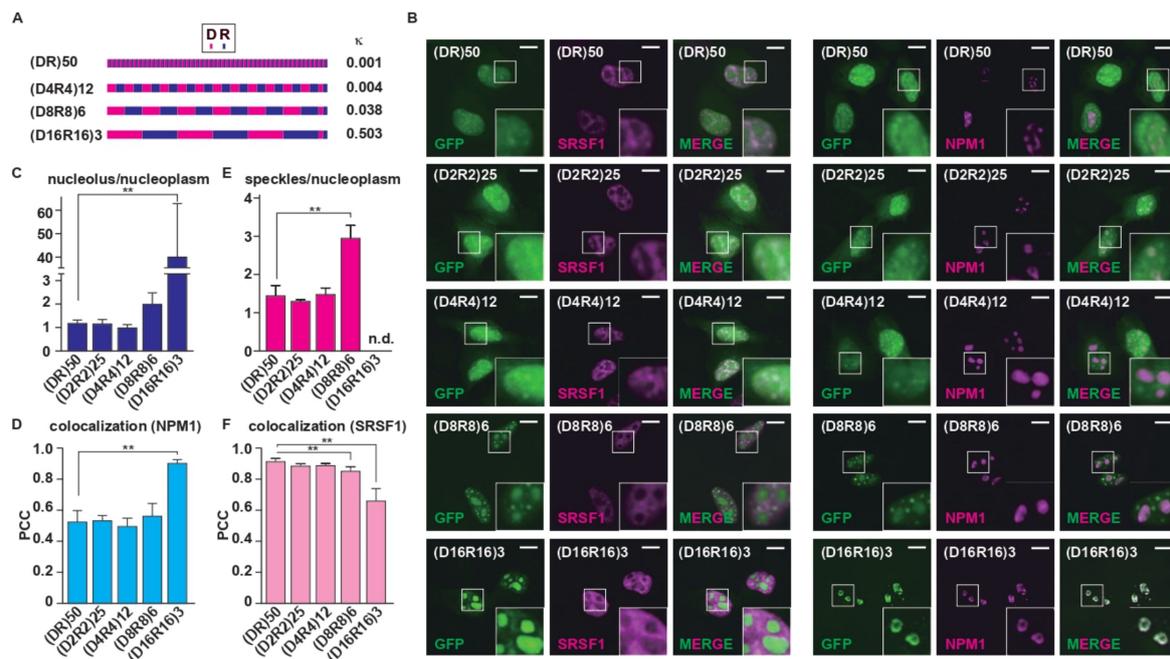
In this study, we first tested a series of artificial R-MCD variants with different periodicities or different proportions to determine how the charge position and net charge affect the distribution of proteins to the MLOs. When the positive and negative charges were separated, the localization of R-MCDs shifted from nuclear speckles to the nucleolus, even though the net charge remained neutral. Furthermore, we showed that the incorporation of R-MCD into MLOs was enhanced when their net charge was positive. We performed interactome analysis by proximity-dependent biotin labeling and showed that proteins from the nucleolar granular components (GCs) were enriched in the interactome of R-MCD with blocky positive/negative charges. Lastly, the R-MCD peptide with opposite charges segregated was found to be susceptible to phase separation and associated with RNA, poly-E peptide, and nucleophosmin (NPM1), a nucleolar protein. These results indicate that the net charge of the LCD and the position of the charged amino acids in the LCD determine the degree of phase separation and the localization of proteins to specific MLOs.

## 2. Results

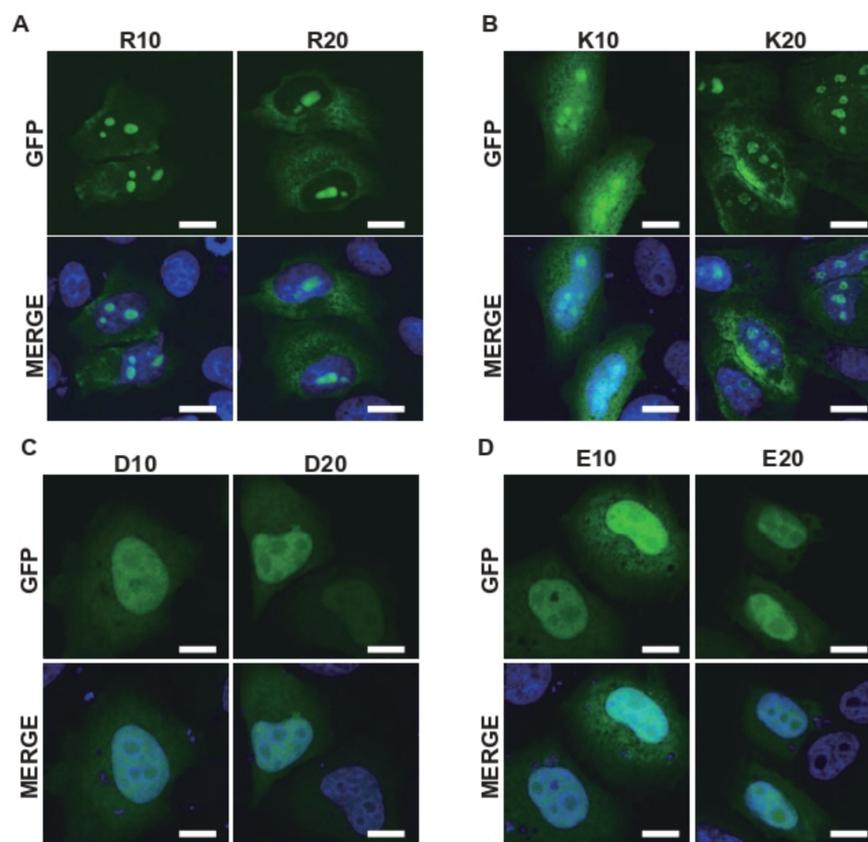
### 2.1. Charge Patterning in the R-MCDs Influences Their Distribution to MLOs

To investigate how the position of oppositely charged residues in the R-MCDs affects their subcellular localization, we synthesized (DR)<sub>50</sub> variants with different charge distributions as models for R-MCD (Figure 1A). All (DR)<sub>50</sub> variants carried the same number of R ( $n = 50$ ) and D ( $n = 50$ ) residues, and the charge patterning was the only difference, as indicated by the charge patterning parameter  $\kappa$  (Figures 1A and S1) [14].

As previously reported, when overexpressed in HeLa cells, green fluorescent protein (GFP)-(DR)<sub>50</sub> colocalized with serine/arginine-rich splicing factor 1 (SRSF1), a nuclear speckle marker (Figure 1B) [15]. When we changed the charge patterning without changing the net charge, (D<sub>8</sub>R<sub>8</sub>)<sub>6</sub> tended to accumulate in the nucleolus, and (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> localized almost exclusively to the nucleolus, as shown by a good Pearson correlation coefficient (PCC) with a nucleolar marker, NPM1-DsRed [16] (Figure 1B–F). Reportedly, the localization of proteins to the nucleolus requires the presence of basic amino acids such as R and K at high levels [7–9]. However, in the case of (DR)<sub>50</sub> variants, even though both (DR)<sub>50</sub> and (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> had the same number of D and R residues and equal net charges, (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> migrated to the nucleolus, whereas (DR)<sub>50</sub> accumulated in nuclear speckles. This suggests that uneven intramolecular charge distribution may be a key determinant of protein localization to MLOs. To elucidate the roles of the acidic and basic blocks of (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> in the nucleolar localization, we tested the localization of contiguous acidic and basic amino acids. When we overexpressed R<sub>10</sub>, R<sub>20</sub>, K<sub>10</sub>, or K<sub>20</sub>, contiguous basic amino acid chains of different lengths showed nucleolar localization (Figure 2A,B). However, D<sub>10</sub>, D<sub>20</sub>, E<sub>10</sub>, or E<sub>20</sub> expressed in HeLa cells showed no specific localization to MLOs (Figure 2C,D), suggesting that the localization of (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> to the nucleolus is owing to the presence of contiguous basic amino acids. These results indicate that the primary sequence of oppositely charged residues determines the MLO protein targets.



**Figure 1.** The periodicity of the (DR)<sub>50</sub> repeat determines the MLOs targeted by the R-MCD. (A) Schematic structures of (DR)<sub>50</sub> variants with different repeat periodicities. D<sub>2</sub>R<sub>2</sub> was fused to the C-terminus of (D<sub>4</sub>R<sub>4</sub>)<sub>12</sub>, (D<sub>8</sub>R<sub>8</sub>)<sub>6</sub>, and (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> to equalize the number of D and R residues. Kappa (κ: charge patterning parameter) was calculated by CIDER [6]. (B) Subcellular localization of GFP-fused R-MCD variants in HeLa cells. Nuclear speckles were visualized by co-expressed serine/arginine-rich splicing factor 1 (SRSF1)-mCherry, and the nucleolus was visualized by co-expressed nucleophosmin (NPM1)-DsRed. Scale bar = 10 μm. The inset images are high-magnification images of squared areas. (C) Ratio of GFP signals in the nucleolus/nucleoplasm, with *n* = 8–10 cells/condition. (D) The degree of colocalization between GFP-(DR)<sub>50</sub> variants and NPM1 quantified by the Pearson correlation coefficient (PCC). (E) The ratio of GFP signals in nuclear speckles/nucleoplasm. The ratio for (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> was not determined because only a few cells showed speckle incorporation. *N* = 8–10 cells/condition. (F) The degree of colocalization between GFP-(DR)<sub>50</sub> variants and SRSF1, quantified by the PCC. The error bars show ± SD. The asterisks indicate significant differences obtained via one-way analysis of variance (ANOVA) with Dunnett’s test; \*\* *p* < 0.01.

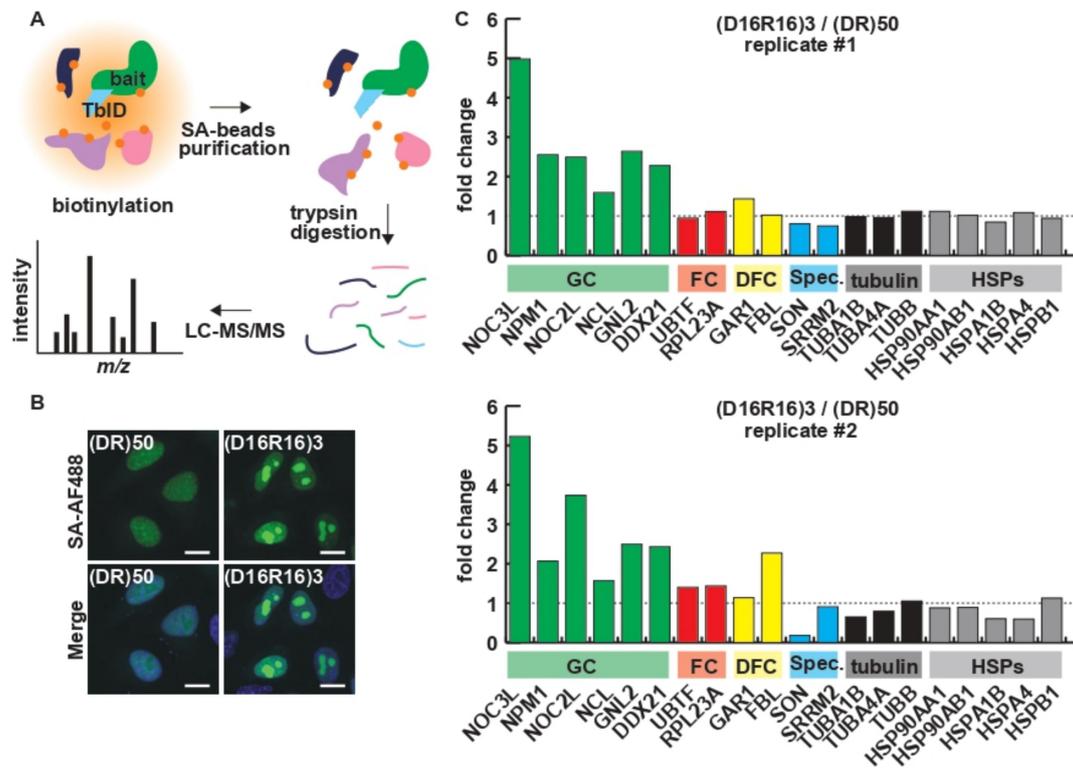


**Figure 2.** The presence of consecutive basic amino acids is responsible for the nucleolar distribution of proteins. (A–D) The subcellular localization of ten or twenty consecutive basic amino acids (R for (A) and K for (B)) or acidic amino acids (D for (C) and E for (D)) fused to the C-terminus of GFP expressed in HeLa cells. Scale bar = 10  $\mu$ m.

## 2.2. Charge Patterning Affects the Interactome of R-MCDs

As indicated by the incorporation of (DR)<sub>50</sub> into the nuclear speckle and the localization of (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> to the nucleolus, charge patterning affects the protein–protein interaction of R-MCDs. Since R and D are highly charged residues, conventional immunoprecipitation/proteomics analysis, which requires cell lysis, may lead to artificial interactions during the lysis/immunoprecipitation process. To investigate the protein–protein interactions while maintaining spatial information, we performed TurboID-mediated proximity biotinylation of the proteomes in close proximity to each R-MCD (Figure 3A) [17,18]. The visualization of biotinylated proteins using AlexaFluor488-conjugated streptavidin revealed that TurboID-(DR)<sub>50</sub> exclusively biotinylated proteins in nuclear speckles, and TurboID-(D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> biotinylated the proteins in the nucleolus, as expected (Figure 3B). Biotinylated proteomes were analyzed using quantitative liquid chromatography–mass spectrometry (LC–MS) [19]. Signals from nucleolar markers and speckle markers showed that the segregation of oppositely charged residues in the R-MCD increased the cohesion of the nucleolar GC, where rRNA and proteins assemble in the ribosomal subunit, and reduced the cohesion to speckle markers (Figure 3C). Some nuclear and cytoplasmic MLOs have been shown to require RNA to maintain their integrity [20]. RNA has a negative charge and can interact with basic amino acids via electrostatic interactions, and RNA depletion by intranuclear expression of RNase induces morphological changes in MLOs [20]. Furthermore, the inhibition of RNA transcription by actinomycin D (ActD) is known to cause the enlargement of nuclear speckles and dwarfing of the nucleolus [21–23]. A decrease in the levels of negatively charged RNAs alters the subcellular localization of proteins with a high basic amino acid content [24]. We investigated how the ActD-mediated inhibition of RNA

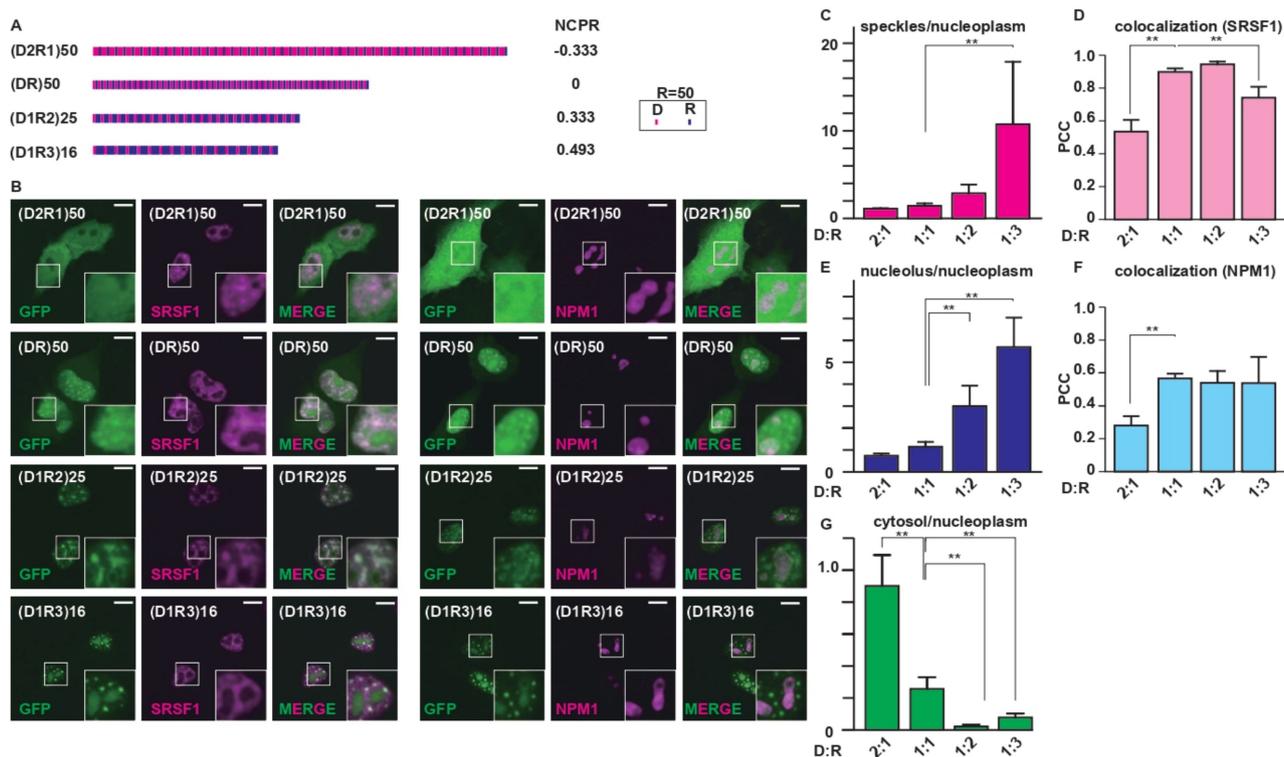
synthesis affects the localization of R-MCD. The distribution of (DR)<sub>50</sub>, which localizes to nuclear speckles, showed no significant change in colocalization with SRSF1 after ActD treatment (Figure S2A–C). However, the distribution of (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub>, which localizes to the nucleolus, was significantly affected upon ActD treatment and showed a granular pattern in the nucleoplasm (Figure S2D–F), suggesting that electrostatic interaction of (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> with RNA may be one of the drivers of its distribution to the nucleolus.



**Figure 3.** Proximity labeling analysis revealed that charge patterning affects the interactome of R-MCDs. (A) Schematic illustration of TurboID (TbID)-mediated proximity biotinylation. (DR)<sub>50</sub> or (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> was used as the bait in the experiment. SA: streptavidin. (B) Protein biotinylation by TurboID-(DR)<sub>50</sub> or TurboID-(D<sub>16</sub>R<sub>16</sub>)<sub>3</sub>, followed by visualization with AlexaFluor488-labeled streptavidin (SA-AF488). Scale bar = 10 μm. (C) Fold change in the signals of markers for the granular component (GC), fibrillar center (FC), or dense fibrillar component (DFC) of the nucleolus, nuclear speckles (Spec.), tubulin species, and heat-shock proteins (HSPs). Proximity labeling analysis followed by liquid chromatography–tandem mass spectrometry (LC–MS/MS) was performed twice for independently prepared sample sets.

### 2.3. A Net Positive Charge in R-MCDs Enhances Their Incorporation into MLOs

Next, we examined how the net charge of an R-MCD affects its subcellular localization. To this end, we synthesized R-MCD variants containing the same number of R residues ( $n = 50$ ) with different net charges (Figure 4A). The net charge per residue (NCPR) of these R-MCD variants ranged from  $-0.333$  to  $0.493$ . An increase in the NCPR through R up to  $+0.2$  was shown to enhance the incorporation of (DR)<sub>50</sub> mutants into nuclear speckles [13]. GFP-(D<sub>2</sub>R<sub>1</sub>)<sub>50</sub> (NPCR =  $-0.333$ ) showed a diffused distribution pattern and low cohesion with nuclear speckles (Figure 4B). When the proportion of R-MCD changed and the number of acidic amino acids was reduced to impart a net positive charge, the condensation of R-MCDs increased not only in nuclear speckles but also in the nucleolus (Figure 4B–G). This effect was more evident when the net charge increased, which suggests that an increase in the net positive charge of R-MCDs enhances their condensation in MLOs (Figure 4C,E). Note that when the net charge of R-MCDs was negative, their nuclear localization ceased and they diffused into the nucleus and cytoplasm (Figure 4G).



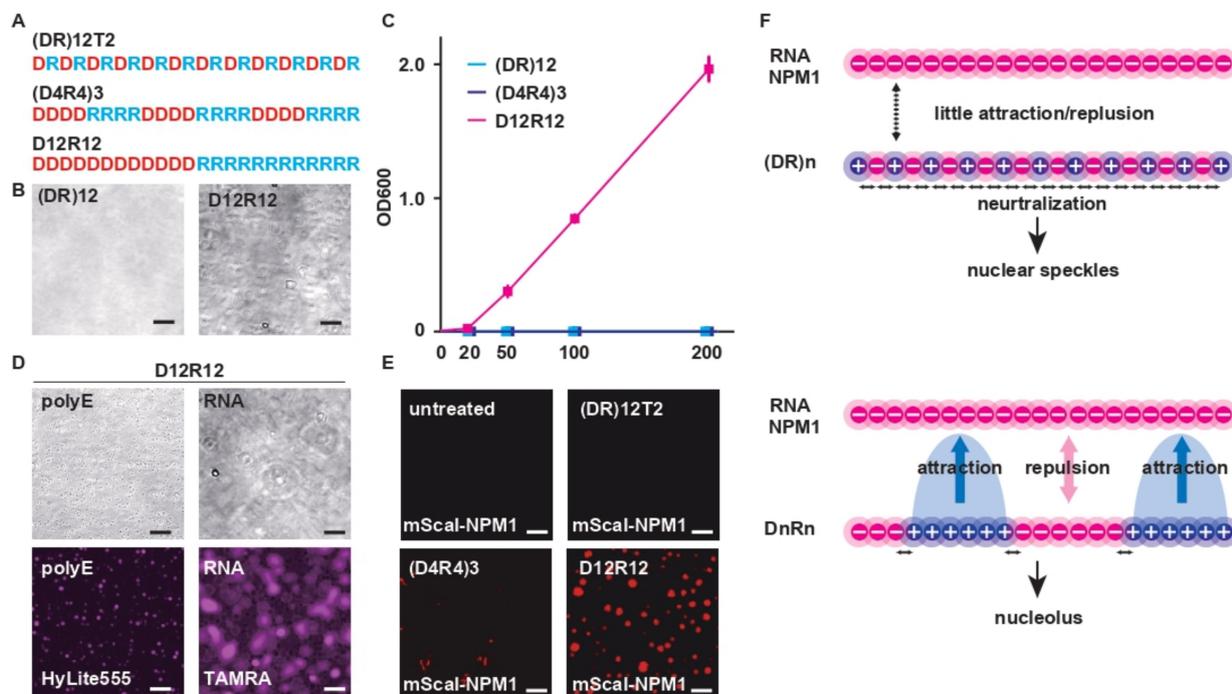
**Figure 4.** Imparting a net positive charge to the R-MCD enhances its ability to target MLOs. (A) Schematic structures of (DR)<sub>50</sub> proportional variants with different net charges per residue (NCPR) [6]. (B) Subcellular localization of GFP-fused (DR)<sub>50</sub> proportional variants in HeLa cells. Nuclear speckles were visualized by co-expressed SRSF1-mCherry, and the nucleolus was visualized by co-expressed NPM1-DsRed. Scale bar = 10  $\mu$ m. The inset images are high-magnification images of squared areas. (C) The ratio of GFP signals in nuclear speckles/nucleoplasm, with  $n = 10\sim 12$  cells/condition. (D) The degree of colocalization between GFP-(DR)<sub>50</sub> proportional variants and SRSF1 quantified by PCC. (E) The ratio of GFP signals in the nucleolus/nucleoplasm, with  $n = 10\sim 12$  cells/condition. (F) The degree of colocalization between GFP-(DR)<sub>50</sub> proportional variants and NPM1 quantified by PCC. (G) The ratio of GFP signals in the cytosol/nucleoplasm, with  $n = 10\sim 12$  cells/condition. The error bars show  $\pm$  SD. The asterisks indicate significant differences derived using one-way ANOVA with Dunnett's test; \*\*  $p < 0.01$ .

#### 2.4. Charge Patterning in Natural R-MCDs Determines Protein Distribution to Specific MLOs

Next, we tested the effect of charge patterning on the R-MCDs of natural nuclear speckle-associated proteins. Wild-type (wt) full lengths of small nuclear ribonucleoprotein U1 subunit 70 (SNRNP70-FL) and wt-SNRNP70-MCD were localized to nuclear speckles (Figure 5A–C). When oppositely charged amino acids were segregated without changing the net charge, the site of localization shifted from the nuclear speckles to the nucleolus. Similarly, the site of localization of the full-length protein and MCD of negative elongation factor complex member E (NELFE) also changed from the nuclear speckles to the nucleolus when the distribution of charged amino acids was changed (Figure 5D–F). These results show that the charge patterning effects observed in the artificial R-MCD are similar to those in natural proteins and are a means of controlling the localization of proteins to MLOs in cells.



GFP-(D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> with a digitonin solution containing different concentrations of NaCl. After cell permeabilization, (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> localized to the nucleolus even in the presence of 200 mM NaCl, but (DR)<sub>50</sub> lost its localization to nuclear speckles in the presence of 100 mM NaCl, suggesting that (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> binds to the surrounding molecules via stronger electrostatic forces than (DR)<sub>50</sub> (Figure S3A,B).



**Figure 6.** Charge patterning determines the phase-separation properties of R-MCD peptides. (A) The structure of (DR)<sub>12</sub> variants with different periodicities. (B) D<sub>12</sub>R<sub>12</sub>, but not (DR)<sub>12</sub>, underwent simple coacervation. Scale bar: 10 μm. (C) The turbidity of (DR)<sub>12</sub> variants with different periodicities at different concentrations determined using the OD<sub>600</sub>, with *n* = 3. The error bars show ± SD. (D) The D<sub>12</sub>R<sub>12</sub> droplets mixed with 10 nM TAMRA-labeled rA15 RNA or 10 ng/μL of HyLite555-labeled polyE peptide. (E) Recombinant mScarlet-I-NPM1 protein mixed with each (DR)<sub>12</sub> variant peptide (10 μM). (F) Scheme of intramolecular neutralization by oppositely charged amino acids in the vicinity of (DR)<sub>n</sub> and DnRn.

### 3. Discussion

MLOs form and disappear dynamically in the nucleus and cytoplasm and play essential roles in cellular homeostasis. The appropriate distribution of each MLO-specific component to the designated MLOs is essential for the maintenance of their functions, but the mechanism underlying their distribution is unknown. In this study, we showed that the charge pattern in R-MCD determines the phase separation of proteins to MLOs and that segregation of residues with opposite charges influences the interactome of R-MCD, promoting phase separation and shifting their distribution from nuclear speckles to the nucleolus.

The (DR)<sub>50</sub> variants with different periodicities used in this study were fused with enhanced GFP and had an isoelectric point of 5.6. Although EGFP and EGFP-(DR)<sub>50</sub>, with different periodicities, have similar isoelectric points (pI = 5.6), the EGFP-(DR)<sub>50</sub> variants are localized to nuclear speckles or the nucleolus. Thus, even though the net charge of the R-MCD is zero, the abundance of basic amino acids acts as a nuclear-localizing signal and induces nuclear protein distribution.

Nuclear speckles are the nuclear MLOs with an important role in mRNA splicing, storage, maturation, and splicing-factor modification [25]. The mechanism of nuclear speckle formation remains unknown. Recently, SRRM2 and SON were identified as speckle

scaffold proteins [12]. While SRRM2 and SON were shown to be essential for nuclear speckle formation, the process by which SRRM2 and SON form nuclear speckles has not been determined. Many splicing factors that localize to nuclear speckles harbor D/R or S/R repeats [26–28]. The D/R and S/R repeats are reportedly essential for the localization of factors to nuclear speckles [13]. Poly(PR) and poly(GR) dipeptides produced from the mutant *C9ORF72* gene, which is associated with ALS [29,30], are structurally similar to (DR)<sub>50</sub> in that they contain alternating R residues; however, their net charge is positive and their targeted MLO is the nucleolus [31], suggesting that the neutralization of the charge on R by neighboring acidic amino acids is important for the localization of factors to nuclear speckles. This is consistent with the fact that when phosphorylated, serine residues in the S/R repeat sequence act as acidic amino acids, to neutralize R-MCDs and induce their localization to nuclear speckles [13].

Peptides with an alternating sequence of acidic and basic amino acids are known as zwitterionic peptides. Zwitterionic peptides are easily hydrated and interact poorly with proteins and charged molecules [32,33]. Because many strongly charged molecules such as RNA are present in the nucleus, natural R-MCDs use their D/R-alternating zwitterionic structure to reduce their susceptibility to the strong attractive force of RNA and localize to nuclear speckles. Conversely, once the oppositely charged residues are segregated and the local charge unevenness is established, the zwitterionic peptide loses its properties and accumulates in the nucleolus by changing its molecular binding mode. Although both the nucleolus and nuclear speckle contain RNA, the nucleolus contains much more abundant RNA than the nuclear speckles, as shown by strong signals when intracellular RNA is stained [34], and the abundant RNA-derived electrostatic forces might attract the blocky charged sequences, which (DR)<sub>50</sub> can escape. On the other hand, the mechanism of how (DR)<sub>50</sub> specifically condenses to nuclear speckles remains unclear. The zwitterionic structure may be advantageous for specific interaction with SON, SRRM2, or other components of nuclear speckles. The intra-organellar environmental differences such as hydrophobicity and density of aromatic compounds might also have an influence. Supporting the idea that the (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> has stronger electrostatic interactions than (DR)<sub>50</sub>, (D<sub>16</sub>R<sub>16</sub>)<sub>3</sub> remained in the nucleolus when cells were permeabilized with digitonin and treated with salts that shield electrostatic forces, whereas (DR)<sub>50</sub> lost its localization to the nuclear speckle (Figure S3). Nevertheless, the possibility that blocked charged residues localize to the nucleolus by interactions with specific nucleolar molecules cannot be ruled out. It also remains unclear why a net positive charge in (DR)<sub>50</sub> variants enhances their incorporation into nuclear speckles as well as the nucleolus. The NCPR of (D<sub>1</sub>R<sub>3</sub>)<sub>50</sub> is 0.493 (Figure S1), which is similar to that of the ALS-causing *C9ORF72*-encoded R-rich dipeptides, poly(GR) and poly(PR) (NPCR = 0.5 for both), which exclusively localize to the nucleolus [31]. Therefore, we believe that the positioning of the anionic D residue next to the cationic R residue is necessary for protein localization to the speckle; however, further studies are warranted.

Only few studies have reported the role of the charge pattern of R-MCD in the protein localization to MLOs. Greig et al. reported that modest unevenness of charge distribution in the R-MCD enhances cohesion for both the nucleolus and nuclear speckles, and also a shift in the net charge of R-MCD to positive increases the size of the nuclear speckle and triggers retention of mRNA in the nucleus [13]. Therefore, the charge distribution of the R-MCD plays an important role not only in the subcellular localization but also in the protein function.

Research on the mechanism of localization to each MLO has only commenced recently. MLOs are involved in various intracellular functions, and their disruption has been implicated in various diseases such as neurodegenerative diseases and malignant tumors. However, only a few drugs target MLOs. Klein et al. recently showed that small molecules such as cisplatin are incorporated into the droplets of super-enhancers; however, their controlled delivery to specific MLOs is yet to be achieved [35]. The targeting of specific MLOs by charge patterning, as shown here, may help develop novel drug delivery systems (DDS) in the future.

In conclusion, we found that charge patterning is important for the targeted distribution of proteins in specific MLOs and that it regulates the protein interactome as well as the manner of phase separation in MCDs. We believe that these data are of significance from both biological and medical perspectives and will help develop DDS that can target specific MLOs.

#### 4. Materials and Methods

##### 4.1. Recombinant Plasmids Construction

All MCD constructs and cDNAs for SNRNP70 (full length and MCD), NELFE (full length and MCD), and TurboID were synthesized and subcloned into pcDNA3.1-N-eGFP vector or pcDNA3.1 vector (for TurboID) by GenScript (Piscataway, NJ, USA). NPM1-DsRed was a gift from Mary Dasso (Addgene plasmid # 34,553; <http://n2t.net/addgene:34553> (accessed on 6 July 2022); RRID:Addgene\_34553). For bacterial expression, human NPM1 tagged with mScarlet-I cDNA was synthesized and subcloned into pET28a vector by GenScript. Human SRSF1 cDNA was PCR-amplified from a HeLa-cell cDNA pool with a sense primer (GGATCCATGTCGGGAGGTGGTGTG) and an antisense primer (GAATTCCTATGTACGAGAGCGAGATCTG) using PrimeSTAR Max DNA polymerase (Takara Bio, Tokyo, Japan) and subcloned into pmCherry-C1 vector (Clontech, Mountain View, CA, USA).

##### 4.2. Transfection of HeLa Cells and Image Acquisition by Confocal Microscopy

HeLa cells (RIKEN BRC, Tsukuba, Japan) plated on a chambered glass slide (Matsunami, Osaka, Japan) were transfected with each GFP-MCD construct in the presence or absence of an MLO marker using Lipofectamine 2000 (Thermo Fisher Scientific, Waltham, MA, USA), in accordance with the manufacturer's instructions. Twenty-four hours after transfection, the cells were fixed with formalin-PBS, and coverslips were mounted using ProLong Gold-DAPI antifade reagent (Thermo Fisher Scientific). For ActD treatment, the transfected HeLa cells were treated with 100 nM of ActD (Sigma-Aldrich, St. Louis, MO, USA) for the indicated period. For the digitonin permeabilization assay, the transfected HeLa cells were treated with a buffer containing 10 mM HEPES (pH 7.4), 0.1% digitonin (Nacalai, Kyoto, Japan), 300 mM sucrose, and 100 mM or 200 mM NaCl for 10 min before fixation. The cells were imaged with FV10i confocal microscopy (Olympus, Tokyo, Japan). Image analyses were performed using the ImageJ Fiji software (<http://rsbweb.nih.gov/ij>, accessed on 6 July 2022).

##### 4.3. Peptide Synthesis and Purification of Recombinant mScarlet-I-NPM1

All (DR)<sub>12</sub> variant peptides were synthesized by GenScript. Trifluoroacetic acid was substituted with acetic acid. For the expression of the recombinant mScarlet-I-NPM1 protein, *Escherichia coli* BL21 cells were transformed with pET28a-mScarlet-I-NPM1, and expression of NPM1 was induced by incubation of the cells in 1mM IPTG for 16 h at 25 °C. The recombinant protein was purified using a HisTalon gravity column (Clontech) and dialyzed with Tris-buffered saline.

##### 4.4. Phase Separation of Peptides

The (DR)<sub>12</sub> or D<sub>12</sub>R<sub>12</sub> peptide (1 mM) was dissolved in water and diluted in a phase separation buffer (composed of 10 mM HEPES (pH 7.4) and 100 mM NaCl) at 100 μM. The droplets were observed by FV10i confocal microscopy, and the OD<sub>600</sub> value was measured using a NanoDrop One (Thermo Fisher). The droplets were mixed with 10 nM TAMRA-labeled rA15 RNA (Fasmac, Kanagawa, Japan) or 10 ng/μL of HyLite555-labeled polyE peptide [34]. For phase separation with mScarlet-I-NPM1, 1 μM mScarlet-I-NPM1 was mixed with each (DR)<sub>12</sub> variant peptide (10 μM) in the phase separation buffer.

#### 4.5. Proximity Labeling and Quantitative LC–MS

TurboID-mediated proximity labeling was performed as previously described [17]. Briefly, HeLa cells plated on a 15 cm dish were transfected with TurboID-(DR)<sub>50</sub> or TurboID-(D<sub>16</sub>R<sub>16</sub>)<sub>3</sub>. Twenty-four hours after transfection, the cells were incubated with 50  $\mu$ M biotin for 2 h. After the cells were washed four times with ice-cold PBS, they were lysed using RIPA buffer containing cOMplete protease inhibitor cocktail (Roche). Excess biotin was removed by ultrafiltration using an Amicon Ultra filter. For purifying biotinylated proteins, cell lysates were incubated with Dynabeads MyOne streptavidin beads (Thermo Fisher Scientific) in PBS for 2 h. The beads were sequentially washed with RIPA buffer, 1 M KCl, 0.1 M Na<sub>2</sub>CO<sub>3</sub>, 2 M urea in 10 mM Tris-HCl (pH 7.5), and RIPA buffer. The proteins were eluted by boiling with 2 $\times$  Laemmli sample buffer (Bio-Rad, Hercules, CA, USA) and subjected to quantitative LC–MS [19]. Briefly, the extracted samples were reduced with 10 mM TCEP at 100 °C for 10 min, alkylated with 50 mM iodoacetamide at room temperature for 45 min, and separated using SDS-PAGE. Protein bands, visualized via Coomassie brilliant blue staining, were excised, destained, and cut finely before in-gel digestion with Trypsin/Lys-C Mix (Promega) at 37 °C for 12 h. The digested peptides were analyzed with an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) in combination with an UltiMate 3000 RSLC nano-flow HPLC system (Thermo Fisher Scientific) in the HCD MS/MS mode. Peptides were identified and quantified using Proteome Discoverer 2.4 software (Thermo Fisher Scientific), where the MS/MS spectra were searched against the *Homo sapiens* protein database in SwissProt (<https://www.uniprot.org/>, accessed on 6 July 2022), with a false discovery rate of 1% as an identification cutoff.

#### 4.6. Colocalization Analysis

The colocalization analysis of GFP-(DR)<sub>50</sub> variants with NPM1 or SRSF1 was performed with Fiji ImageJ software using the EzColocalization plugin [36], and the PCC for each combination was calculated.

#### 4.7. Bioinformatic Analysis of Biochemical Properties of Peptides and Proteins

The biochemical properties of peptides and proteins used in this study were analyzed using the CIDER web server (<http://pappulab.wustl.edu/CIDER/analysis/>, accessed on 6 July 2022) [6]. The following components were analyzed and are shown in Figure S1: FCR, fraction of charged residues; NCPR, net charge per residue; Kappa ( $\kappa$ ), charge patterning parameter; and hydropathy, the Kyte–Doolittle hydropathy score for the sequence on a scale of 0–9.

#### 4.8. Statistic Analysis

Data are represented as means  $\pm$  standard deviations. Statistical analysis of the data was performed using SPSS software 28 (IBM).

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23147658/s1>.

**Author Contributions:** Conceptualization, T.M. and K.K.; data curation, T.M., K.U. and K.K.; formal analysis, T.M. and K.K.; funding acquisition, S.N., Y.H., H.U.-i., M.K. and K.K.; investigation, T.M., R.Y., K.U. and K.K.; methodology, T.M., K.U., S.N., Y.H., H.U.-i. and K.K.; project administration, K.K.; resources, K.K.; supervision, M.K. and K.K.; validation, K.K.; visualization, K.K.; writing—original draft, T.M. and K.K.; writing—review and editing, T.M., S.N., Y.H., H.U.-i., M.K. and K.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was financially supported by the JSPS KAKENHI (20H03593 to K.K., 21H02706 to M.K., 21H04634 to H.U., 19H03627 to S.N., 20H02564 to Y.H.).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be provided by the authors without undue reservation.

**Acknowledgments:** We thank Mary Dasso for providing us with NPM1-DsRed (Addgene: # 34553).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Su, Q.; Mehta, S.; Zhang, J. Liquid-liquid phase separation: Orchestrating cell signaling through time and space. *Mol. Cell* **2021**, *81*, 4137–4146. [[CrossRef](#)] [[PubMed](#)]
2. Wright, P.E.; Dyson, H.J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29. [[CrossRef](#)] [[PubMed](#)]
3. Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **2019**, *176*, 419–434. [[CrossRef](#)] [[PubMed](#)]
4. Lee, K.-H.; Zhang, P.; Kim, H.J.; Mitrea, D.M.; Sarkar, M.; Freibaum, B.D.; Cika, J.; Coughlin, M.; Messing, J.; Molliex, A.; et al. C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* **2016**, *167*, 774–788.e17. [[CrossRef](#)]
5. Babinchak, W.M.; Surewicz, W.K. Liquid-Liquid Phase Separation and Its Mechanistic Role in Pathological Protein Aggregation. *J. Mol. Biol.* **2020**, *432*, 1910–1925. [[CrossRef](#)]
6. Holehouse, A.S.; Das, R.K.; Ahad, J.N.; Richardson, M.O.; Pappu, R.V. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys. J.* **2017**, *112*, 16–21. [[CrossRef](#)]
7. Martin, R.M.; Ter-Avetisyan, G.; Herce, H.D.; Ludwig, A.K.; Lattig-Tunnemann, G.; Cardoso, M.C. Principles of protein targeting to the nucleolus. *Nucleus* **2015**, *6*, 314–325. [[CrossRef](#)]
8. Musinova, Y.R.; Kananykhina, E.Y.; Potashnikova, D.M.; Lisitsyna, O.M.; Sheval, E.V. A charge-dependent mechanism is responsible for the dynamic accumulation of proteins inside nucleoli. *Biochim. Biophys. Acta* **2015**, *1853*, 101–110. [[CrossRef](#)]
9. Emmott, E.; Hiscox, J.A. Nucleolar targeting: The hub of the matter. *EMBO Rep.* **2009**, *10*, 231–238. [[CrossRef](#)]
10. Stenstrom, L.; Mahdessian, D.; Gnann, C.; Cesnik, A.J.; Ouyang, W.; Leonetti, M.D.; Uhlen, M.; Cuylen-Haering, S.; Thul, P.J.; Lundberg, E. Mapping the nucleolar proteome reveals a spatiotemporal organization related to intrinsic protein disorder. *Mol. Syst. Biol.* **2020**, *16*, e9469. [[CrossRef](#)]
11. Galganski, L.; Urbanek, M.O.; Krzyzosiak, W.J. Nuclear speckles: Molecular organization, biological function and role in disease. *Nucleic Acids Res.* **2017**, *45*, 10350–10368. [[CrossRef](#)]
12. Ilik, I.A.; Malszycki, M.; Lubke, A.K.; Schade, C.; Meierhofer, D.; Aktas, T. SON and SRRM2 are essential for nuclear speckle formation. *Elife* **2020**, *9*, e60579. [[CrossRef](#)]
13. Greig, J.A.; Nguyen, T.A.; Lee, M.; Holehouse, A.S.; Posey, A.E.; Pappu, R.V.; Jedd, G. Arginine-Enriched Mixed-Charge Domains Provide Cohesion for Nuclear Speckle Condensation. *Mol. Cell* **2020**, *77*, 1237–1250.e4. [[CrossRef](#)]
14. Das, R.K.; Pappu, R.V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13392–13397. [[CrossRef](#)]
15. Tripathi, V.; Song, D.Y.; Zong, X.; Shevtsov, S.P.; Hearn, S.; Fu, X.D.; Dundr, M.; Prasanth, K.V. SRSF1 regulates the assembly of pre-mRNA processing factors in nuclear speckles. *Mol. Biol. Cell* **2012**, *23*, 3694–3706. [[CrossRef](#)]
16. Yun, C.; Wang, Y.; Mukhopadhyay, D.; Backlund, P.; Kolli, N.; Yergey, A.; Wilkinson, K.D.; Dasso, M. Nucleolar protein B23/nucleophosmin regulates the vertebrate SUMO pathway through SENP3 and SENP5 proteases. *J. Cell Biol.* **2008**, *183*, 589–595. [[CrossRef](#)]
17. Branon, T.C.; Bosch, J.A.; Sanchez, A.D.; Udeshi, N.D.; Svinkina, T.; Carr, S.A.; Feldman, J.L.; Perrimon, N.; Ting, A.Y. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **2018**, *36*, 880–887. [[CrossRef](#)]
18. Liu, F.; Morderer, D.; Wren, M.C.; Vetteson-Trutza, S.A.; Wang, Y.; Rabichow, B.E.; Salemi, M.R.; Phinney, B.S.; Oskarsson, B.; Dickson, D.W.; et al. Proximity proteomics of C9orf72 dipeptide repeat proteins identifies molecular chaperones as modifiers of poly-GA aggregation. *Acta Neuropathol. Commun.* **2022**, *10*, 22. [[CrossRef](#)]
19. Ueda, K.; Saichi, N.; Takami, S.; Kang, D.; Toyama, A.; Daigo, Y.; Ishikawa, N.; Kohno, N.; Tamura, K.; Shuin, T.; et al. A Comprehensive Peptidome Profiling Technology for the Identification of Early Detection Biomarkers for Lung Adenocarcinoma. *PLoS ONE* **2011**, *6*, e18567. [[CrossRef](#)]
20. Decker, C.J.; Burke, J.M.; Mulvaney, P.K.; Parker, R. RNA is required for the integrity of multiple nuclear and cytoplasmic membrane-less RNP granules. *EMBO J.* **2022**, *41*, e110137. [[CrossRef](#)]
21. Boulon, S.; Westman, B.J.; Hutten, S.; Boisvert, F.M.; Lamond, A.I. The nucleolus under stress. *Mol. Cell* **2010**, *40*, 216–227. [[CrossRef](#)] [[PubMed](#)]
22. Shav-Tal, Y.; Blechman, J.; Darzacq, X.; Montagna, C.; Dye, B.T.; Patton, J.G.; Singer, R.H.; Zipori, D. Dynamic sorting of nuclear components into distinct nucleolar caps during transcriptional inhibition. *Mol. Biol. Cell* **2005**, *16*, 2395–2413. [[CrossRef](#)] [[PubMed](#)]
23. Raina, K.; Rao, B.J. Mammalian nuclear speckles exhibit stable association with chromatin: A biochemical study. *Nucleus* **2022**, *13*, 58–73. [[CrossRef](#)] [[PubMed](#)]

24. Bursac, S.; Brdovcak, M.C.; Pfannkuchen, M.; Orsolich, I.; Golomb, L.; Zhu, Y.; Katz, C.; Daftuar, L.; Grabusic, K.; Vukelic, I.; et al. Mutual protection of ribosomal proteins L5 and L11 from degradation is essential for p53 activation upon ribosomal biogenesis stress. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 20467–20472. [[CrossRef](#)] [[PubMed](#)]
25. Spector, D.L.; Lamond, A.I. Nuclear speckles. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a000646. [[CrossRef](#)] [[PubMed](#)]
26. Zheng, X.; Peng, Q.; Wang, L.; Zhang, X.; Huang, L.; Wang, J.; Qin, Z. Serine/arginine-rich splicing factors: The bridge linking alternative splicing and cancer. *Int. J. Biol. Sci.* **2020**, *16*, 2442–2453. [[CrossRef](#)] [[PubMed](#)]
27. Cazalla, D.; Zhu, J.; Manche, L.; Huber, E.; Krainer, A.R.; Caceres, J.F. Nuclear export and retention signals in the RS domain of SR proteins. *Mol. Cell. Biol.* **2002**, *22*, 6871–6882. [[CrossRef](#)]
28. Staknis, D.; Reed, R. Members of a family of proteins (the RD family) detected by a U1 70K monoclonal antibody are present in spliceosomal complexes. *Nucleic Acids Res.* **1995**, *23*, 4081–4086. [[CrossRef](#)]
29. Renton, A.E.; Majounie, E.; Waite, A.; Simón-Sánchez, J.; Rollinson, S.; Gibbs, J.R.; Schymick, J.C.; Laaksovirta, H.; van Swieten, J.C.; Myllykangas, L.; et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **2011**, *72*, 257–268. [[CrossRef](#)]
30. DeJesus-Hernandez, M.; Mackenzie, I.R.; Boeve, B.F.; Boxer, A.L.; Baker, M.; Rutherford, N.J.; Nicholson, A.M.; Finch, N.A.; Flynn, H.; Adamson, J.; et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **2011**, *72*, 245–256. [[CrossRef](#)]
31. Kanekura, K.; Harada, Y.; Fujimoto, M.; Yagi, T.; Hayamizu, Y.; Nagaoka, K.; Kuroda, M. Characterization of membrane penetration and cytotoxicity of C9orf72-encoding arginine-rich dipeptides. *Sci. Rep.* **2018**, *8*, 12740. [[CrossRef](#)]
32. Ederth, T.; Lerm, M.; Orihuela, B.; Rittschof, D. Resistance of Zwitterionic Peptide Monolayers to Biofouling. *Langmuir* **2019**, *35*, 1818–1827. [[CrossRef](#)]
33. Beyer, C.D.; Thavalingam, S.; Guseva, T.; Scharadt, L.; Zimmermann, R.; Werner, C.; Dietze, P.; Bandow, J.E.; Metzler-Nolte, N.; Rosenhahn, A. Zwitterionic Peptides Reduce Accumulation of Marine and Freshwater Biofilm Formers. *ACS Appl. Mater. Interfaces* **2021**, *13*, 49682–49691. [[CrossRef](#)]
34. Chen, C.; Yamanaka, Y.; Ueda, K.; Li, P.; Miyagi, T.; Harada, Y.; Tezuka, S.; Narumi, S.; Sugimoto, M.; Kuroda, M.; et al. Phase separation and toxicity of C9orf72 poly(PR) depends on alternate distribution of arginine. *J. Cell Biol.* **2021**, *220*, e202103160. [[CrossRef](#)]
35. Klein, I.A.; Boija, A.; Afeyan, L.K.; Hawken, S.W.; Fan, M.; Dall’Agnese, A.; Oksuz, O.; Henninger, J.E.; Shrinivas, K.; Sabari, B.R.; et al. Partitioning of cancer therapeutics in nuclear condensates. *Science* **2020**, *368*, 1386–1392. [[CrossRef](#)]
36. Stauffer, W.; Sheng, H.; Lim, H.N. EzColocalization: An ImageJ plugin for visualizing and measuring colocalization in cells and organisms. *Sci. Rep.* **2018**, *8*, 15764. [[CrossRef](#)]