國立臺灣大學醫學院分子醫學研究所

碩士論文

Institute of Molecular Medicine College of Medicine
National Taiwan University

Master Thesis

# Studying Y14-RNA interaction and co-phase separation

楊謦潞

Sabrina Yeo Samuel

指導教授：譚婉玉 博士

Advisor: Woan-Yuh Tarn, Ph.D.

中華民國 112 年 7 月

July 2023

doi:10.6342/NTU202302064

**Acknowledgments**

I would like to express my heartfelt gratitude to my parents and family members, whose unwavering support and encouragement have been instrumental throughout my academic journey. Their love, guidance, and sacrifices have been a constant source of motivation, and I am deeply grateful for their belief in me.

I am also indebted to Dr. Woan Yuh Tarn, Dr. Hsin Yue Tsai, Dr. Huey Nan Wu, and Dr. Kevin Tsai for their invaluable contributions as members of my thesis committee. Their expertise, guidance, and insightful feedback have significantly shaped the direction and quality of my research. To my seniors in the lab, I extend my sincere appreciation for their mentorship and guidance. Their expertise, patience, and willingness to share their experiences have been invaluable in shaping my research skills. Their constant support and constructive feedback have played a crucial role in my development as a researcher.

Lastly, I would like to extend my deepest gratitude to the Common Equipment Core of the Institute of Biomedical Sciences at the Institute of Biomedical Sciences, Academia Sinica, for their invaluable technical assistance in light microscopy. I would also like to thank Dr. Ching-Shu Suen from the BioIT-Biostatistics Department at the Institute of Biomedical Sciences, Academia Sinica, for her enthusiastic help and support for bioinformatic analyses.



doi:10.6342/NTU202302064

**Abstract**

Y14 is a multifunctional RNA processing factor that plays an essential role in RNA metabolism and also participates in non-homologous end joining (NHEJ)-mediated repair of DNA double-strand breaks (DSBs). Our recent study has uncovered a long non-coding RNA (lncRNA) HOTAIRM1 that mediates the interaction between Y14 and the NHEJ factors. In this study, I first examined the molecular interaction between Y14 and RNA, in particular, HOTAIRM1. RNA electrophoretic mobility shift assay (REMSA) revealed that Y14 bound to multiple regions of HOTAIRM1. Our recent study also revealed that Y14 has the capability to undergo liquid-liquid phase separation (LLPS), and can co-phase-separate with HOTAIRM1. Therefore, I evaluated whether RNA structure or modification affects the LLPS properties of Y14. Moreover, since DNA damage induced m6A modification of HOTAIRM1, I also evaluated whether RNA methylation also affects the LLPS of Y14. My apparent data revealed that methylation enhanced Y14 LLPS. Since LLPS may involve electrostatic interactions of intrinsically disordered domains, I evaluated whether specific charge patterns can modulate a protein's LLPS capacity. The multi-functionality of Y14 and its biochemical properties make it a valuable target for further research in understanding its role in RNA metabolism and DSB repair.

Keywords: liquid-liquid phase separation, electrostatic interaction, RNA processing factor, DNA damage repair, N6-methyladenosine modification, intrinsically disordered domains



doi:10.6342/NTU202302064

**Table of Contents**





doi:10.6342/NTU202302064

doi:10.6342/NTU202302064

**List of Figures/Appendixes**





doi:10.6342/NTU202302064

## Introduction

### 1. Liquid-liquid phase separation

Eukaryotic cells are very densely packed with proteins and RNAs, making it challenging to regulate various biochemical reactions happening in close spatial proximity. Thus, efficient spatiotemporal organization of proteins and RNAs is vital to ensure appropriate cell function, resulting in the formation of membrane-bound and membrane-less organelles. Membrane-less organelles, also known as biomolecular condensates, are capable of selective concentration or exclusion of specific molecules to maintain biological function (Espinosa et al., 2020). Examples of known membrane-less organelles include the P bodies, stress granules, and Cajal bodies.

Biomolecular condensate formation results from liquid-liquid phase separation, a thermodynamic process that pushes multicomponent cellular mixtures into distinct phases of differing concentrations, leading to separated coexisting microenvironments (Espinosa et al., 2020). LLPS is generally driven by weak multivalent interactions such as electrostatic interactions or π–π interactions between intrinsically disordered regions (IDRs) or low complexity domains (LCDs) of proteins (Fung et al., 2018; Alberti et al., 2019). Furthermore, IDRs have been observed to phase separate on their own and even recruit other IDRs and proteins to create dynamic assemblies (Lin et al., 2015). Short-length amino acid motifs, like the charged residues Arginine/Lysine (R/K), IDRs, and LCDs, are found to follow a molecular grammar that modulates the phase separation properties of a protein (Wang et al., 2018). In addition, crowding of the microenvironment has also been shown to influence LLPS, although the exact mechanisms remain to be understood (André& Spruijt, 2020).

LLPS can be distinguished into three phases: nucleation, growth (either by coalescence or inclusion), and coarsening (Hyman et al., 2014). Nucleation can occur



doi:10.6342/NTU202302064

spontaneously via a random fluctuation of protein concentrations or induced by a specific incident, such as DNA damage. During growth, more components are concentrated in the condensate, causing it to be increasingly viscous. Finally, smaller droplets coalesce to form larger condensates, resulting in an increase in nucleated foci and a decrease in volume. In the absence of a membrane, the boundary of the condensates supports a flexible exchange of molecules between distinct phases, regulating biochemical reactions while still being constrained spatially (Banani et al., 2017). The flexible exchange of molecules allows the condensate formation to be reversible, so the cell can quickly turn on and off specific intracellular functions for highly efficient cell signaling. However, condensates can also undergo further irreversible transitions under certain conditions, leading to the formation of hydrogels or protein aggregations such as amyloid-like fibers (Alberti et al., 2019).

LLPS can be modulated by RNA in various ways. RNA can lower the concentration threshold for LLPS, and RNA has also been found to influence the size and morphology of condensates (Lin et al., 2015; Navarro et al., 2019). Furthermore, RNA structure can also influence LLPS (Langdon et al., 2018). Highly structured long non-coding RNAs (lncRNAs) like *Neat1* (nuclear paraspeckle assembly transcript 1) can act as a scaffold to promote the nucleation of biological condensates even in LLPS-inhibiting conditions (Maharana et al., 2018). In addition, RNA is also capable of phase separation in isolated events, suggesting that RNA can promote LLPS via RNA-protein and RNA-RNA interactions (Van Treeck et al., 2018). Recent studies have also shown the emerging role of N6-methyladenosine ($m^6A$) modification of RNA in LLPS (Wang et al., 2020; Lee et al., 2022). Methylation can affect the properties and structure of RNA, leading to changes in RNA-protein and RNA-RNA interactions and thus modulating LLPS.



doi:10.6342/NTU202302064

## 2. Phase separation regulates DNA damage repair

Living organisms have developed highly efficient mechanisms for dealing with intrinsic and exogenous sources of DNA damage. This is because the failure to properly repair DNA damage at the cellular level can have serious consequences, such as genomic instability, apoptosis, or senescence, which can significantly compromise the organism's survival ability.

DNA damage is repaired through different mechanisms depending on the type and severity of the damage. For example, mismatch repair (MMR) machinery can recognize and repair base misincorporation or deletion during DNA synthesis and replication. In contrast, nucleotide excision repair (NER) machinery can identify and replace abnormal bases with normal ones. Base excision repair (BER) removes single nucleotides to correct small base lesions. For double-strand breaks (DSBs), cells exploit non-homologous end-joining (NHEJ) or homologous recombination (HR), in part upon the cell cycle stage (Ciccia & Elledge, 2010).

In 2006, Bekker-Jensen et al. observed that compartmentalization of DDR factors would occur in the nucleus after DSB, a phenomenon we now define as the DNA repair foci (Bekker-Jensen et al., 2006). DNA repair foci can create the optimum microenvironment for ensuring a precise and efficient repair at the DNA damage site. Furthermore, DNA repair foci are also highly dynamic and capable of assembly and disassembly within a short time frame, allowing for the recycling of all the various repair proteins needed for DDR (Lisby et al., 2004). Previous reports have also implicated the contribution of RNA binding proteins (RBPs) to initiating phase separation at DNA damage sites to kickstart the DDR mechanisms (D'Alessandro & d'Adda di Fagagna, 2017).



doi:10.6342/NTU202302064

Recent studies suggest that DNA damage response involves LLPS. In 2017, researchers discovered a rapid increase of poly(ADP)-ribosylation (PAR), a type of post-translational modification, at the DNA damage sites (Ray Chaudhuri & Nussenzweig, 2017). DSBs will induce the recruitment of PARP1, a PAR polymerase, which will add the negatively charged PAR to proteins and thus attract the positive LCD of DDR factors such as FUS and induce LLPS. Another well-researched example is the FUS protein, an RNA processing factor that has also been implicated in DNA repair. FUS possesses an LCD with prion-like properties that easily promote LLPS even in the absence of RNA, giving rise to aggregates and amyloid-like fibrils associated with neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) (Sukhanova et al., 2020). FUS is also recruited to the DNA damage foci to interact with DDR signaling and repair cascade (Levone et al., 2021).

### 3. Role of RNA and RNA processing factors in DNA damage repair

RNA participates in various cellular processes, such as gene expression, translation, and epigenetic regulation. For DSB repair, while proteins are the primary players, RNA also plays a critical role in DDR and DNA repair (Hawley et al., 2017). When DSBs occur, the MRN complex recruits RNA Polymerase II to the damage site and triggers the formation of damage-induced lncRNAs (dilncRNAs). These dilncRNAs are subsequently processed by DROSHA/DICER to form DNA damage response RNA (DDRNAs) that will form hybrids with nascent dilncRNAs to recruit DDR factors to the damage site (Michelini et al., 2017). dilncRNAs can also pair with resected DNA ends in S/G2-phase cells to recruit HR factors (Whelan et al., 2018). In addition, lncRNAs can also interact with BRCA1 and other DDR factors to mediate DDR (Sharma et al., 2015; Haemmig et al., 2020).



doi:10.6342/NTU202302064

Accumulating evidence has also revealed that RNA processing factors play a pivotal role in regulating DNA damage response, either by regulating the mRNAs coding for DNA repair proteins or by directly interacting with the DNA repair proteins (Meng et al., 2020). For example, RBM14 is an RBP associated with transcription and RNA splicing. It can also interact with PARP1 to join the PARP-dependent DSB repair process (Kai, 2016). RNA splicing factors such as THRAP3 and BCLAF1 can also contribute to DDR by regulating the splicing and export of genes related to the HR repair pathway (Vohhodina et al., 2017).

**4. Y14 is an RNA processing factor involved in DNA damage repair**

Y14 protein (RBM8A) is an RNA processing factor, mostly known for its role as a core factor of the exon junction complex (EJC). While associated with the regulation of post-transcriptional mRNA, the EJC is best known for its function in nonsense-mediated mRNA decay by recognizing and degrading mRNAs with premature stop codons (Boehm & Gehrig, 2016). Y14 also contributes to other essential cellular functions as well, including the modulation of mitotic cell-cycle RNA splicing (Fukumura et al., 2016), regulation of apoptotic regulators (Michelle et al., 2012), and DNA damage repair (Lu et al., 2017; Chuang et al., 2019).

In particular, Y14 is crucial in maintaining genome stability through various pathways. Through its role as an EJC factor, Y14 helps to regulate the expression of several DDR factors for repairing DNA damage and preventing R-loops from accumulating during transcription. In addition, Y14, but not other EJC factors, can directly participate in double-strand break (DSB) repair by RNA-dependent interactions with Ku and other DDR factors (Chuang et al., 2019). Moreover, *HOTAIRM1* has also been identified via immunoprecipitation-RNA-seq as a lncRNA that mediates the Y14 and NHEJ complex interaction. *HOTAIRM1* accumulates at DSBs and is essential for the



doi:10.6342/NTU202302064

efficient recruitment of NHEJ and nonsense-mediated decay (NMD) factors such as Upf1/SMG6 to DSBs and subsequent DSB repair (Chuang et al., 2023). *HOTAIRM1* promotes the LLPS of Y14, further cementing its role in the Y14-mediated DNA damage repair pathway (Yu et al., 2023). However, the exact mechanism of how RNA can mediate this process remains to be investigated.



doi:10.6342/NTU202302064

## Results

### RNA Binding of Recombinant Human Y14 Protein

Previous research has already implicated the role of Y14 in the DNA damage repair pathway, likely due to its association with the lncRNA *HOTAIRM1* (Chuang et al., 2019; Chuang et al., 2023). To further understand the interaction between Y14 and *HOTAIRM1*, we first expressed recombinant wild-type Y14 in *E. coli*. We also generated two other Y14 mutants, namely the WV mutant, which has been previously identified to have a decreased RNA binding capacity (Chuang et al., 2016), and the ΔC mutant, which has been found to disrupt LLPS formation (Yu et al., 2023). The proteins were dual-tagged with an S peptide on the N-terminus to increase protein solubilization and a polyhistidine (His) tag on the C-terminus for easy purification. The recombinant S-Y14-His was then purified using Ni Sepharose 6 Fast Flow medium. Purified protein was detected via SDS-PAGE as a band around 25 kDa (Appendix 1). The recombinant protein is then used to evaluate its RNA-binding properties.

We then generated five *HOTAIRM1* fragments and labeled them with $^{32}P$ for electrophoretic mobility shift assay (EMSA). The Y14-*HOTAIRM1* complex migrated slower than the free RNA probes. As shown in Fig 1., a shift was detected upon an increasing concentration of recombinant protein. The result shows that wild-type S-Y14-His bound to all HOTAIRM1 fragments but had a higher binding affinity to the HM1-4 fragment (Fig 1A). The ΔC mutant showed a slightly higher binding affinity to all five *HOTAIRM1* fragments than the wild-type (Fig 1B), while the WV mutant did not bind as expected (Fig 1C). Due to the removal of its positively charged C terminus, the protein is less prone to undergo LLPS.

We also evaluated the binding of Y14 to RNA homopolymers. The shift was not as noticeable as the HOTAIRM1 fragments; however, a similar trend can be observed.



doi:10.6342/NTU202302064

Both wild-type and ΔC mutant bound to the $U_{25}$ fragment, with the ΔC mutant showing a higher binding affinity, while the WV mutant showed no binding (Fig 1D).

Recent research has revealed the role of $N^6$-methyladenosine ($m^6A$) in the regulation of DNA repair and genome integrity for DNA damage response (Qu et al., 2021). Using photo-crosslinking-assisted $m^6A$ sequencing (PA-$m^6A$-seq) (Chen et al., 2015), we identified A827 as an m6A site likely induced by DNA damage. Therefore, we wanted to investigate if RNA methylation can affect RNA binding. Thus, we generated a methylated (HM1-$m^6A$) and non-methylated (HM1-A) RNA oligo mimicking the $m^6A$ site on *HOTAIRM1* and performed EMSA. Similar to previous results, both WT and ΔC mutant bound to the RNA oligos, with ΔC mutant showing a higher binding affinity (Fig 1E). Unlike previous results, however, the WV mutant bound to RNA oligos. In addition, all three S-Y14-His proteins had a higher binding affinity to HM1-$m^6A$ oligo than HM1-A oligo. We hypothesized that this is due to a potential secondary structure formation in that area that might affect its binding affinity to the protein, which is then exacerbated with the presence of an $m^6A$. Taken together, we concluded that while Y14 has a non-specific binding affinity with RNA, it had a stronger binding affinity with specific RNA sequences or structures.

**Characterization of Y14 LLPS**

In our previous research, we have already observed that Y14 can form condensates independently (Yu et al., 2023). We also observed droplet formation of S-Y14-His (20 μM) in buffers ranging from pH 6 to 7.5 using a differential interference contrast (DIC) microscope and recorded similar results. S-Y14-His will form spherical condensates after 2 hours of incubation (Fig 2A). The condensate number and size decreased gradually as pH increased.



doi:10.6342/NTU202302064

To avoid any interference caused by the dual tags on our recombinant proteins, we generated a non-tagged recombinant Y14 protein. We engineered a fusion containing GB1, 6xHis tag, and a TEV protease recognition site. Non-tagged Y14 was purified according to the procedure (Appendix 2C), which Hsin-Hong Yeh kindly gifted. We then compared the droplet formation ability of S-Y14-His and non-tagged Y14 using the *in vitro* LLPS assay. As seen in Fig 2B-D., both recombinant S-Y14-His and Y14 proteins are able to form condensates, with Y14 proteins forming bigger condensates compared to S-Y14-His proteins.

Next, we evaluated the droplet formation of increasing amounts of Y14. The size of the condensates formed increases as the protein concentration increases but will lose its ability to undergo LLPS at high protein concentrations (Fig 2E). At high concentrations (80 μM), Y14 immediately formed very small condensates upon placement on the coverslip but formed solid aggregates after 5 minutes. Y14 is also capable of forming large multiphase condensates. However, it is unable to maintain them for long, as the smaller condensates are highly mobile and have a tendency to merge if close enough (Fig 2F).

**Different RNAs modulate Y14 LLPS differently**

RNA can modulate the LLPS ability of Y14 (Yu et al., 2023). We next investigated how different types of RNA modulate Y14 LLPS. We prepared four different types of RNA: Cy5-labeled ssRNA ($A_{25}$), Cy5-labeled dsRNA ($U_{25}$+$A_{25}$), $m^6A$ oligo, and non-$m^6A$ oligo, and compared Y14 droplet formation in the presence of RNA. Similar to the above result, Cy5-labeled ssRNA and dsRNA co-phase separated with S-Y14-His droplets. ssRNA had no significant effect on S-Y14-His droplet formation (Fig 3A). In contrast, the addition of dsRNA resulted in a decrease in droplet number but an increase in average droplet size.



doi:10.6342/NTU202302064

However, the addition of HM1-m$^6$A and HM1-A oligo resulted in a significant decrease in the average size of droplets (Fig 3B). However, despite the decrease in average size, the relative coverage rate of droplets formed after the addition of HM1-m$^6$A oligo showed no significant difference from the control.

Then, we repeated the same experiment using non-tagged Y14 and titrated the RNA concentrations. As seen in Fig 4AB, the addition of RNA up to 0.1 μM had no significant effects on Y14 LLPS ability. Higher concentrations of dsRNA inhibited LLPS formation, while HM1-m$^6$A RNA oligo resulted in a dramatic increase in both condensate size and coverage rate. Conversely, ssRNA and HM1-A RNA oligo only caused a slight increase in coverage rate.

**RNA Methylation promotes Y14 LLPS**

As one of the most prevalent mRNA modifications, m$^6$A is involved in various biological functions. Notably, the YTH family of proteins (YTHDFs) bind m6A. YTHDF proteins can phase separate *in vitro*, whereas m6A RNA can enhance LLPS of YTHDF proteins (Wang et al., 2020). Y14 does not have a YTH domain, but its LLPS was promoted by HM1-m$^6$A RNA oligo. Therefore, we wondered whether m$^6$A generally enhances Y14 LLPS. We generated four different unstructured oligos with varying numbers of m$^6$A in repeated sequences: m$^6$A0, m$^6$A1, m$^6$A3, and m$^6$A5. To visualize a more pronounced change in droplet formation, we performed in vitro phase separation assay at pH 7.12, resulting in smaller droplet formation compared to pH 7 (Fig 2D).

As shown in Fig 5, 0.2 μM RNA has no significant difference in condensate formation, regardless of the number of m6A, consistent with previous results. However, increasing the concentration of RNA up to 1.6 uM revealed significant changes. 1.6 uM of 0 m$^6$A had an inhibitory effect on Y14 LLPS, similar to unspecific RNA A25 oligo (results not shown). The result revealed that m$^6$A caused a notable increase in the



doi:10.6342/NTU202302064

coverage rate, *i.e.,* increasing amounts of $m^6A$ led to higher relative coverage. Specifically, $m^6A3$ and $m^6A5$ oligos further increased the coverage rate of droplets. Unexpectedly, a single $m^6A$-containing HM1 yielded higher relative coverage than $m^6A3$ and $m^6A5$ oligos. Although $m^6A$ enhanced Y14 phase separation, excess numbers of m6A had no further effects on LLPS. Nevertheless, our result showed that m6A RNA promotes the LLPS of a non-YTH protein.

**Bioinformatics analysis of charged IDRs in RNA processing and DNA repair factors**

Since charged IDRs significantly contribute to the LLPS of Y14, we attempted to assess whether other known LLPS proteins and DNA repair factors also contain charged IDRs. We first identified proteins that can undergo LLPS from LLPSDB v1.0 (Li et al., 2020), a well-curated database of proteins involved with LLPS. We then filtered these proteins through the Classification of Intrinsically Disordered Ensemble Regions (CIDER) (Holehouse et al., 2017), which analyzes a range of physicochemical properties of intrinsically disordered proteins (IDPs). We were particularly interested in the diagram of states, a predictive analysis that uses the fractions of positively and negatively charged residues to divide IDP sequences into five different conformational classes. These different conformational classes can predict the ability of IDP/IDRs to fold into three-dimensional structures, which can affect their ability to undergo LLPS. Proteins in Regions 1 and 2 possess a lower fraction of charged residue and, thus, a lower net charge per residue (NCPR), so these IDP sequences generally lend themselves to a tighter conformation. However, proteins in Region 3 have a higher fraction of charged residues and, thus, a higher NCPR value, which allows for a more open conformation. On the other hand, proteins in Regions 4 and 5 are sequences that are very positively or negatively charged in general.



doi:10.6342/NTU202302064

We identified 20 human proteins that can undergo LLPS and compared these proteins to Y14 using CIDER (Fig 6A-B). Most proteins fell under Regions 1 and 2, while only a few fell under Region 3, including Y14. Other proteins in Region 3 include HP1a, HP1b, and hNPM1, which are all well-characterized LLPS proteins. To analyze whether the structure suggested by the diagram of states would play a role in governing LLPS ability, we then identified other proteins with well-characterized LLPS abilities and compared them with Y14 (Fig 6C). The two extensively-studied LLPS proteins, FUS and G3BP1, are predicted to be in Regions 1 and 2. While the predicted open structure could not explain why Y14 can undergo LLPS, we hypothesized it could explain why Y14 can bind to RNA with no sequence. On a similar note, previous research has shown that NPM1, a known LLPS protein in Region 3, can bind to DNA and RNA with no sequence specificity (Hingorani et al., 2000). Hence, we were interested in whether a protein's ability to recognize and bind with RNA could affect its LLPS ability. We identified all the RRM-containing proteins with IDRs from various online databases, including MobiDB, PhasePro, Disprot, LLPSDB, and IUPred2, and analyzed the IDRs using CIDER (Fig 6D-E). However, as expected, the majority of IDRs were located in Regions 1 and 2 on the diagram of state. We hypothesized that protein function might play a role in determining a protein's position on the diagram of states and the optimal conditions for its LLPS.

Thus, we wanted to know if proteins with other domains could have a different distribution on the diagram of state. While most domains tend to have a larger share of proteins falling under Regions 1 and 2, some domains have a different distribution (Fig 6F). The IDRs of proteins with SAP, PWI, and DEAD domains tend to be distributed to Region 3 more than other protein domains, most likely due to its function. SAP proteins are associated with transcription, DNA repair, or RNA processing. PWI proteins are



doi:10.6342/NTU202302064

associated with splicing, while DEAD proteins are associated with RNA metabolism. The function of these proteins requires binding to various non-specific RNA, which can be helped due to their more open conformation. Meanwhile, KH domains tend to have a larger share of proteins in Regions 1 and 2, as KH proteins tend to bind to specific DNA or RNA. Thus, having a tighter conformation can help to increase binding specificity. However, as a whole, a large majority (70%) of IDRs are located in Regions 1 and 2.

Previous research has uncovered the HOTAIRM1-mediated DSB repair, including the recruitment of Y14 and other NHEJ factors to the DNA damage site. The ability of Y14 to phase separate is also thought to help facilitate the recruitment and increase the efficiency of damage repair. Hence, we wanted to find out whether other NHEJ proteins can affect the LLPS ability of Y14. We identified 23 NHEJ-related proteins and their 41 IDRs through literature research and the DisProt database and analyzed them through CIDER (Fig 6G, Table 2). In particular, we focused on Ku70, a key player in the NHEJ pathway that can bind to Y14. Ku70 has two IDRs, one with an overall negative charge and the other with a mixture of positively and negatively charged blocks (Fig 6H). We generated these two peptides and performed a phase separation assay with Y14 (data not shown). Both Ku70 peptides co-phase separated with Y14 and affected the LLPS ability of Y14. We hypothesized that proteins that can co-phase separate with Y14 may have charged IDRs, *i.e.,* in region 3 (or 4/5) of the state diagram, to facilitate electrostatic interaction.



doi:10.6342/NTU202302064

**Discussion**

LLPS is a dynamic process driven by weak, reversible interactions between proteins and/or nucleic acids, resulting in the formation of distinct liquid-phase compartments within the cell. In this study, we characterized the RNA binding ability of Y14 and its effect on phase separation. Y14 has been shown to undergo LLPS, and its RNA binding ability significantly affects this process. Y14 can form phase-separated droplets independently of RNA. Consistent with our previous research, both the wild-type and WV mutant of Y14 are able to form droplets, but only Y14 WT is able to co-condensate with RNA. As has been observed from our previous research and for other known LLPS proteins such as FUS (Maharana et al., 2018), low RNA:protein ratios can promote LLPS, while high RNA:protein ratios will inhibit LLPS. Moreover, experimental evidence shows that RNA modifications or structures, can promote and stabilize the formation of these condensates. In conclusion, this study implies that RNA can modulate LLPS based on its structure.

The methylation of adenosine in RNA has long been implicated in the regulation of RNA metabolism. However, it was not until recently that $m^6A$ was implied in the context of DNA damage repair by regulating the accumulation of R-loops (DNA:RNA hybrids formed due to DNA damage) (Abakir et al., 2019). Y14 has been shown to have a higher binding affinity to methylated RNA. This interaction is believed to play a role in enhancing the specificity and efficiency of recruitment and localization of Y14 to DNA damage sites. The recognition of methylated RNA by Y14 may enhance their multivalent interactions with the protein, leading to the formation of larger and more stable protein-RNA condensates. These condensates can then facilitate the spatiotemporal regulation of repair factors and enable more precise control over the repair process.



doi:10.6342/NTU202302064

Charge segregation is essential for LLPS. Our previous research has revealed that charge segregation within Y14, influenced by its unique distribution of charged residues, plays a significant role in LLPS behavior. However, factors such as post-translational modifications or interactions with other biomolecules can influence this charge segregation. Our results (data not shown) indicate that the addition of charged peptides may cause the disruption of charge segregation and alter the balance of attractive and repulsive forces in Y14, thus impacting its LLPS behavior. However, proteins with charged IDRs can reinforce the attractive interactions and facilitate the assembly of Y14 condensates. Further experiments are required to understand the interplay between charge segregation, electrostatic interactions, and LLPS in Y14, which can provide valuable insights into the underlying mechanisms of cellular processes involving Y14 in the context of DNA repair.

The interplay between Y14, RNA, and LLPS offers a fascinating perspective on the intricate molecular mechanisms underlying DNA damage repair. Further research in this field can contribute to our understanding of RNA-mediated processes and provide deeper insight into the spatial organization of cellular components and the translocation of RNA concerning DNA damage repair.



doi:10.6342/NTU202302064

## Experimental procedures

### Protein expression and purification

Both wild-type and mutant S-Y14-His were overexpressed in E. coli strain BL21 (DE3). Transformed bacteria were grown at 37°C in Luria broth (LB) containing kanamycin. At $OD_{600}$ 0.4–0.6, 1 mM isopropyl β-d-1-thio-galactopyranoside (IPTG) was added. Bacteria were subsequently shifted to 15°C for 12 –16 h. Cell pellets were harvested and resuspended with the resuspension buffer (0.5 M NaCl, 20 mM Tris- HCl, pH 7.9) and sonicated, followed by centrifugation at 13,000 rpm for 30 min at 4°C. The supernatant was passed over a Ni Sepharose 6 Fast Flow medium (GE Healthcare), washed with the resuspension buffer containing 60 mM imidazole, and eluted with the same buffer containing 500 mM imidazole. Purified Y14 was dialyzed against the storage buffer (50 mM KCl, 20 mM HEPES pH 7.9, 20% glycerol) and stored at −80°C. Proteins were analyzed on SDS-poly-acrylamide gel electrophoresis followed by TOOLStart Blue (BIOTOOLS) staining. Nucleic acid contamination was evaluated on a Nanodrop.

### Fluorescent labeling of proteins

Wild-type and mutant S-Y14-His and Y14 were labeled using Alexa Fluor 488 $C_5$ Maleimide (Thermo Fisher Scientific) according to the manufacturer's instructions. For visualization, the unlabeled proteins were mixed with the labeled protein with a ratio of 100:1 unless otherwise stated.

### Preparation of RNA probes

For in vitro transcription, the pGEM vectors, each encoding a HOTAIRM1 fragment, was linearized with NotI (for Sp6 transcription) or SacII (for T7 transcription) (New England Lab) and used as a template. Radiolabeled HOTAIRM1 fragments were synthesized using



doi:10.6342/NTU202302064

T7 or Sp6 RNA polymerase (Promega) according to the manufacturer's instructions. In brief, each transcription reaction contained 0.5 mM ATP, CTP and GTP, 0.1 mM UTP, and 3.3 μM of $\alpha^{32}$P-UTP (3000 Ci/mmol, 10 mCi/ml, PerkinElmer). RNA was gel purified through a 14% polyacrylamide gel containing 7 M urea. The length of HOTAIRM1 fragments containing the flanking linker regions ranged from 214 to 289 nt. The specific radioactivity was $3 \times 10^6$ cpm/μg.

For end-labeling, RNA oligos are incubated with T4 PNK (New England Lab), 10x kinase buffer, and 9.9 μM of $\gamma^{32}$P-ATP (3000 Ci/mmol, 10 mCi/ml, PerkinElmer) for 1 hour. According to the manufacturer's instructions, 0-5 m6A oligos were first treated with alkaline phosphatase (Roche). Free nucleotides are removed using a microspin G-25 column (GE Healthcare). The specific radioactivity was $3 \times 10^6$ cpm/μg.

**Electrophoretic mobility shift assay (EMSA)**

For EMSA, $5\times10^4$ cpm of (0.78-6.25nM) $^{32}$P-labeled RNA oligos or (17–26 nM) $^{32}$P-labeled HOTAIRM1 fragment was incubated with Y14 in a buffer containing 50 mM HEPES (pH 7.0), 150 mM NaCl, 1 mM DTT, 10% glycerol, 0.5 mg/mL BSA on ice for 20 min. The reaction was mixed with the loading dye containing 2M KCl, 1M Tris (pH 7.6), 50% glycerol, 0.01% xylene cyanol, and 0.01% bromophenol blue before being separated by electrophoresis on a 4-10% polyacrylamide gel in Tris-Borate-EDTA buffer. Gels were vacuum-dried and imaged on a Typhoon scanner (GE Healthcare). Quantification of radiolabeled RNA was performed using ImageQuantTL software (GE Healthcare). The Kd value was determined from the concentration of which 50% of the probes were bound.



doi:10.6342/NTU202302064

**In-vitro transcription of methylated oligos**

0-5 m6A DNA oligos are first annealed to a T7 oligo to create a hybridized template. According to the manufacturer's instructions, methylated oligos were synthesized using T7 RNA polymerase (Promega). Each transcription reaction contained 0.1mM GTP, 0.1mM CTP, 0.1mM UTP, and 0.05mM $m^6$ATP. RNA was gel-purified through a 6% polyacrylamide gel containing 8 M urea. Purified RNA was loaded onto a 8% TBE-polyacrylamide gel, electrophoresed, and stained with GelRed. The length of methylated oligos was 42 nt.

**Phase separation assay**

Recombinant S-Y14-His proteins were mixed with buffers containing 20 mM MES/HEPES at different pHs. Recombinant non-tagged Y14 proteins were diluted in a storage buffer containing 50 mM HEPES and 300 mM KCl (pH 8.0) before being mixed with a buffer containing 20 mM HEPES (pH 6.7).

For protein/RNA co-condensation analysis, proteins and RNA were mixed before dilution into the LLPS reaction buffer. After dilution, samples were incubated for 2 hrs at room temperature and then placed onto the coverslip for visualization using a Leica microscope (DMI6000B) with a 100× oil objective under the differential interference contrast (DIC) or fluorescence mode (Cy5 or FITC/GFP). Representative images were captured after 5 or 10 minutes for S-Y14-His and Y14 proteins, respectively, unless otherwise stated.



doi:10.6342/NTU202302064

**Table 1: RNA used in this study (5' → 3')**

| | |
|---|---|
| HM1-1<br>(1-217) | **GAAUACUCAAGCUAGGCCGCACUAGUGAUU**AAAAGUUUGC<br>CGGCUUCCGCAGUGAUGGAUCACCGUUUUAGUGGCAUUUAA<br>AUCCCCGGCGCUCCGCCGUCUAGGUGACGCGCAGUCGCCCCC<br>CCAGGCAGCCUAGGCGGCGGCAGCUGCUGCGGCGACUGCAAA<br>GGCCGAUUUGGAGUGCUGGAGCGAAGAAGAGCAAAAGCUGC<br>GUUCUACGUCAACAGCCGACUCCGCUGCCCGCCCCGCCAGGC<br>CAAUCCCGCGG |
| HM1-2<br>(221-395) | **GAAUACUCAAGCUAGGCCGCACUAGUGAUU**GGGAGGUGGG<br>GCUGGGAGGCGUCCCCCGCUCCCGCCCCCUCCCCACCGUUCA<br>AUGAAAGAUGAACUGGCGAGAGGACGAAUCGCAUCCAGGAG<br>CUGCGCAGCCCUGGCCGCUGCCGGGACGCCCUGCUCCGCGCU<br>GAGCUUGGGGCCAGAAACCAGCCAUAGUCCCCACACUCCAAU<br>CCCGCGG |
| HM1-3<br>(396-631) | **GGGCGAAUUGGGGAUU**GCCGCCGCAGCUGAGAUUUAGCGG<br>AGGAAGGGGCGAGGGAAGGUAGGGAGCAAACCUAUGAAGAA<br>ACAUCGCGUUGUCAUUGGAACUUCCAAGCCUUUGCUGUUAA<br>GAGCCAGGUUCUUAAAUCAACCCGCCCCACACACAUGUUGCU<br>UACAUGCUGCGUUUUCUCACGGUCUGUUUUGCCUGAACCCA<br>UCAACAGCUGGGAGAUUAAUCAACCACACUGAAAAUGUGGA<br>GGGAUUAAUCACUAGUGCGGCCGC |
| HM1-4<br>(632-882) | **GAAUACUCAAGCUACCGCACUAGUGAUU**UAUGGGGGAGGG<br>GGUUGAAAUGUGGGUGUUUGAAACAAAAGUGUAUAAACAAA<br>UGAAUUGUUGAUAACUUAGUUAUUGACCUGGAGACUGGUAG<br>CUUAUUAAAGAAACUCCGUGUUACUCAUUCCUGGAGUUGGG |



doi:10.6342/NTU202302064

| | |
|---|---|
| | GGUUUCUGUAGGCACUUUAUUUCUCCACUUUCAAGAGCUUG GGCUUGGCCCAAAUCUUAGACUGUCCAAUUCUGCCUCUAUU ACCAAUUUAAAUCUAUGGCUUGAACCUGUGCACUAAUCCCG CGG |
| HM1-5 (882-1044) | **GGGCGAAUUGGGCCC**UGAAAAUCAAAUCCUUUAAAAAGAA AGAGGAGAAUAAGAAGCAAAAAAGAAAGAAAAAACACUUAU UAGAAGCCCUAGUCAUUUUUUGGCUUUCUGUUUUGUUGCUG UCCAUUGAAGACUUUGAACAUGCCGCCUUAAUAAAUGUAUU AAAAUUGAAAAAAGAGCGGCCGC |
| HM1-m$^6$A | CCCAAAUUUAG**m$^6$A**CUGUCCAAUUCU |
| HM1-A | CCCAAAUUUAGACUGUCCAAUUCU |
| m$^6$A5 | UUC**m$^6$A**GGUUUUC**m$^6$A**GGUUUUC**m$^6$A**GGUUUUC**m$^6$A**GGUUUUC **m$^6$A**GGUU |
| m$^6$A3 | UUC**m$^6$A**GGUUUUCUGGUUUUC**m$^6$A**GGUUUUCUGGUUUUC**m$^6$A** GGUU |
| m$^6$A1 | UUCUGGUUUUCUGGUUUUC**m$^6$A**GGUUUUCUGGUUUUCUGGU U |
| m$^6$A0 | UUCUGGUUUUCUGGUUUUCUGGUUUUCUGGUUUUCUGGUU |

(Restriction enzyme sites are underlined while m6A and transcription sites are shown in bold)



doi:10.6342/NTU202302064

## References


Abakir, A., Giles, T. C., Cristini, A., Foster, J. M., Dai, N., Starczak, M., Li, M., Eleftheriou, M., Crutchley, J., Flatt, L., Young, L., Gaffney, D. J., Denning, C., Dalhus, B., Emes, R. D., Gackowski, D., Corrêa, I. R., Garcia -Perez, L., J., Klungland, A., Gromak, N., Ruzov, A. (2020). N6-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. *Nature Genetics*, *52*(1), 48-55. https://doi.org/10.1038/s41588-019-0549-x

Alberti, S., Gladfelter, A. S., & Mittag, T. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell*, *176*(3), 419–434. https://doi.org/10.1016/j.cell.2018.12.035

André, A. A. M., and Spruijt, E. (2020). Liquid -liquid phase separation in crowded environments. *International Journal of Molecular Sciences*, *21*(16), 5908–5908. https://doi.org/10.3390/ijms21165908

Banani, S. F., Lee, H., Hyman, A. A., & Rosen, M. J. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology*, *18*(5), 285–298. https://doi.org/10.1038/nrm.2017.7

Bekker-Jensen, S., Lukas, C., Kitagawa, R., Melander, F., Kastan, M. B., Bartek, J., & Lukas, J. (2006). Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *Journal of Cell Biology*, *173*(2), 195–206. https://doi.org/10.1083/jcb.200510130

Boehm, V., & Gehring, N. H. (2016). Exon Junction Complexes: Supervising the Gene Expression Assembly Line. *Trends in Genetics*, *32*(11), 724–735. https://doi.org/10.1016/j.tig.2016.09.003

Chuang, T.W., Lee, K.M., Lou, Y.C., Lu, C.C., & Tarn, W.Y. (2016). A Point Mutation in the Exon Junction Complex Factor Y14 Disrupts Its Function in mRNA Cap




doi:10.6342/NTU202302064

Binding and Translation Enhancement. *Journal of Biological Chemistry*, *291*(16), 8565-8574. https://doi.org/10.1074/jbc.M115.704544

Chuang, T.W., Lu, C.C., Su, C.H., Wu, P.Y., Sarasvathi E., Lee, C.C., Kuo, H.C., Hung, K.Y., Lee, K.M., Tsai, C.Y., & Tarn, W.Y. (2019). The RNA Processing Factor Y14 Participates in DNA Damage Response and Repair. *IScience, 13*, 402–415. https://doi.org/10.1016/j.isci.2019.03.005

Chen, K., Lu, Z., Wang, X., Fu, Y., Luo, Z., Liu, N., Han, D., Dominissini, D., Dai, Q., Pan, T., & He, C. (2015). High-Resolution N6-Methyladenosine (m6A) Map Using Photo-Crosslinking-Assisted m6A Sequencing. *Angewandte Chemie (International ed. in English)*, *54*(5), 1587. https://doi.org/10.1002/anie.201410647

Chuang, T.W., Su, C.H., Wu, P.Y., Chang, Y.M., & Tarn, W.Y. (2023). LncRNA *HOTAIRM1* functions in DNA double-strand break repair via its association with DNA repair and mRNA surveillance factors. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkad143

Ciccia, A., & Elledge, S. J. (2010). The DNA Damage Response: Making It Safe to Play with Knives. *Molecular Cell*, *40*(2), 179–204. https://doi.org/10.1016/j.molcel.2010.09.019

Diaz-Lagares, A., Crujeiras, A. B., Lopez-Serra, P., Soler, M., Setien, F., Goyal, A., Sandoval, J., Hashimoto, Y., Martínez -Cardús, A., Gomez, A., Heyn, H., Moutinho, C., Espada, J., Vidal, A., Paules M., Galán M., Sala, N., Akiyama, Y., Martínez -Iniesta M., & Farre, L. (2016) Epigenetic inactivation of the p53-induced long noncoding RNA TP53 target 1 in human cancer. *Proceedings of the National Academy of Sciences, 113*(47), E7535-E7544. https://doi.org/10.1073/pnas.1608585113



doi:10.6342/NTU202302064

D'Alessandro, G., & d'Adda di Fagagna, F. (2017). Transcription and DNA Damage: Holding Hands or Crossing Swords? *Journal of Molecular Biology*, *429*(21), 3215–3229. https://doi.org/10.1016/j.jmb.2016.11.002

Espinosa, J. R., Joseph, J. A., Sanchez-Burgos, I., Adiran Garaizar, Frenkel, D., & Collepardo-Guevara, R. (2020). Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(24), 13238–13247. https://doi.org/10.1073/pnas.1917569117

Fay, M. M., & Anderson, P. J. (2018). The Role of RNA in Biological Phase Separations. *Journal of Molecular Biology*, *430*(23), 4685-4701. https://doi.org/10.1016/j.jmb.2018.05.003

Fukumura, K., Wakabayashi, S.-I., Kataoka, N., Sakamoto, H., Suzuki, Y., Nakai, K., Mayeda, A., & Inoue, K. (2016). The Exon Junction Complex Controls the Efficient and Faithful Splicing of a Subset of Transcripts Involved in Mitotic Cell-Cycle Progression. *International Journal of Molecular Sciences*, *17*(8), 1153–1153. https://doi.org/10.3390/ijms17081153

Fung, H. Y. J., Birol, M., & Rhoades, E. (2018). IDPs in macromolecular complexes: the roles of multivalent interactions in diverse assemblies. *Current Opinion in Structural Biology*, *49*, 36–43. https://doi.org/10.1016/j.sbi.2017.12.007

Haemmig, S., Yang, D., Sun, X., Das, D., Ghaffari, S., Molinaro, R., Chen, L., Deng, Y., Freeman, D., Moullan, N., Tesmenitsky, Y., Khyrul Wara, K. M., Simion, V., Shvartz, E., Lee, J. F., Yang, T., Sukova, G., Marto, J. A., Stone, P. H., Lee, W. L., Auwerx J., Libby P., Feinberg, M. W. (2020). Long noncoding RNA SNHG12 integrates a DNA-PK–mediated DNA damage response and vascular senescence. *Science Translational Medicine*. https://doi.org/aaw1868



doi:10.6342/NTU202302064

Hawley, B. R., Lu, W., Wilczynska, A., & Bushell, M. (2017). The emerging role of RNAs in DNA damage repair. *Cell Death & Differentiation*, *24*(4), 580-587. https://doi.org/10.1038/cdd.2017.16

Hingorani, K., Attila Szebeni, & Olson, M. E. (2000). *Mapping the Functional Domains of Nucleolar Protein B23*. *275*(32), 24451–24457. https://doi.org/10.1074/jbc.m003278200

Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O., & Pappu, R. V. (2017). CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophysical Journal*, *112*(1), 16-21. https://doi.org/10.1016/j.bpj.2016.11.3200

Hyman, A. A., Weber, C. A. & Jülicher, F. (2014) Liquid-Liquid Phase Separation in Biology *Annual Review of Cell and Developmental Biology 30*, 39–58. https://doi.org/10.1146/annurev-cellbio-100913-013325

Kai, M. (2016). Roles of RNA-Binding Proteins in DNA Damage Response. *International Journal of Molecular Sciences*, *17*(3). https://doi.org/10.3390/ijms17030310

Langdon, E. M., Qiu, Y., Niaki, A. G., McLaughlin, G. A., Weidmann, C. A., Gerbich, T. M., Smith, J. A., Crutchley, J. M., Termini, C. M., Weeks, K. M., Myong, S., & Gladfelter, A. S. (2018). mRNA structure determines specificity of a polyQ-driven phase separation. *Science*, *360*(6391), 922-927 https://doi.org/10.1126/science.aar7432

Lee, H. G., Kim, J., & Seo, P. J. (2022). N6-methyladenosine–modified RNA acts as a molecular glue that drives liquid–liquid phase separation in plants. *Plant Signaling & Behavior*, *17*(1). https://doi.org/10.1080/15592324.2022.2079308



doi:10.6342/NTU202302064

Levone, B. R., Lenzken, S. C., Antonaci, M., Maiser, A., Rapp, A., Conte, F., Reber, S., Mechtersheimer, J., Ronchi, A. E., Mühlemann, O., Leonhardt, H., Cardoso, M. C., Ruepp, M.-D., & Barabino, S. M. L. (2021). FUS-dependent liquid–liquid phase separation is important for DNA repair initiation. *Journal of Cell Biology*, *220*(5). https://doi.org/10.1083/jcb.202008030

Li, Q., Peng, X., Li, Y., Tang, W., Zhu, J., Huang, J., Qi, Y., & Zhang, Z. (2020). LLPSDB: A database of proteins undergoing liquid–liquid phase separation in vitro. *Nucleic Acids Research*, *48*(D1), D320-D327. https://doi.org/10.1093/nar/gkz778

Lin, Y., Protter, D. S., Rosen, M. K., & Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell*, *60*(2), 208-219. https://doi.org/10.1016/j.molcel.2015.08.018

Lisby, M., Barlow, J. H., Burgess, R. C., & Rothstein, R. (2004). Choreography of the DNA Damage Response. *Cell*, *118*(6), 699–713. https://doi.org/10.1016/j.cell.2004.08.015

Lu, C.-C., Lee, C.-C., Tseng, C.-T., & Tarn, W.-Y. (2017). Y14 governs p53 expression and modulates DNA damage sensitivity. *Scientific Reports*, *7*(1). https://doi.org/10.1038/srep45558

Maharana, S., Wang, J., Papadopoulos, D. K., Richter, D., Pozniakovsky, A., Poser, I., Bickle, M., Rizk, S., Guillén -Boixet, J., Franzmann, T. M., Jahnel, M., Marrone, L., Chang, T., Sterneckert, J., Tomancak, P., Hyman, A. A., & Alberti, S. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science, 360*, 918-921. https://doi.org/10.1126/science.aar7366



doi:10.6342/NTU202302064

Meng, X., Yang, S., & Camp, V. J. A. (2020). The Interplay Between the DNA Damage Response, RNA Processing and Extracellular Vesicles. *Frontiers in Oncology*, *9*. https://doi.org/10.3389/fonc.2019.01538

Michelle, L., Cloutier, A., Toutant, J., Lulzim Shkreta, Thibault, P., Durand, M., Garneau, D., Gendron, D., Lapointe, E., Couture, S., Hervé Le Hir, Klinck, R., Sherif Abou Elela, Panagiotis Prinos, & Chabot, B. (2012). Proteins Associated with the Exon Junction Complex Also Control the Alternative Splicing of Apoptotic Regulators. *Molecular and Cellular Biology*, *32*(5), 954–967. https://doi.org/10.1128/mcb.06130-11

Michelini, F., Pitchiaya, S., Vitelli, V., Sharma, S., Gioia, U., Pessina, F., Cabrini, M., Wang, Y., Capozzo, I., Iannelli, F., Matti, V., Francia, S., Shivashankar, G. V., & Walter, N. G. (2017). Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. *Nature cell biology*, *19*(12), 1400. https://doi.org/10.1038/ncb3643

Navarro, M. G., Kashida, S., Chouaib, R., Souquere, S., Pierron, G., Weil, D., & Gueroui, Z. (2019). RNA is a critical element for the sizing and the composition of phase-separated RNA–protein condensates. *Nature Communications*, *10*(1), 1-13. https://doi.org/10.1038/s41467-019-11241-6

Qu, F., Tsegay, P. S., & Liu, Y. (2021). N6-Methyladenosine, DNA Repair, and Genome Stability. *Frontiers in Molecular Biosciences*, *8*, 645823. https://doi.org/10.3389/fmolb.2021.645823

Ray Chaudhuri, A., & Nussenzweig, A. (2017). The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. *Nature Reviews Molecular Cell Biology*, *18*(10), 610-621. https://doi.org/10.1038/nrm.2017.53



doi:10.6342/NTU202302064

Sharma, V., Khurana, S., Kubben, N., Abdelmohsen, K., Oberdoerffer, P., Gorospe, M., & Misteli, T. (2015). A BRCA1-interacting lncRNA regulates homologous recombination. *EMBO Reports*, *16*(11), 1520-1534. https://doi.org/10.15252/embr.201540437

Sukhanova, M. V., Singatulina, A. S., Pastré, D., & Lavrik, O. I. (2020). Fused in Sarcoma (FUS) in DNA Repair: Tango with Poly(ADP-ribose) Polymerase 1 and Compartmentalisation of Damaged DNA. *International Journal of Molecular Sciences*, *21*(19). https://doi.org/10.3390/ijms21197020

Wang, J., Choi, J. M., Holehouse, A. S., Lee, H. O., Zhang, X., Jahnel, M., et al. (2018). A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell 174*, 688–699. https://doi.org/10.1016/j.cell.2018.06.006

Wang, J., Wang, L., Diao, J., Shi, Y. G., Shi, Y., Ma, H., & Shen, H. (2020). Binding to m6A RNA promotes YTHDF2-mediated phase separation. *Protein & Cell*, *11*(4), 304-307. https://doi.org/10.1007/s13238-019-00660-2

Whelan, D. R., Howard, S. M., Vitelli, V., Renaudin, X., Adamowicz, M., Iannelli, F., Winston, C., Lee, M., Matti, V., Lee, W. T., Morten, M. J., Venkitaraman, A. R., Cejka, P., & Rothenberg, E. (2018). BRCA2 controls DNA:RNA hybrid level at DSBs by mediating RNase H2 recruitment. *Nature Communications*, *9*(1), 1-17. https://doi.org/10.1038/s41467-018-07799-2

Van Treeck, B., Protter, D. S., Matheny, T., Khong, A., Link, C. D., & Parker, R. (2018). RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proceedings of the National Academy of Sciences*, *115*(11), 2734-2739. https://doi.org/10.1073/pnas.1800038115



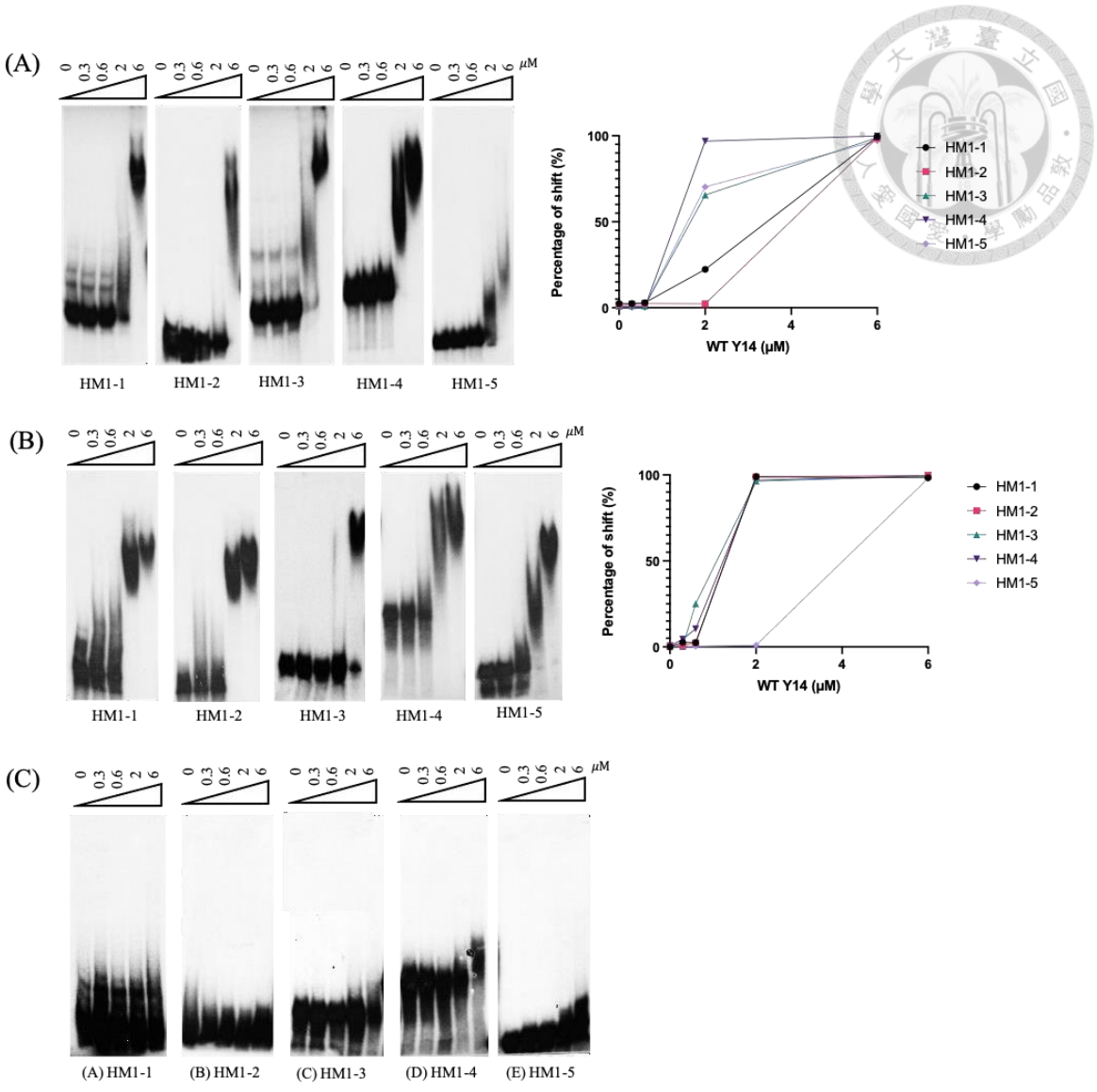
doi:10.6342/NTU202302064

Vohhodina, J., Barros, E. M., Savage, A. L., Liberante, F. G., Manti, L., Bankhead, P., Cosgrove, N., Madden, A. F., Harkin, D. P., & Savage, K. I. (2017). The RNA processing factors THRAP3 and BCLAF1 promote the DNA damage response through selective mRNA splicing and nuclear export. *Nucleic Acids Research*, *45*(22), 12816-12833. https://doi.org/10.1093/nar/gkx1046

Yu, C.L., Chuang, T.W., Samuel, S. Y., Lou, Y.C., & Tarn, W.Y. (2023). Co-phase separation of Y14 and RNA in vitro and its implication for DNA repair. *RNA*, rna.079514.122. https://doi.org/10.1261/rna.079514.122

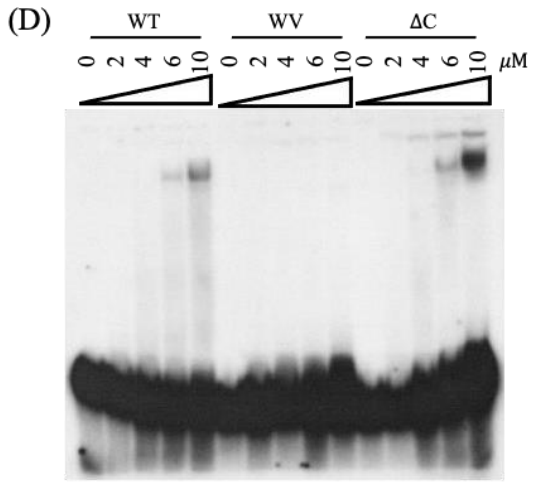


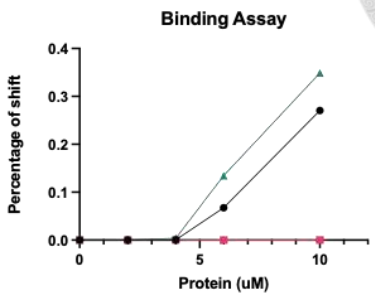
doi:10.6342/NTU202302064

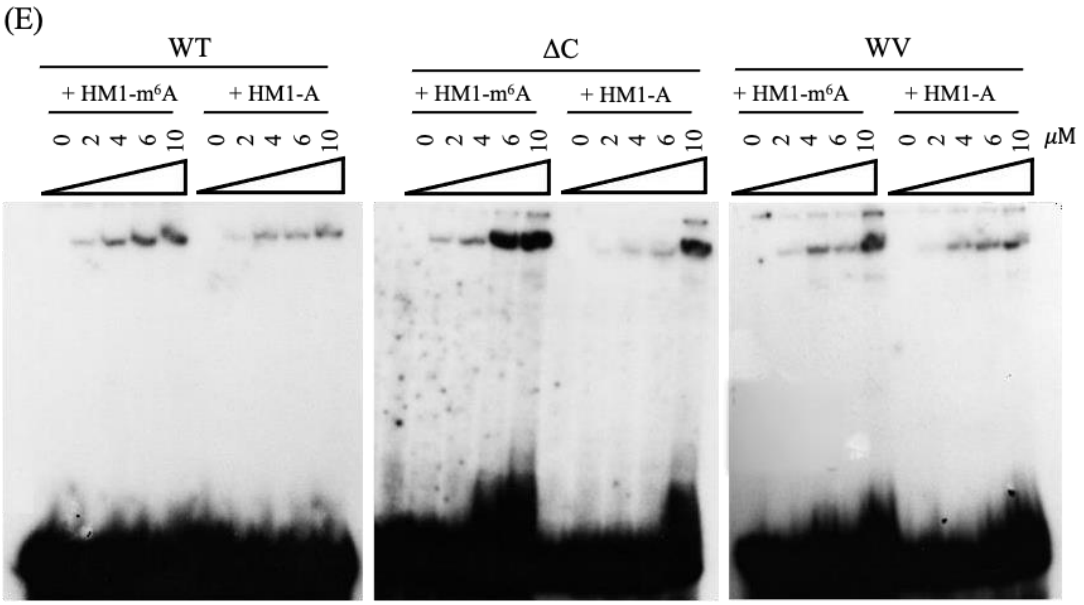


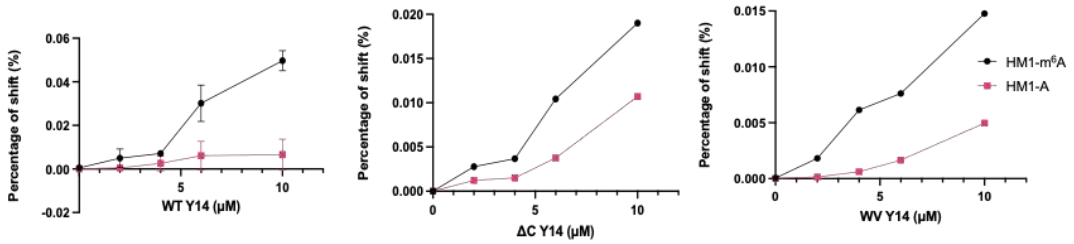



doi:10.6342/NTU202302064






doi:10.6342/NTU202302064

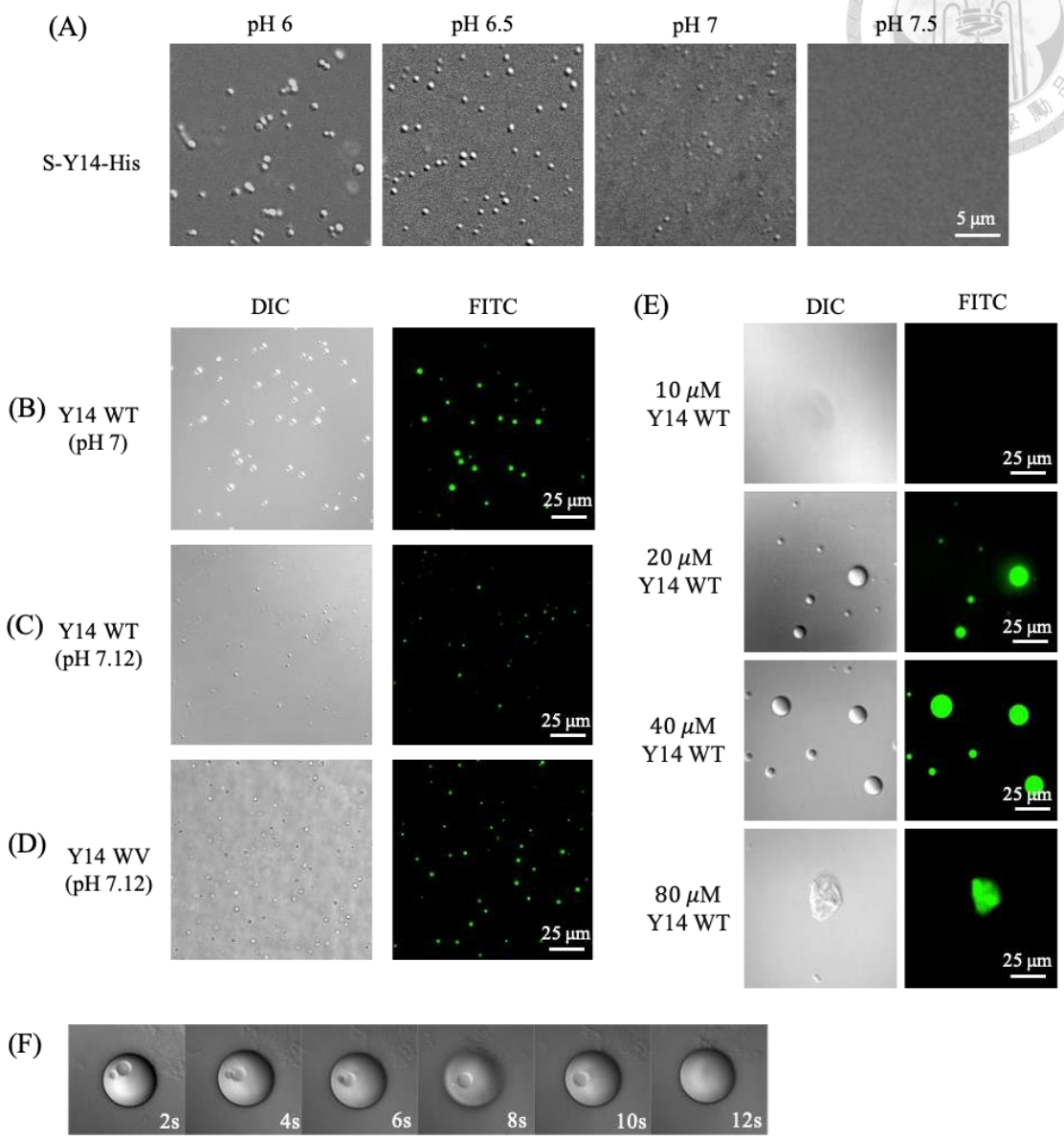
**Fig 1. RNA binding of Recombinant Human Y14 Protein.**

EMSA shows all 5 HOTAIRM1 (HM1) fragments bind to S-Y14-His WT and ΔC mutant, but not to WV mutant. Increasing concentration of recombinant S-Y14-His (A) wild-type, (B) ΔC mutant and (C) WV mutant was incubated with $^{32}$P-labeled HM1-1 (19 nM), HM1-2 (22 nM), HM1-3 (18 nM), HM1-4 (nM), and HM1-5 (26 nM) fragments. (D) EMSA shows U25 oligo binds S-Y14-His WT and ΔC mutant, but not WV mutant. Increasing concentration of recombinant S-Y14-His wild-type, Δ C mutant and WV mutant was incubated with 0.78 nM of $^{32}$P-labeled $U_{25}$ oligo. (E) EMSA shows HM1-m$^6$A oligo binds better to all S-Y14-His proteins compared to HM1-A oligo. An increasing concentration of recombinant S-Y14-His WT was incubated with 6.25 nM of $^{32}$P-labeled HM1-m$^6$A and HM1-A oligo, respectively.

Percentage of RNA probes shifted is shown graphically. (N = 2)



doi:10.6342/NTU202302064






doi:10.6342/NTU202302064

**Fig 2. Characterization of Y14 LLPS.**

(A) DIC images of phase separation of S-Y14-His WT (20$\mu$M) in four different pH level buffers. The representative images (200x200 pixels) show droplets after 2h of incubation. A mixture of ALEXA Fluor 488-conjugated (0.2 $\mu$M) and unlabeled wild-type Y14 (20$\mu$M) was incubated in a (B) pH 7 and (C) pH 7.12 buffer. (D) A mixture of ALEXA Fluor 488-conjugated (0.2 $\mu$M) and unlabeled Y14 WV mutant (20$\mu$M) was incubated in a pH 7.12 buffer. (E) A mixture of ALEXA Fluor 488-conjugated and increasing concentrations of unlabeled wild-type Y14 at the molar ratio of 1:100 was incubated in a pH 7 buffer. (G) DIC images showing the droplets formation of wild-type Y14 over a time period to show the disappearance of multiphase droplets.

The representative images show condensates after 2h of incubation.



doi:10.6342/NTU202302064

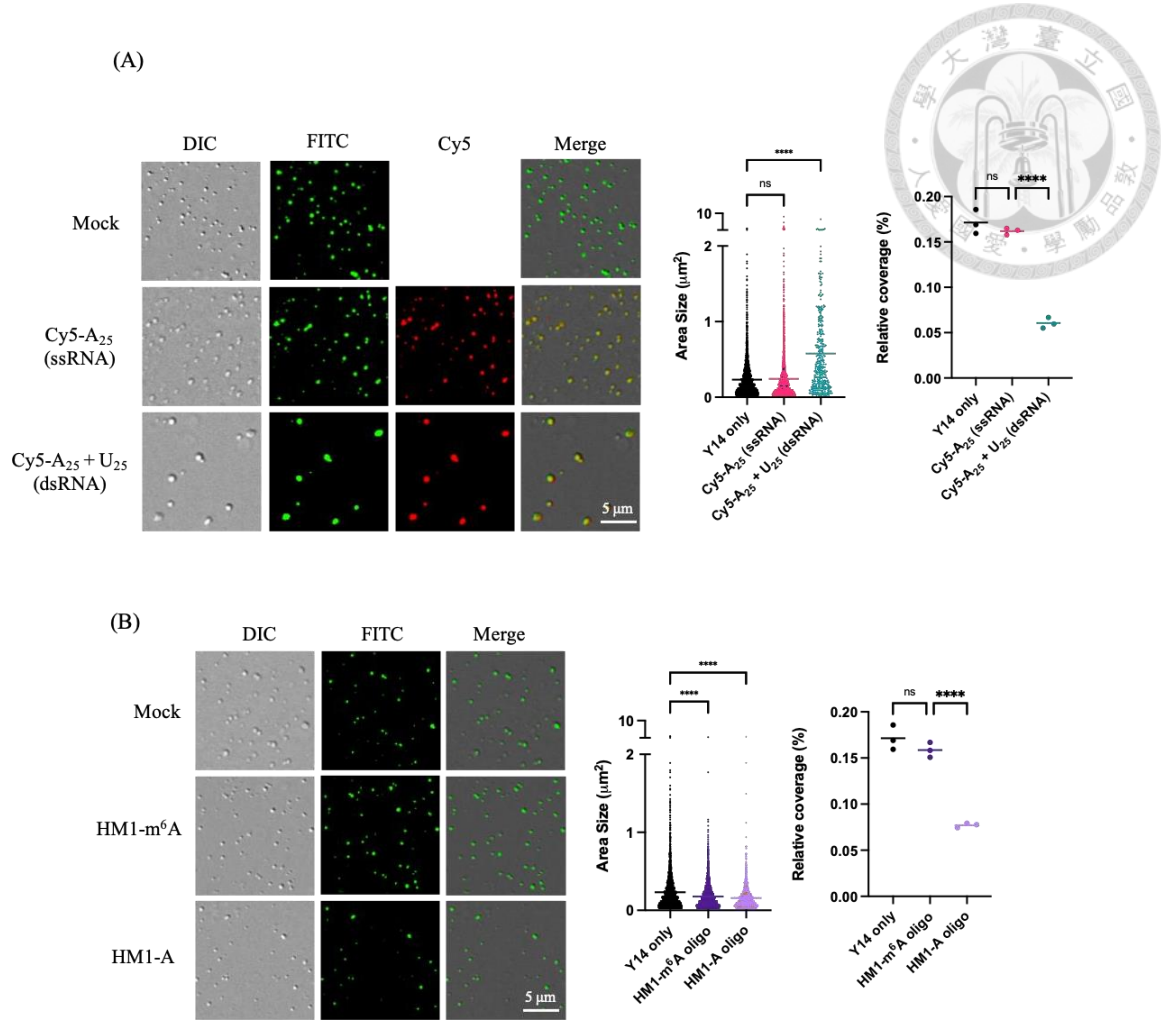


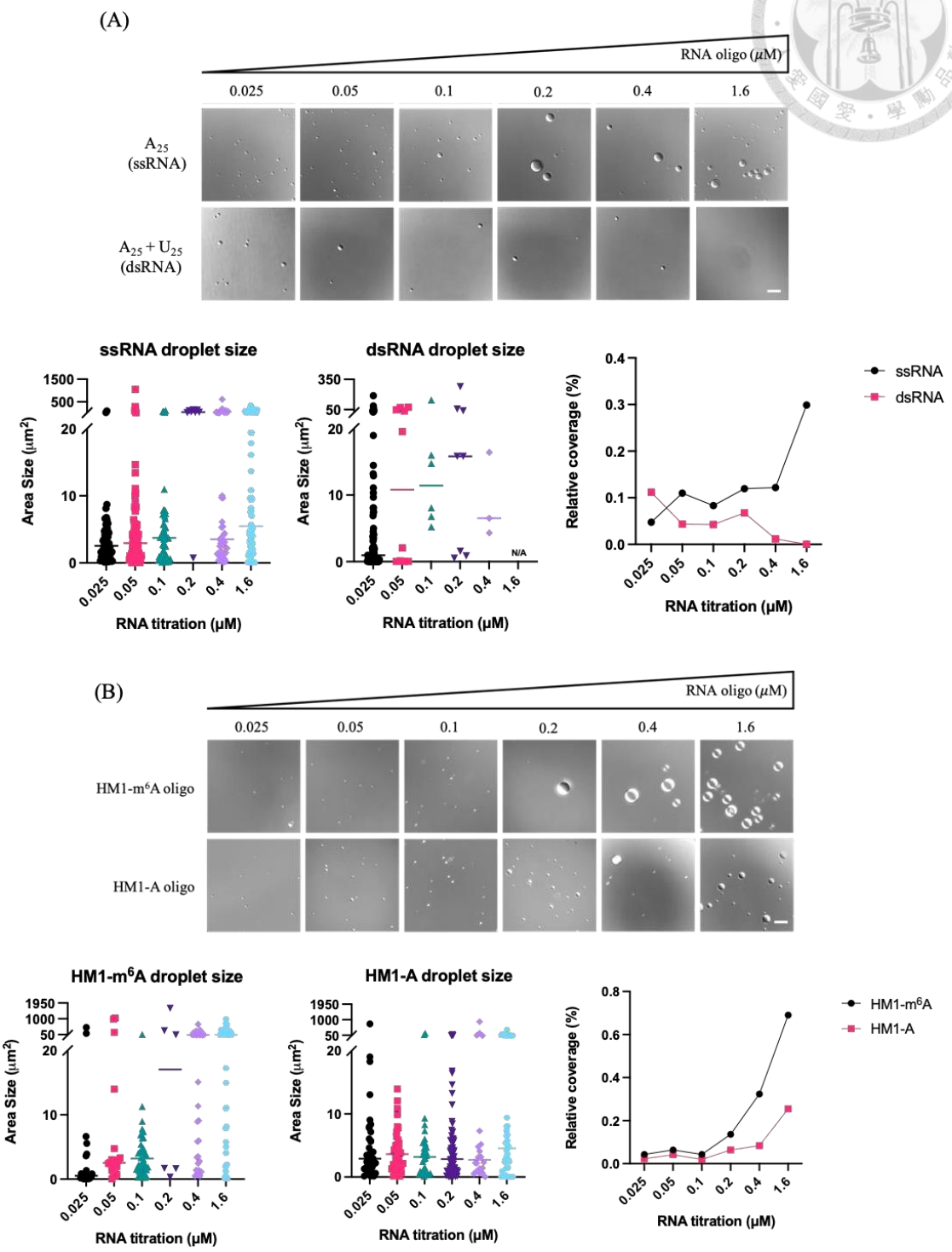

**Fig 3. Different RNA modulates S-Y14-His LLPS differently.**

(A) A mixture of ALEXA Fluor 488-conjugated (0.2 $\mu$M) and unlabeled S-Y14-His WT (20 $\mu$M) was mock-incubated or incubated with 0.2 $\mu$M of ssRNA $U_{25}$ or dsRNA ($U_{25}$+$A_{25}$) in a pH 6.5 buffer. (B) A mixture of ALEXA Fluor 488-conjugated (0.2 $\mu$M) and unlabeled S-Y14-His WT (20 $\mu$M) was mock-incubated or incubated with 0.2 $\mu$M of HM1-$m^6$A oligo or HM1-A oligo in a pH 6.5 buffer.

Mean size and relative coverage rate of the condensates are shown in the dot graph. The representative images (200x200 pixels) show condensates after 2h of incubation. Bar, 5 μm. (N = 3, *p*-value <0.0001)



doi:10.6342/NTU202302064






doi:10.6342/NTU202302064

**Fig 4. Different RNA modulates Y14 LLPS differently.**

(A) A mixture of ALEXA Fluor 488-conjugated (0.2 $\mu$M) and unlabeled Y14 WT (20 $\mu$M) was mock-incubated or incubated with 0.2 $\mu$M of ssRNA $A_{25}$ or dsRNA ($U_{25}$+$A_{25}$) in a pH 7 buffer. (B) A mixture of ALEXA Fluor 488-conjugated (0.2 $\mu$M) and unlabeled Y14 WT (20 $\mu$M) was mock-incubated or incubated with increasing concentrations of HM1-A or HM1-$m^6$A oligo in a pH 7 buffer.

Mean size and relative coverage rate of the condensates are shown in the dot graph. The representative images (1004x1002 pixels) show condensates after 2h of incubation. Bar, 25 μm. (N = 3)



doi:10.6342/NTU202302064

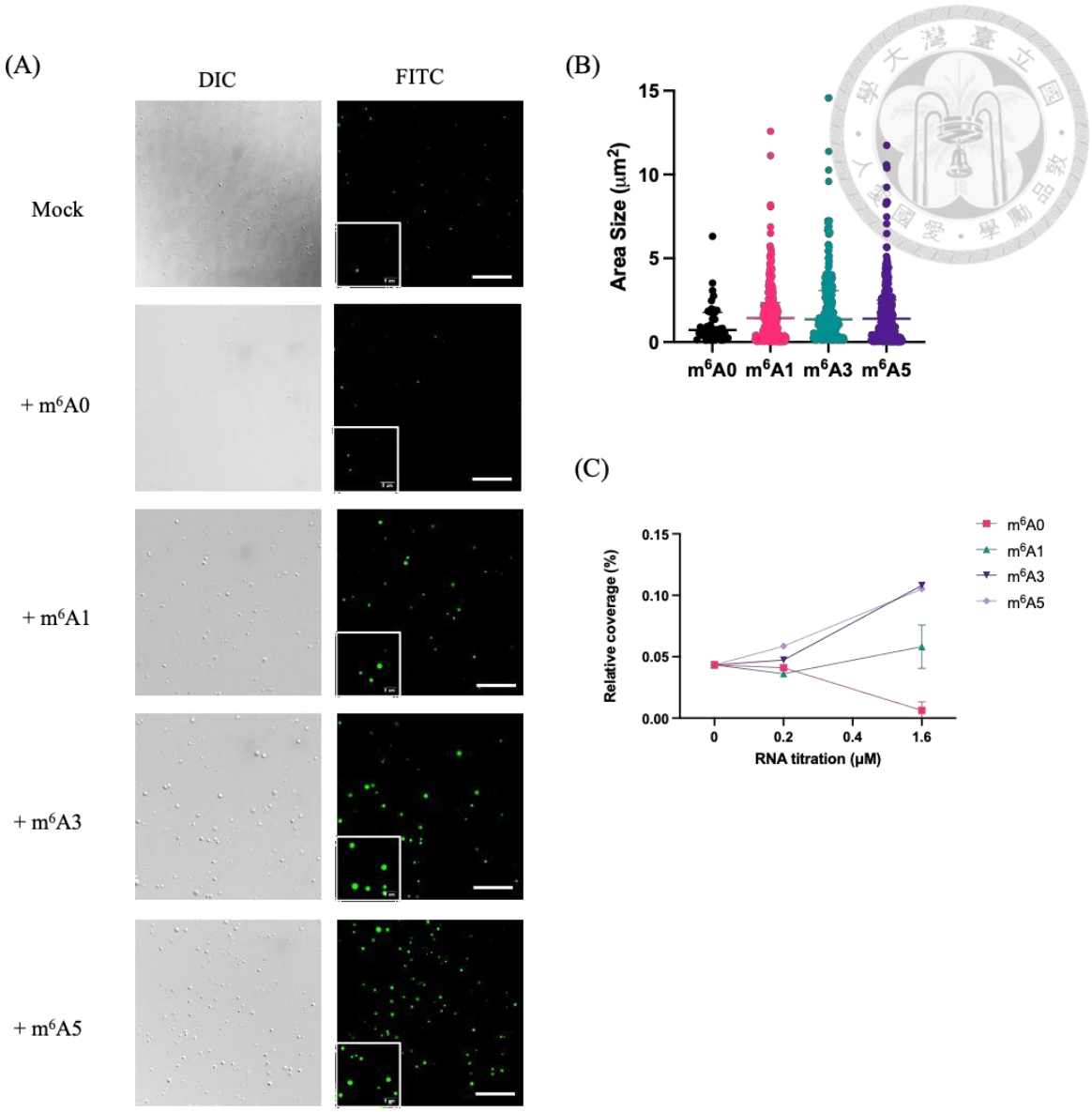



**Fig 5. RNA Methylation promotes Y14 LLPS.**

(A) A mixture of ALEXA Fluor 488-conjugated (0.2 $\mu$M) and unlabeled Y14 WT (20 $\mu$M) was mock-incubated or incubated with 1.6 $\mu$M of m$^6$A0, m$^6$A1, m$^6$A3, and m$^6$A5 oligo at pH 7.12 buffer. The representative images show condensates after 2h of incubation. Bar, 25 μm. Inset shows magnified images (200x200 pixels) of a selected area. Mean size (B) and relative coverage rate (C) of the droplets are shown in the dot graph.



doi:10.6342/NTU202302064

(A)

LLPSDB v1.0

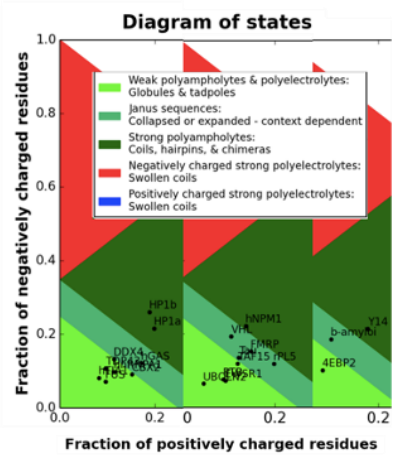
↓

Proteins that undergo LLPS: **62** proteins

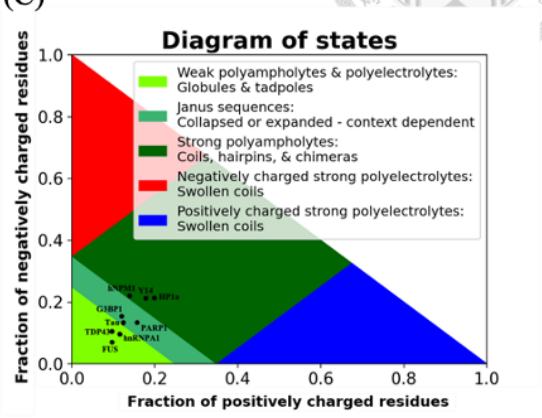
↓

Homo Sapiens/Natural proteins: **20** proteins

(B)




(C)



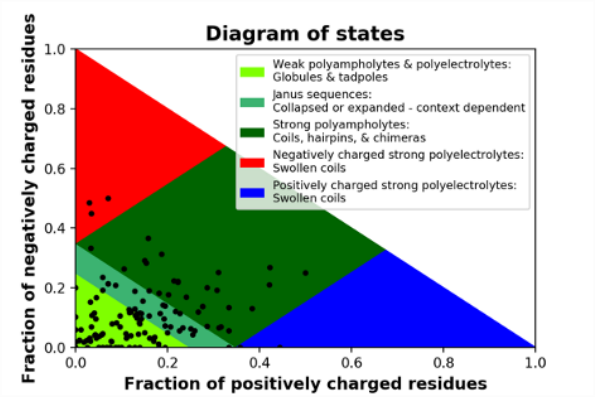

(D)

| | Protein number | IDP/IDR number |
|---|---|---|
| **MobiDB** | 142 | - |
| **PhasePro** | 121 | - |
| **DisProt** | 1042 | 6576 |
| **LLPSDB** | 192 | - |
| **IUPred2** | 793 | 1195 |
| **TOTAL** | 1141 | 2012 |

↓

**60** RRM-containing protein
**127** IDRs

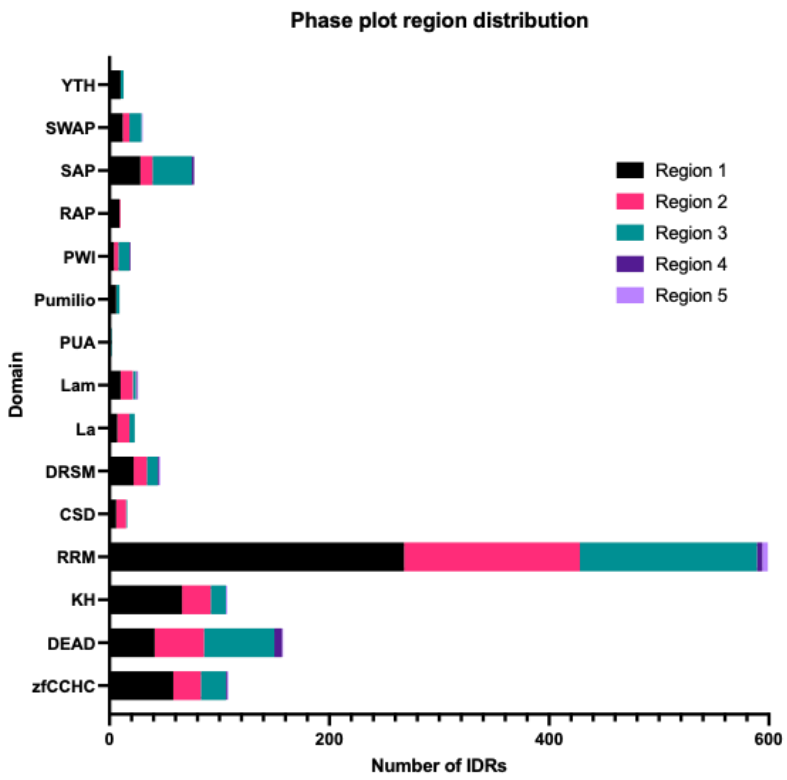
(E)






doi:10.6342/NTU202302064

(F)

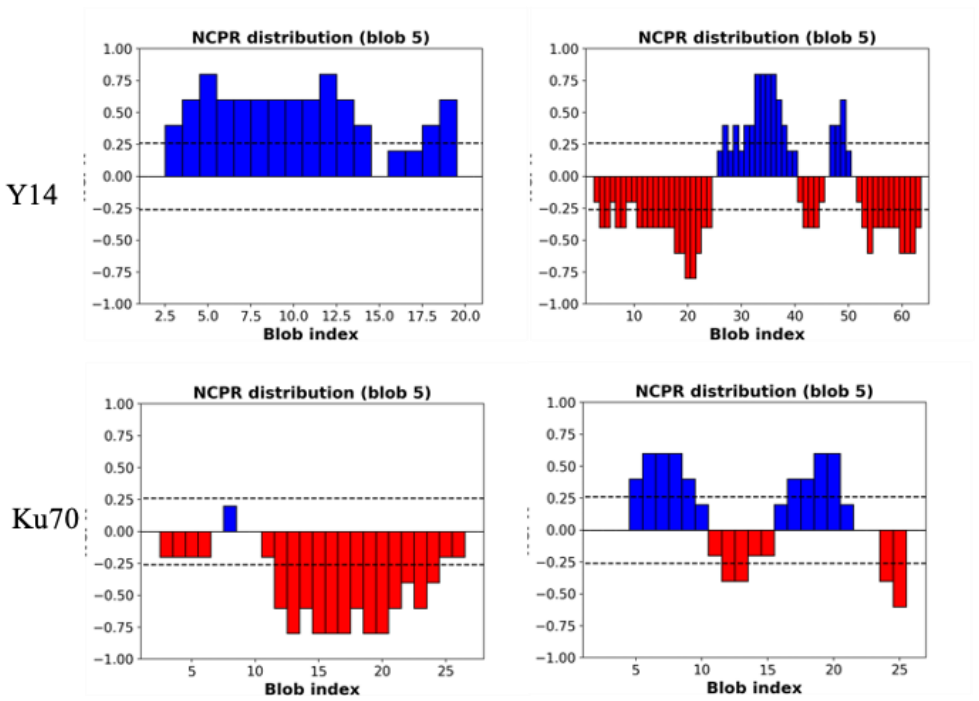



(G) Literature research + DisProt

↓

NHEJ proteins:
**23** proteins, **41** IDRs

(H)





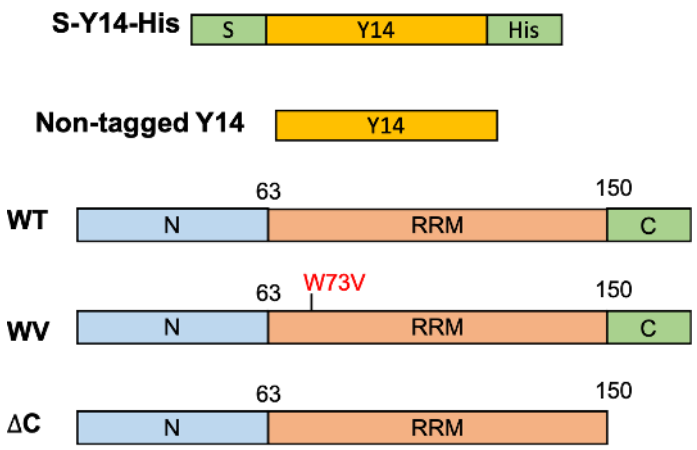

doi:10.6342/NTU202302064

**Fig 6. Bioinformatics analysis of charged IDRs in RNA processing and DNA repair factors.**

(A) Identification of the proteins that have LLPS properties from the LLPSDB database. (B) Diagram shows the charge states of known LLPS proteins. (C) Diagram of the charge states of Y14 and disease-associated LLPS proteins. (D) Identification of RRM-containing IDRs and their (E) charge states. (F) Diagram of the charge state distribution of LLPS proteins with various domains. (G) Identification of NHEJ-related proteins and the (H) NCPR distribution of Y14 and Ku70 IDRs. NPCR represents Net Charge Per Residue.



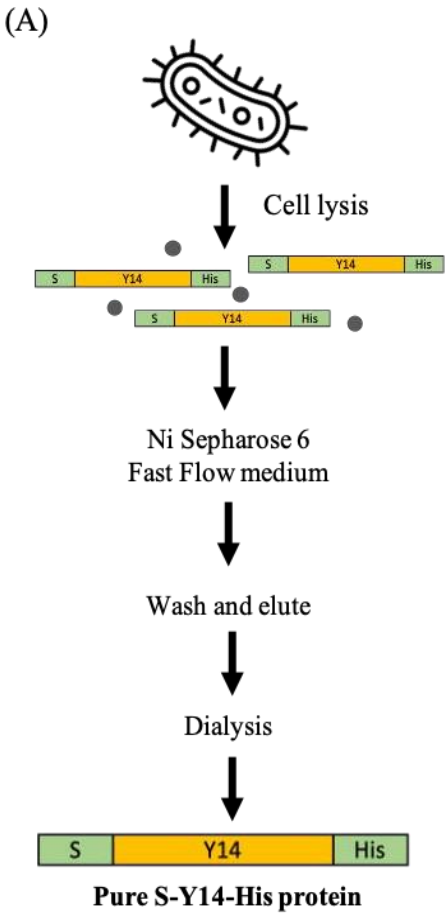
doi:10.6342/NTU202302064

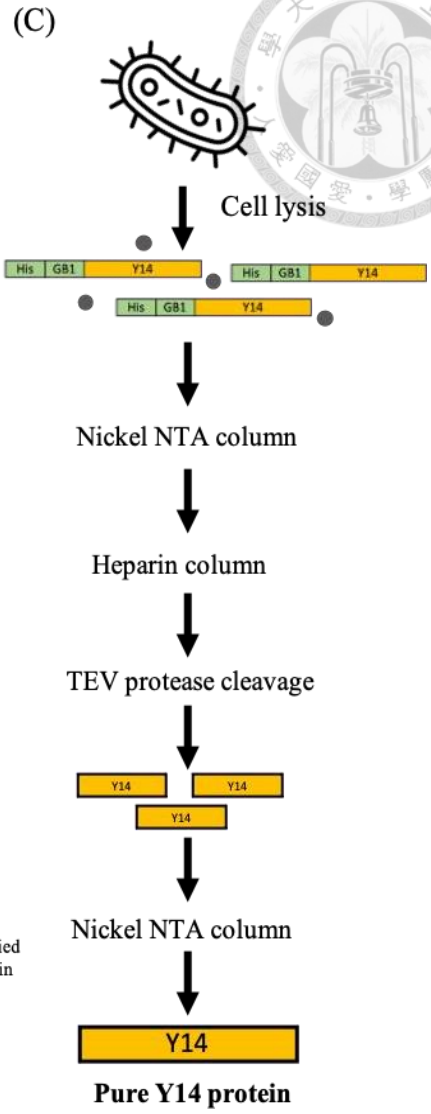


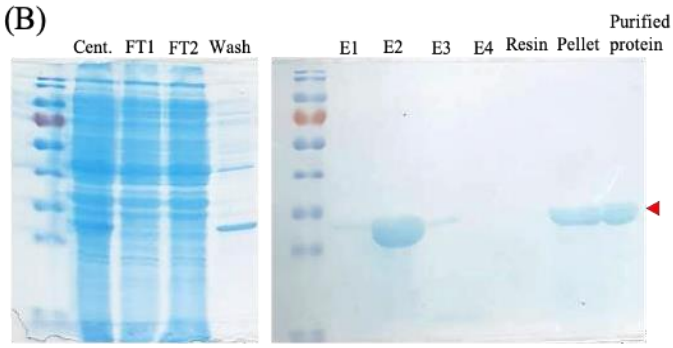

**Appendix 1. Illustration showing recombinant S-Y14-His and non-tagged Y14, and domains of wild-type and mutant (WV and ΔC) Y14.**

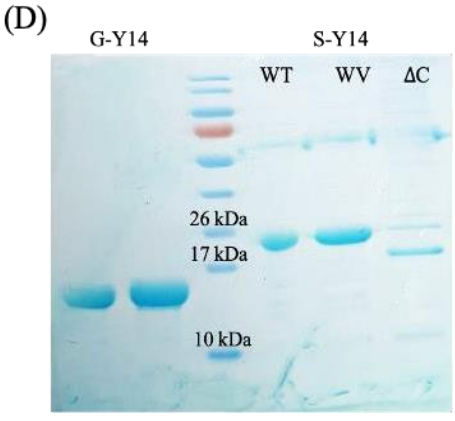


doi:10.6342/NTU202302064





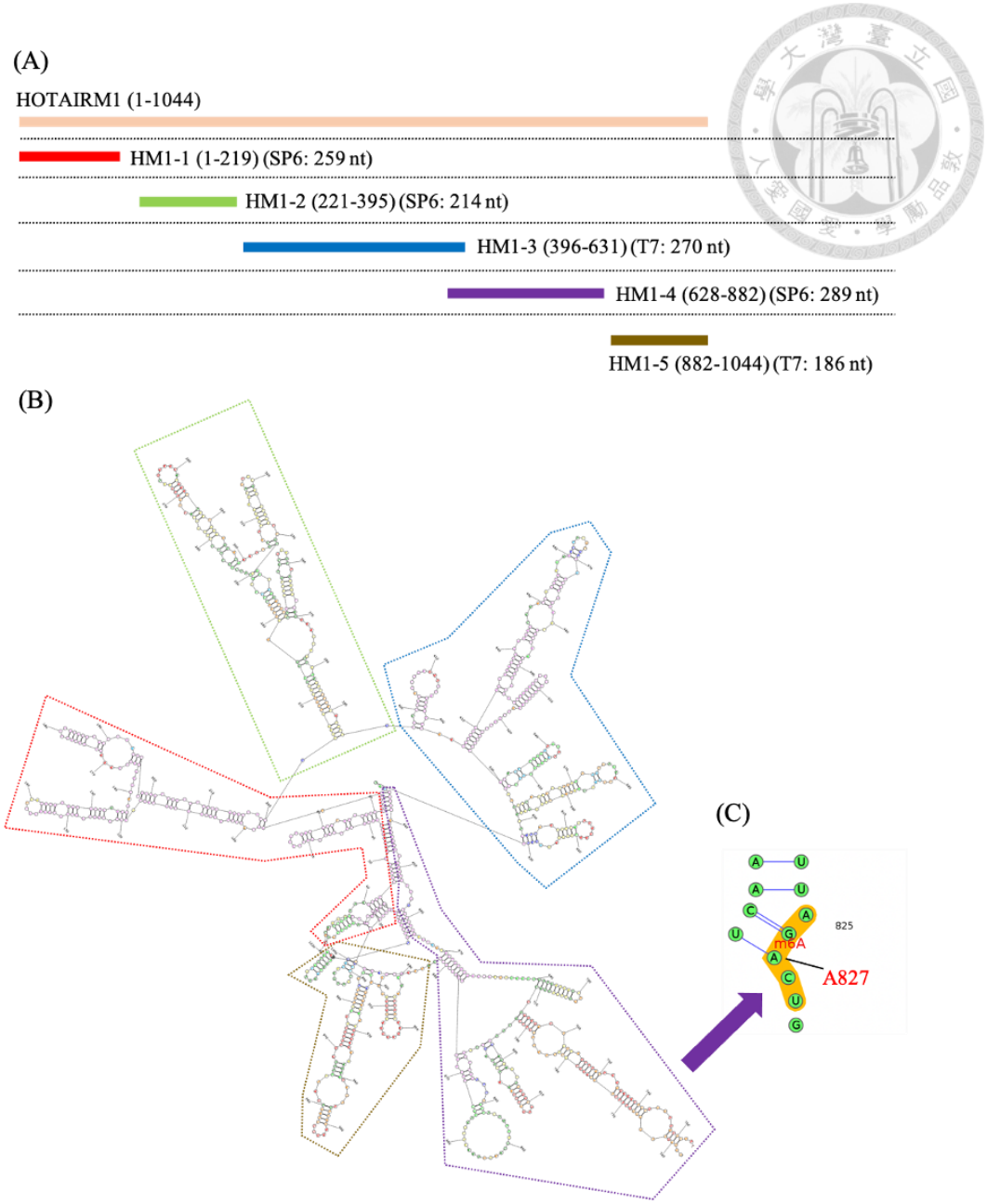

doi:10.6342/NTU202302064

**Appendix 2. Diagram showing Y14 purification process and results.**

(A) Illustration depicting S-Y14-His purification steps and (B) SDS PAGE results showing S-Y14-His purification results. (C) Illustration depicting Y14 purification steps. (D) SDS PAGE results showing purified recombinant Y14 WT (left) and S-Y14-His WT, WV, and ΔC mutant (right) proteins. Protein bands were stained with ToolStart Blue Staining Reagent.



doi:10.6342/NTU202302064

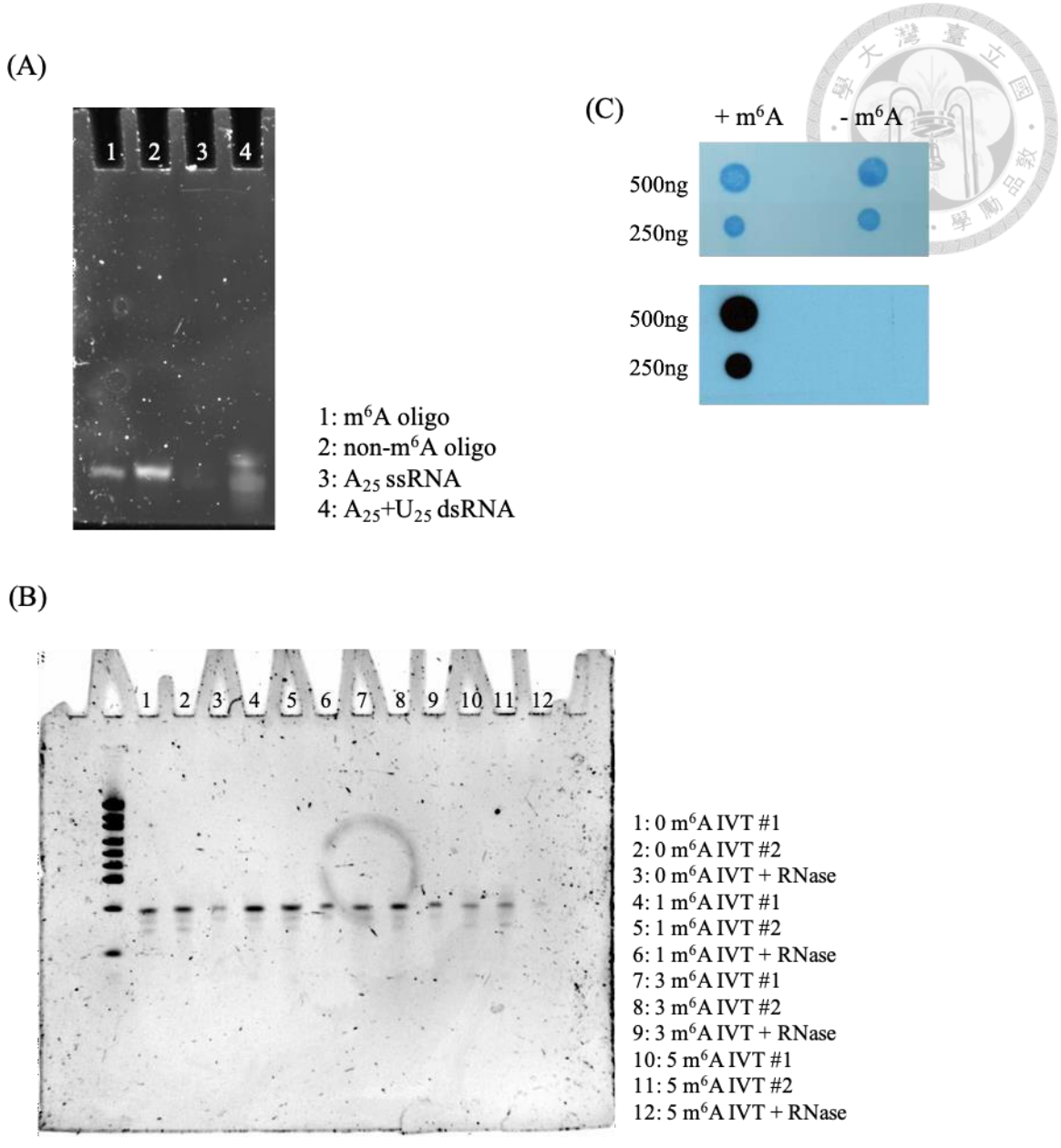



**Appendix 3. Diagram showing lncRNA HOTAIRM1.**

(A) HOTAIRM1 fragments, (B) its structure, and (C) methylation site.



doi:10.6342/NTU202302064

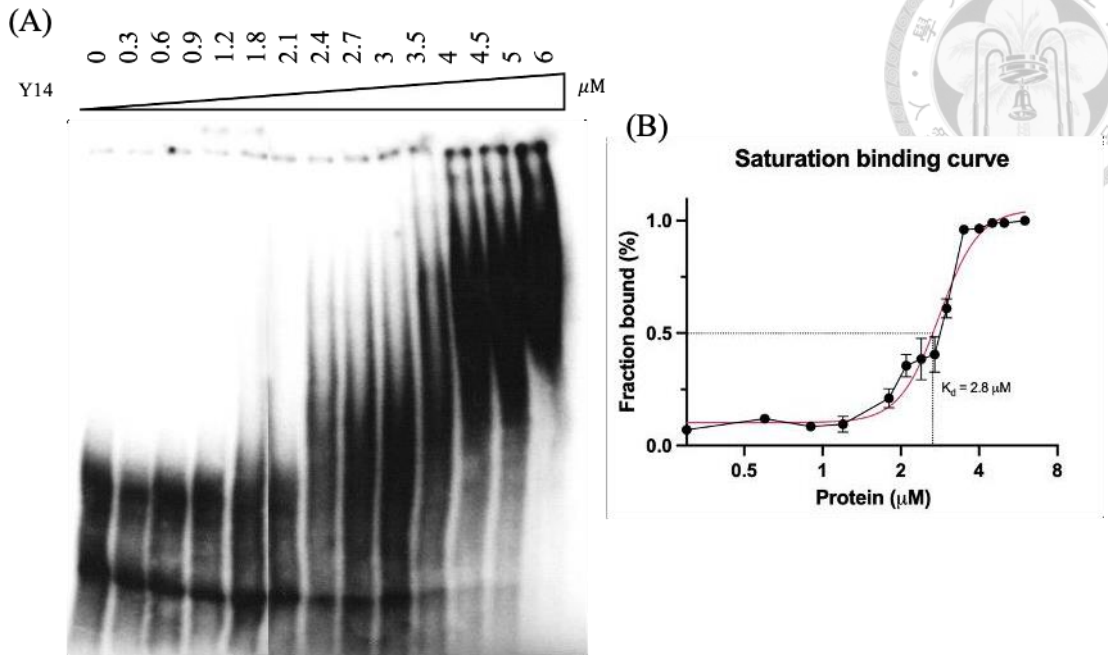



**Appendix 4. RNA materials used in experiments.**

6% Urea-polyacrylamide gel electrophoresis after PAGE GelRed stain showing (A) RNA fragments and (B) *in vitro* transcription of methylated fragments. (C) Dot blot assay for the screening of methylation of m6A oligo and non-m6A oligo.



doi:10.6342/NTU202302064




**Appendix 5. Saturation binding curve of HM1-4 fragment to S-Y14-His.**

(A) EMSA binding analysis of 10 nM HM1 fragment 632-882 with increasing concentrations of recombinant S-Y14-His WT. (B) The percentage of HM1 fragments bound at each protein concentration is shown graphically. $K_d$ is estimated via a saturation binding curve.



doi:10.6342/NTU202302064

**Table 2: NHEJ Protein IDRs and its various properties**

| Protein (Gene) | IDR | Size | Position | FCR | NCPR | Kappa | Hydropathy | Phase Plot region | pI value |
|---|---|---|---|---|---|---|---|---|---|
| Ku70 (XRCC6) | MSGWESYYKT EGDEEAEEEQ EENLEASG | 28 | 1-28 | 0.42857 | -0.35714 | 0.4787 | 2.78929 | 4 | 3.59 |
| | PEGKVTKRKH DNEGSGSKRP KVEYSEE | 27 | 536-562 | 0.48148 | 0.03704 | 0.2918 | 2.4037 | 3 | 8.72 |
| DNA-PKcs (PRKDC) | QSYSYSSQDP RPATGRFRRR EQRD | 24 | 2050-2073 | 0.375 | 0.125 | 0.22962 | 2.27083 | 3 | 10.74 |
| | LPEDNSMNVD QDGDPSDRME VQE | 23 | 3200-3222 | 0.3913 | -0.30435 | 0.15584 | 2.94783 | 3 | 3.54 |
| Artemis (DCLER1C) | LENFPSSTVA GGSQSPKLFS DSDGESTHIS SQNSSQSTHI TEQGSQGWDS QS | 52 | 504-555 | 0.13462 | -0.09615 | 0.28339 | 3.5 | 1 | 4.29 |
| | STNADSQSSS DFEVPSTPEA ELPKR | 25 | 640-664 | 0.28 | -0.12 | 0.26003 | 3.32 | 2 | 4.18 |
| XRCC4 | EGETAICSEM TADRDPVYDE STDEESENQT DLSGLASA | 38 | 212-249 | 0.34211 | -0.28947 | 0.23667 | 3.55 | 2 | 3.43 |
| DNA Pol λ (POLL) | LLQTALSPPP PPTRPVSPPQ KAKEAPNTQA QPISDDEASD GEETQ | 45 | 161-205 | 0.22222 | -0.08889 | 0.44885 | 3.36444 | 1 | 4.2 |

| Protein (Gene) | IDR | Size | Position | FCR | NCPR | Kappa | Hydropathy | Phase Plot region | pI value |
|---|---|---|---|---|---|---|---|---|---|
| DNA Pol μ (POLM) | MLPKRRRARV GSPSGDAASS TPPS | 24 | 1-24 | 0.25 | 0.16667 | 0.86426 | 3.58333 | 2 | 12 |
| | TDIAPSRKRR QRMQRNLGTE PKMAPQENQL QEKENSRPDS SLPETSKKEH ISAENMSLET LRNSSPEDLF DEI | 73 | 264-336 | 0.35616 | -0.0274 | 0.2535 | 3.02055 | 3 | 5.38 |
| XLF (NHEJ1) | NQPEQLVSSA PTLSAPEKES TGTSGPLQRP QLSKVKRKKP RGLFS | 45 | 255-299 | 0.24444 | 0.11111 | 0.50669 | 3.48444 | 1 | 10.55 |
| Tyrosyl DNA phospho-diesterase 1 (TDP1) | MSQEGDYGRW TISSSDESEE EKPKPDKPST SSLLCARQGA ANEPRYTCSE AQKAAHKRKI SPVKFSNTDS VLPPKRQKSG SQEDLGWCLS SSDDELQPEM P | 101 | 1-101 | 0.30693 | -0.0297 | 0.30173 | 3.30792 | 2 | 5.16 |
| PNKP | TRTPESQPDT PPGTPLVSQD EKRDAELPKK RMRKS | 35 | 109-143 | 0.4 | 0.05714 | 0.26007 | 2.74857 | 3 | 9.52 |
| Aprataxin (APTX) | GLETHRKRKR SGNSDSIERD AAQEAEAGTG LEPGSNSGQC SVPLKK | 46 | 122-167 | 0.32609 | 0.02174 | 0.36967 | 3.25435 | 2 | 8.16 |
| PAXX | EETAVSPRKS PRPAGPQLFL PDPDPQRGGP GPGVRRRCPG ESLINPGFKS KKPAGGVDFD ET | 62 | 143-204 | 0.29032 | 0.03226 | 0.36544 | 3.47258 | 2 | 9.12 |
| TDP43 (TARDBP) | EPKHNSNRQL ERSGRFGGNP GGFGNQGGFG NSRGGGAGLG NNQ | 43 | 261-303 | 0.16279 | 0.06977 | 0.13749 | 3.09302 | 1 | 11.54 |
| | ASQQNQSGPS GNNQNQGNMQ REPNQAFGSG NNS | 33 | 341-373 | 0.06061 | 0 | 0.18456 | 2.63939 | 1 | 6.05 |
| APLF | KSQLNTTQQG RRQLISSGSS ENTSAEQDTG EECKNTDQEE STISSKEMPQ SFSAITLSNT EMNNIKTNAQ RNKLPIEELG KVSKHKIATK RTPHKEDEAM SCSENCSSAQ GDSLQDESQG SHSESSSNPS NPETLHAKAT DSVLQGSE | 148 | 223-370 | 0.25676 | -0.05405 | 0.23984 | 3.28378 | 2 | 5 |
| | LDEDNDNVGQ PNEYDLNDSF LDDEEEDYEP TDEDSDWEPG KEDEEKEDV | 49 | 449-497 | 0.55102 | -0.46939 | 0.25689 | 2.42653 | 4 | 3.28 |

| Protein (Gene) | IDR | Size | Position | FCR | NCPR | Kappa | Hydropathy | Phase Plot region | pI value |
|---|---|---|---|---|---|---|---|---|---|
| WRN | ERNLGLPTKE EEEDDENEAN EGEEDDDKDF LW | 32 | 498-529 | 0.59375 | -0.40625 | 0.47828 | 2.4125 | 4 | 3.71 |
| P53 (TP53) | IEQWFTEDPG PDEAPRMPEA APPVAPAPAA PTPAAPAPAP SWPLSSS | 47 | 50-96 | 0.14894 | -0.10638 | 0.29457 | 4.04894 | 1 | 3.77 |
| | RRTEEENLRK KGEPHHELPP GSTKRALPNN TSSSPQPKKK PLDG | 44 | 282-395 | 0.36364 | 0.09091 | 0.24208 | 2.60227 | 3 | 9.99 |
| | KDAQAGKEPG GSRAHSSHLK SKKGQSTSRH KKLMFKTEGP DSD | 43 | 351-393 | 0.34884 | 0.11628 | 0.26245 | 2.9 | 2 | 10 |
| Nibrin (NBS1) | QGHPSTGLKT TTPGPSLSQG V | 21 | 326-346 | 0.04762 | 0.04762 | 1 | 3.8381 | 1 | 8.76 |
| | LSQDAPTVKE SCKTSSNNNS | 20 | 396-415 | 0.2 | 0 | 0.18545 | 3.325 | 1 | 6.06 |
| | QLSPTKLPSI NKSKDRASQQ QQTNSIRNYF QPSTKKRERD EENQEMSSC | 49 | 430-478 | 0.30612 | 0.06122 | 0.19705 | 2.7449 | 2 | 9.43 |
| MRE11 | TRQKNTNEED DEVREAMTRA RALRSQSEES ASAF | 33 | 507-540 | 0.41176 | -0.05882 | 0.40934 | 2.97059 | 3 | 4.86 |
| | NDSDDSISAA TNKGRGRGRG RRGGRGQNSA SRGGSQRGRA DTGLETSTRS RNSKTAVSA | 59 | 556-614 | 0.30508 | 0.13559 | 0.3461 | 3.00169 | 2 | 11.92 |
| | VEEDIFPTTS KTDQRWSSTS SSKIMSQSQV SKGVDFESSE DDDDDPFMNT SSLRRNRR | 58 | 651-708 | 0.34483 | -0.06897 | 0.44009 | 3.21552 | 2 | 4.51 |



doi:10.6342/NTU202302064

| Protein (Gene) | IDR | Size | Position | FCR | NCPR | Kappa | Hydropathy | Phase Plot region | pI value |
|---|---|---|---|---|---|---|---|---|---|
| BRCA1 | ETDVTNTEHH QPSNNDLNTT EKRAAERHPE KYQGSSVSNL H | 41 | 230-270 | 0.26829 | -0.07317 | 0.08553 | 2.77805 | 2 | 5.68 |
| | NKSKQPGLAR SQHNRWAGSK ETCNDRRTPS TEK | 33 | 306-338 | 0.33333 | 0.15152 | 0.09723 | 2.52727 | 2 | 10.44 |
| | QGTNQTEQNG QVMNITNSGH ENKTKGDSIQ NEKNPNP | 37 | 534-570 | 0.18919 | -0.02703 | 0.13178 | 2.64865 | 1 | 5.53 |
| | KYNQMPVRHS RNLQLMEGKE PATGAKKSNK PNEQTSKRHD SDTFPELKLT NAPGSF | 56 | 654-709 | 0.28571 | 0.07143 | 0.19074 | 3.08393 | 2 | 9.78 |
| | VQKGELSRSP SPFTHTHLAQ GYRRGAKKLE SSEENL | 36 | 1181-1216 | 0.27778 | 0.05556 | 0.33178 | 3.35 | 2 | 9.52 |
| | KQMRHQSESQ GVGLSDKELV SDDEERGTGL EENNQEEQSM DSNLGEAASG CESETSVSED CSGLSS | 66 | 1322-1387 | 0.31818 | -0.19697 | 0.17886 | 3.31667 | 2 | 3.97 |
| | EDLRNPEQST SEKAVLTSQK SSEYPISQNP EGLSADKFEV SADSSTSKNK EPGVERSSPS KCPSLD | 66 | 1440-1505 | 0.30303 | -0.06061 | 0.08535 | 3.31515 | 2 | 4.65 |
| | ESGISLFSDD PESDPSEDRA PESARVGNIP SS | 32 | 1565-1596 | 0.3125 | -0.1875 | 0.2866 | 3.54063 | 2 | 3.85 |
| BRCA2 | SEAPPYNSEP AEESEHKNNN YEPNLFKTPQ RK | 32 | 37-68 | 0.3125 | -0.0625 | 0.31476 | 2.47187 | 2 | 5.1 |
| | VEPNDTDPLD SNVANQKPFE SGSD | 24 | 358-381 | 0.29167 | -0.20833 | 0.2006 | 3.23333 | 2 | 3.66 |
| | ENRQKQNIDG HGSDDSKNKI N | 21 | 2430-2450 | 0.38095 | 0 | 0.21976 | 2.22381 | 3 | 6.85 |
| | EQESSQASTE ECEKNKQDTI TTKKYI | 26 | 3393-3418 | 0.38462 | -0.07692 | 0.19853 | 2.81538 | 3 | 4.72 |