Molecular Cell

CellPress

Voices

Is transcriptional regulation just going through a phase?

Geeta J. Narlikar
University of California, San Francisco, United States

Differences matter: Characterization before definition

The formation of nuclear structures such as Cajal bodies and nucleoli has been long accepted as arising from phase-separation processes. However, liquid-liquid phase separation is now being implicated in a multitude of additional nuclear processes. This includes transcriptional repression via heterochromatin and transcriptional activation via enhancer-promoter cooperation. The seemingly widespread presence of phase-separation in biology has generated much excitement in research and biotechnology. Part of the excitement appears to derive from the intuitive nature of phase separation: it's easy to explain using simple examples such as oil in water.

Many findings of phase separation have been made in a test-tube using purified components and sometimes under far from physiological conditions. Therefore, there is understandable skepticism about biological relevance. Well-intentioned attempts to impose rigor have led to definitions for what constitutes a liquid phase-separated domain in a cell. However, I feel that imposing definitions at this stage in the field's development is premature. Definitions that constrain may blind us to the beautiful complexity underlying the mechanisms and structures of phase-separated states in cells.

I have found it useful to compare the emerging discussions in this field to historical discussions in other fields, for example, after the proposal of Watson-Crick base-pairing. In his excellent historical review, Dr. Hashim Al-Hashimi (DOI 10.1002/bip.22334) describes how the DNA double-helix model proposed by Watson and Crick, although biologically meaningful, was met with skepticism. This is because direct experimental evidence took another two decades. In the meantime, Hoogsteen rather than Watson-Crick base-pairing was seen in solved structures. As discussed in this review, Hoogsteen base-pairs are now viewed as a means to "expand the structural and functional versatility of duplex DNA." Importantly, the biological importance of both types of base-pairing was inferred from studies in a test-tube using purified components.

I'm applying a few lessons learned from such stories to the phase-separation field. First, I am keeping an open mind to the potential diversity of phase-separation mechanisms in biology compared to simple phase-separating systems studied by chemists and physicists. Even well-established phase-separated states like the nucleolus contain several protein and nucleic acid components. These compoents have more diversity of molecular structure than simple oil-in-water or polyion-based systems. Second, just as careful biophysical studies of DNA revealed biologically meaningful structural variation, in-depth biophysical study of putative phase-separating systems may uncover unexpected biological relevance. We stumbled into phase separation because of keen observations made by graduate student Adam Larson about the biophysical behavior of the human HP1α protein. Further biophysical characterization of HP1 proteins in a test-tube has led us to auto-inhibition-based control of phase separation, deformation of chromatin as a driver of phase-separation, and examples of how DNA's polymer properties can be co-opted to form stable genomic territories. Many of these properties go beyond existing definitions of phase-separated states being discussed in the field. Yet, analogous to the DNA double-helix story, these properties do help explain heterochromatin's diverse biological roles.

At the same time, finding ways to impose rigor on interpretation, which underlies the efforts of crafting definitions, is extremely important. Our approach to ensuring rigor has been to (1) carry out quantifiable biophysical experiments under conditions that are close to physiological and (2) test the biological relevance of biochemically identified interactions through mutagenesis in cells. I anticipate that, in some cases, meso-scale droplets observed in a test-tube will not be seen in a cell, but the underlying interactions will still play significant biological roles. In the future, I look forward to discussing with

my colleagues the different types of phase separation in cells and using the differences to better appreciate the broader potential of phase separation in biology.

Sua Myong
Johns Hopkins University, United States

The condensate is the context

The transcription field recently had an ''aha'' moment. In what seemed like one fell swoop, transcription factors, coactivators, and RNA polymerase II were all caught in the act of transcribing a gene from the inside of a liquid droplet, i.e., biomolecular condensate. For years, many researchers wondered why many transcription factors harbor an intrinsically disordered domain despite being otherwise highly structured. The discovery that these domains can promote liquid-liquid phase separation (LLPS) offered a likely answer. Despite some controversies in the field, this new framework makes sense for an active process like transcription, which needs to recruit multiple components of the transcription machinery and keep them all together in time and space with the RNA template. Adding to the excitement, Rick Young's laboratory recently revealed that transcription can occur in a burst-like manner, tuned by the local RNA concentration, offering new insight into the role of noncoding RNA production in the vicinity of a gene it regulates. Beyond just a new way of understanding a fundamental biological process, the formation of liquid-like transcription condensates has been implicated in cancer and neurodegenerative diseases. Thus, elucidating the molecular details of this new form of condensate biology may help create avenues for developing novel therapeutics.

Much attention has been directed at the phenomenon of LLPS, but the condensate is the *context*, not the *content*. It is a transient and dynamic framework that empowers the cell to concentrate a well-defined set of molecules for a specific purpose, at a certain time, and in a certain location. But condensates can't explain everything and often tend to underestimate the bafflingly intricate sequence of molecular events that entail stoichiometric complex formation, specific conformational changes of protein and RNA constituents, distinct roles of structured domains interspersed with the disordered segments, and kinetic and thermodynamic parameters that govern intracondensate interactions. It is axiomatic that if we could build a tiny microscope to look deep inside these condensates, we would see fine-tuned molecular players and hierarchical principles masked by the somewhat simplistic LLPS description. Ultra-precise molecular tools including super resolution microscopes, single particle tracking, cutting edge proteomics and transcriptomics, along with quantitative microrheology test platforms are promising detection methods that will empower the field to map out the true molecular details lurking within each droplet. Most likely, we will find multiple routes of molecular assembly including LLPS, LSPS (liquid-solid phase separation), PPPS (polymer-polymer phase separation), micro phase separation, and beyond.

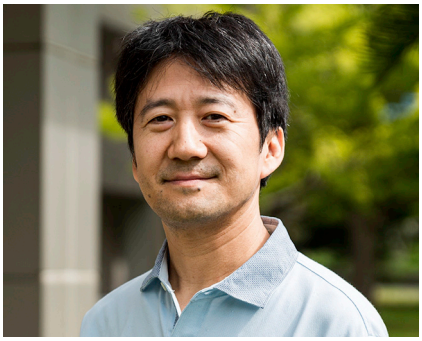
Daniel Larson
National Cancer Institute, NIH, United States

Spatial heterogeneity in the nucleus: Just another phase?

Spatial heterogeneity of the transcription and splicing machinery in the mammalian nucleus has been studied for decades and has been characterized extensively by light microscopy in living and fixed cells. Whether imaging sequence-specific transcriptional activators, RNA binding proteins, components of the general transcription and splicing machinery, and even RNA polymerase II, a common feature is that one observes spots in the fluorescence microscope. These spots have been called ''foci,'' ''factories,'' ''hubs,'' ''clusters,'' ''speckles,'' ''condensates,'' and ''phases.'' The proliferation of vocabulary indicates that the origin and function of these spots remain an enduring question in cell biology. What is clear is that these spots are dynamic, exchanging constituents with the surrounding nucleoplasm, and their spatial dimension is similar to the spatial resolution of light microscopy of ~300 nm. The time resolution of fluorescence microscopy (~milliseconds to days) is sufficient to capture most biologically relevant dynamics, but the spatial resolution is frustratingly close to the phenomena of interest and largely insufficient for directly observing spatial features of transcriptional activation and splicing.

However, temporal heterogeneity can severely complicate our ability to interpret spatial heterogeneity. We know that genes transcribe sporadically in bursts, meaning local accumulation of the transcription machinery due to simple kinetics of binding to DNA would be expected: a difference in residence time between specific and non-specific binding can be sufficient to generate a visible spot in the fluorescence microscope. Moreover, the techniques for interrogating spots are indirect and leverage trade-offs in time or space. For example, sparse imaging (either by partial labeling, excitation, or detection) can be used to enhance spatial resolution, but the image is then a projection of many sequential images acquired over time. The null hypothesis—the degree to which transcription and splicing spots are due to the separation in timescales—is seldom considered.

Weak interactions and inefficiency are hallmarks of gene regulation. For example, there are more RNA polymerase II molecules in the vicinity of an active gene than engage in productive elongation. Such interactions are facilitated by intrinsically-disordered regions. However, it is not clear whether there is an emergent property of this inefficient binding that might lead, for example, to phase separation. In cases that have been studied in depth, IDRs form transient low-affinity interactions rather than unstructured liquids. The biophysical concepts of avidity and affinity play essential roles in the interactions between proteins and chromatin. A key measurement is the abundance of factors in spots. However, these estimates in the cell (of RNA polymerase II, CTCF, Mediator, etc.) vary by two orders of magnitude. A cluster of a few molecules transiently associated with chromatin might not be sufficient for LLPS. Considering the experimental caveats, the transient and infrequent nature of transcriptional activation, and the abundance of factors in spots, I find a bottom-up view based on reaction-diffusion dynamics a more helpful paradigm than a thermodynamically-defined phase.



Kazuhiro Maeshima
National Institute of Genetics and SOKENDAI, Japan

New door to understanding transcription, open with caution

Liquid-liquid phase separation (LLPS) is a process where biomolecular droplets/condensates or membrane-less compartments are formed. Recently, the self-organization of RNA polymerase II (RNA Pol II), mediators, and other transcription-related factors into nuclear liquid droplets/condensates by LLPS has been proposed. Such transcription droplets/condensates can contribute to efficient transcription and gene regulation in space and time because LLPS can concentrate certain required factors, while excluding other factors that suppress transcription. Therefore, LLPS mediated transcription is an exciting new concept, which may explain many aspects of transcription.

However, concrete evidence to prove this concept is needed. While we know that RNA Pol II and some transcription factors form droplets/condensates *in vitro*, it remains unclear whether such transcription droplets/condensates are really formed by LLPS or by another process, in the cell. For the moment, techniques to test this concept in the cell remain limited. For instance, while 1,6-hexanediol (1,6-HD) is widely used for melting liquid droplets formed by LLPS *in vitro* and *in vivo*, we revealed by single-nucleosome imaging that 1,6-HD rapidly immobilizes and condenses chromatin in living cells (Itoh et al., 2021). This action was totally distinct from the reported droplet-melting activity of 1,6-HD. Liquid droplet results obtained using 1,6-HD should be carefully interpreted or reconsidered when these droplets are associated with chromatin. Fluorescence recovery after photobleaching (FRAP) is another technique used to image fluorescently tagged protein components in the droplets/condensates. However, FRAP only indirectly shows changes in mobility and cannot examine whether these protein components behave as "liquid" or not (McSwiggen et al., 2019).

Higher-resolution quantitative analyses of liquid droplets/condensates in living cells is needed to decipher their molecular behavior as component molecules diffuse within and can also transition in and out of these condensates. To this end, single-molecule imaging and tracking is a promising solution. Indeed, this technique revealed the molecular behavior of RNA polymerase I (Pol I) in the nucleolus, a nuclear body with multiphase liquid droplets (Ide et al., 2020). As previously observed for RNA Pol II, active RNA Pol I forms clusters and constrains ribosomal DNA chromatin in the nucleoli during

transcription. But when transcription was inhibited, active RNA Pol I dissociated from ribosomal DNA and moved like a liquid within newly formed droplets in nucleoli (Ide et al., 2020). Could this be the same case for RNA Pol II mediated transcription? Elucidation of detailed molecular behaviors of the components in the transcription droplets/condensates would provide important answers and give us new insight into the RNA Pol II transcription mechanism, regardless of the involvement of LLPS. Furthermore, given that chromatin locally appears liquid-like, more exciting questions arise: How does liquid-like chromatin interplay with the droplets/condensates? How can liquid-like behaviors contribute to the transcription process and what does it mean for the regulation of transcription? Let's open the door to these unknowns, but I urge a cautious multifaceted approach in finding the answers.

Nicole Francis
Montreal Clinical Research Institute, Canada

An intuitive explanation

The mechanisms underlying physical and functional organization of the nucleus, including the large-scale organization of chromatin, are a challenge in modern biology. How can biochemical events at the nucleosome level affect chromatin organization at a megabase scale? Phase separation provides a mechanistic explanation for how the physical chemical properties of proteins and nucleic acids can lead to self-organization at a macromolecular scale. The weak and dynamic interactions involved in phase separation provide a coherent rationale for the abundance of intrinsically disordered sequence in chromatin proteins, and the central role of weak interactions between proteins and histone post-translational modifications in chromatin regulation.

Phase separation provides an intuitive explanation for how properties of individual molecules drive collective assembly, but this does not mean it is the correct explanation in all cases. Much of the theoretical and experimental work has used highly simplified systems consisting of single proteins and isolated domains. These experiments reveal how phase separation can drive molecular organization, which sequences are required, and how condensate formation influences biochemical activities. They provide a framework for considering phase separation in diverse nuclear processes. The analysis of more realistic multi-component condensates, through live imaging, functional reconstitution, and computational modeling, will be important in bridging core physical principles with biological reality. An elegant example of this is a recent analysis of the histone locus body that forms and dissolves through phase separation in a cell-cycle regulated manner to process the highly transcribed histone mRNAs in S-phase that are needed to maintain cell division in developing *Drosophila* embryos (Hur et al., 2020).

Genomics approaches have provided a detailed description of the three-dimensional organization of chromatin and identified some molecules and mechanisms that control it, but we are still at early stages in understanding the relationship between visible condensates formed in cells and contact maps of genomic interactions. The properties of the chromatin polymer dominate the nuclear environment. These properties, and how they affect or participate in phase separation, are far from being understood. This includes whether chromatin itself can exist in an intermixing liquid state in cells (or under which conditions) (Gibson et al., 2019; Strickfaden et al., 2020). Live-cell single-molecule imaging provides a means to understand biochemical mechanisms in the context of the nuclear environment. Mathematical modeling of mobility states of wild type and mutant versions of the HP1 homolog Swi6 in living yeast cells measured with this approach made it possible to develop a thermodynamic framework matching mobility states to biochemical intermediates. The analysis revealed how formation of higher order assemblies allows weak but specific interactions with methylated histones (H3K9me3) to drive collective assembly of heterochromatin domains (Biswas et al., 2021).

Karsten Rippe
German Cancer Research Center (DKFZ) and Bioquant, Germany

More questions than answers

Invoking a liquid-liquid phase separation (LLPS) process in the context of genome-associated activities is becoming commonplace these days. While there is increasing evidence that the activity of RNA polymerase I in the nucleolus is tightly linked to its LLPS features, there are currently more questions than answers for RNA polymerase II (Pol II) transcription. LLPS is an intriguing explanation for the enigmatic role of intrinsically disordered regions (IDRs) present in Pol II, transcription factors and co-activators as they form liquid droplets *in vitro*. I would be cautious, however, to conclude from these findings that the corresponding full-length proteins will do the same in the cell nucleus when present at endogenous concentrations. Furthermore, IDRs could also mediate the assembly of the active transcription machinery via multivalent interactions in the absence of LLPS. I find it surprising that many current studies on Pol II "transcriptional condensates" do not consider alternative mechanisms and lack a comparison with previously developed models. For example, references and discussions of the work that Peter Cook and others have conducted since the 1990s to characterize "transcription factories" are mostly absent. These studies, together with a large body of research on chromatin, provide mechanistic models and quantitative descriptions for gene activation or silencing. They rationalize formation of Pol II transcription subcompartments by direct and indirect interactions of proteins and RNA with chromatin that folds into distinct 3D structures due to bivalent bridging interactions between distant parts of the nucleosome chain. I consider this the "null hypothesis" against which a potential LLPS mechanism needs to be evaluated. This comparison would also be very valuable to clarify, which aspects of Pol II transcription are specific for a LLPS mechanism. Assessing the potential function of liquid transcription factor droplets to enhance transcription activation would also profit from having a better reference state. Otherwise, determining the contribution of LLPS becomes an apples-and-oranges type of comparison, if transcription factors with different propensities to engage in multivalent interactions are compared. In summary, I hope for more informative control and perturbation experiments in living cells and a critical evaluation against alternative mechanisms to better understand how LLPS might regulate Pol II transcription.

Benjamin Sabari
UT Southwestern Medical Center, United States

Crossing thresholds and specificity of dynamic multivalent interactions

Turning on the right set of genes at the right time is a fundamental biological process that requires the concerted action of dozens of different proteins assembled at specific regions of DNA. While we have known that the expression of some genes such as ribosomal RNA is associated with nuclear compartments, more recently, a combination of imaging and biochemical approaches has revealed that many components of the RNA Polymerase II regulatory machinery are dynamically compartmentalized within the nucleus, often at sites of transcription. Strikingly, these compartments form and dissolve with a large distribution of lifetimes suggesting that there is locus-specific regulation. Multivalent, and often dynamic, interactions among protein, RNA, and DNA molecules drive the assembly of such compartments above threshold concentrations, interaction affinities, and interaction valences. A general framework has begun to emerge for how cells reversibly cross these thresholds at specific loci through active processes such as local RNA synthesis, reversible covalent modifications, and accessibility of regulatory elements. The specific mechanisms by which individual genomic loci assemble these networks, the properties of the emergent compartment, and their dynamic regulation is an exciting and rapidly progressing area of investigation. We should expect that additional mechanisms for traversing these thresholds will be discovered for specific normal or disease-relevant transcriptional events.

Another exciting area of investigation is how dynamic multivalent interaction networks compartmentalize or exclude specific macromolecules, either as obligate co-scaffolds or as clients. How do dynamic multivalent interactions engender specificity required to coordinate a multistep and multicomponent biochemical process like transcription? How do disease-associated mutations, repeat-expansions, fusions, or over-expression events integrate into or disrupt these dynamic interaction networks? As these interaction networks form at specific and context-dependent thresholds, they

have likely been overlooked by standard discovery methods. *In situ* and cell-free methods in which these interactions are maintained or reconstituted will be important for their identification and characterization. The incredible diversity of multivalent interactions inherent in the gene control machinery—intrinsically disordered regions, oligomerization domains, "reader" domains, ligand/PTM-induced interactions, repetitive motifs, and large multi-domain proteins or multi-subunit complexes containing each of the above—begs for further investigation into how the ensemble of heterogenous multivalent interactions leads to specific properties, compositions, and functions of the emergent compartment.

Lucia Strader
Duke University, United States

Linking condensation to function in plants

Mounting evidence suggests that condensate formation and phase separation are integral to regulating multiple aspects of the indispensable process of gene transcription. Chromatin remodeling, RNA polymerase activity, RNA splicing, and transcription factor access to DNA have all been shown to be regulated by phase separation in various contexts. Biomolecular condensates comprised of proteins involved in these aspects of transcriptional regulation have been described. In two such plant biology examples, protein condensation has been shown to attenuate the activity of transcription factors. In one case, condensate formation of the transcriptional activator ARF19 is developmentally regulated and results in reduced competence to respond to the plant hormone auxin. In the other case, temperature-dependent condensation of the transcriptional repressor ELF3 allows for relief of repression to control flowering. In each of these cases, protein condensation is strongly tied to biological function; understanding this connection was facilitated by the ability to identify point mutations or natural variants in which the core transcriptional function was unaltered, but condensation was disrupted. In my opinion, a current limitation to understanding how phase separation regulates transcription lies in our inability, in many cases, to cleanly disrupt phase separation without altering other core protein functions. Indeed, it is difficult to truly know the effects of any mutation on protein interactions in addition to altering phase separation behavior. An increased ability to do so, whether by using exciting new toolkits for regulated phase separation, identification of natural variants, or use of close homologs that do not exhibit phase behavior, will allow for dissection of whether phase separation promotes activity, reduces activity, or acts as a buffering mechanism to regulate transcriptional processes.

Robert Tjian
University of California, Berkeley, United States

Droplets everywhere, but still thirsting for evidence

Classical *in vitro* biochemistry suggested that protein complexes stably assemble with precise stoichiometry via interaction of structured domains. Today, we know this is not always the case—instead, many cellular processes rely on weak, transient, multivalent interactions, including those between intrinsically disordered regions (IDRs). The role of IDRs as transcription factor (TF) activation domains has been known for decades, but we are only now unraveling how interactions between TF IDRs and coactivators mediate transcriptional regulation.

The current revival of interest in liquid-liquid phase separation (LLPS) in biological systems has led some researchers to conflate all IDR driven biomolecular interactions with the special case of LLPS. Although weak, multivalent interactions *may* facilitate LLPS in some cases, interactions between IDRs can also enrich proteins at specific genomic loci in the absence of phase separation. Single-molecule measurements reveal that such IDR-mediated "hubs" are dynamic structures within which proteins exchange on a timescale of seconds to minutes. Phase separation is fundamentally different from other forms of cooperative molecular assembly, as it entails a discrete transition in the number of stably coexisting phases as a function of some control parameter, such as protein concentration. Thus, the distinction between condensates and hubs is not just a matter of semantics. Only in a few cases has such a discrete liquid-liquid phase transition been demonstrated for endogenous proteins under physiologically relevant conditions, and often the biological function of such LLPS remains undetermined.

Many claims of LLPS rest upon experiments that do not assess LLPS within physiologically relevant contexts. *In vitro* droplet-forming assays typically involve non-physiological concentrations of just one or two distinct biomolecules in non-native buffer. Intracellular puncta labeled as "phase-separated condensates" are sometimes displayed with exaggerated image contrast or thresholding without clear descriptions of how images were processed for display. Local protein enrichment tells us little about the biophysical mechanisms mediating such hubs. Moreover, experiments reporting LLPS often involve highly overexpressed proteins, so even if LLPS did occur, it would remain unclear whether it is functionally relevant under physiological conditions.

LLPS is a fundamental physical phenomenon that likely plays a role in some biological processes. Applying this term indiscriminately is a disservice relegating it to a meaningless buzzword. Greater transparency and increased rigor in the criteria for invoking LLPS to describe protein behavior are required to better understand its functional role in biology. Emerging methods to monitor dynamic biomolecular behavior *in vivo* may provide new tools to more rigorously dissect how networks of weak, multivalent interactions assemble molecular complexes to regulate gene expression and other essential biological processes.