Article

Nucleation landscape of biomolecular condensates

https://doi.org/10.1038/s41586-021-03905-5

Received: 1 December 2020

Accepted: 12 August 2021

Published online: 22 September 2021

Check for updates

Shunsuke F. Shimobayashi[1], Pierre Ronceray[2,3], David W. Sanders[1], Mikko P. Haataja[4,5] & Clifford P. Brangwynne[1,5,6]✉

All structures within living cells must form at the right time and place. This includes condensates such as the nucleolus, Cajal bodies and stress granules, which form via liquid–liquid phase separation of biomolecules, particularly proteins enriched in intrinsically disordered regions (IDRs)[1,2]. In non-living systems, the initial stages of nucleated phase separation arise when thermal fluctuations overcome an energy barrier due to surface tension. This phenomenon can be modelled by classical nucleation theory (CNT), which describes how the rate of droplet nucleation depends on the degree of supersaturation, whereas the location at which droplets appear is controlled by interfacial heterogeneities[3,4]. However, it remains unknown whether this framework applies in living cells, owing to the multicomponent and highly complex nature of the intracellular environment, including the presence of diverse IDRs, whose specificity of biomolecular interactions is unclear[5–8]. Here we show that despite this complexity, nucleation in living cells occurs through a physical process similar to that in inanimate materials, but the efficacy of nucleation sites can be tuned by their biomolecular features. By quantitatively characterizing the nucleation kinetics of endogenous and biomimetic condensates in living cells, we find that key features of condensate nucleation can be quantitatively understood through a CNT-like theoretical framework. Nucleation rates can be substantially enhanced by compatible biomolecular (IDR) seeds, and the kinetics of cellular processes can impact condensate nucleation rates and specificity of location. This quantitative framework sheds light on the intracellular nucleation landscape, and paves the way for engineering synthetic condensates precisely positioned in space and time.

To quantify condensate nucleation in living cells, we measure their nucleation rate J, defined as the number of condensates formed per unit volume and per unit time. (Extended Data Fig. 1). Specifically, $J = \mathrm{d}\rho(t)/\mathrm{d}t|_{t=t_0}$, where $\rho(t)$ and t_0 denote the observed number of visible droplets per unit volume and a delay time after which droplets first emerge, respectively. For various endogenous condensates (Fig. 1a), we find that nucleation rates are in the range of 10^{-6} to 10^{-4} μm^{-3} s^{-1} (Extended Data Fig. 1b). To probe the molecular biophysical determinants of intracellular condensate nucleation in a tractable system, we take advantage of Corelets, an optogenetic tool which enables the spatial and temporal control of intracellular liquid–liquid phase separation using light[9]. Corelets is a two-module optogenetic system built around a light-activatable core that mimics the endogenous oligomerization of phase-separating proteins, which are often enriched in IDRs (Fig. 1b, c). Consistent with recent work[9–11], we find that light-activated oligomerization drives phase separation of constructs including FUS or HNRNPA1 IDRs (FUS_N, 1–214; $HNRNPA1_C$, 186–320), model aromatic-rich polar sequences[7,8,12], forming liquid condensates in human osteosarcoma (U2OS) cells; the nucleation probability is near zero within nucleoli, owing to their intrinsic exclusion of Corelets[13] (Extended Data Fig. 2a).

Two different phase-separation modes are observed: nucleation and growth, and connected network-like growth and coarsening akin to spinodal decomposition[9] (Fig. 1d, Extended Data Fig. 3, Supplementary Videos 1–3). These distinct behaviours reflect the location of the cell within the phase diagram, as shown in our experimentally measured phase diagram of FUS_NCorelets as functions of mean core concentration and core-to-IDR ratio (f) (Fig. 1g); the core-to-IDR ratio represents the inverse degree of IDR oligomerization, which affects the entropic contribution to the free energy[14], and is thus analogous to temperature. Weakly supersaturated cells lying close to the binodal display visible condensates only after about 80 s of activation, whereas cells close to the spinodal region show visible condensates in around 5 s or less (Fig. 1d, e, Extended Data Fig. 4, Supplementary Video 1).

To further quantitatively examine this dependence, we plot J against the initial supersaturation S, defined as $S = \log(C_{\mathrm{dil}}/C_{\mathrm{sat}})$, where C_{dil} is the initial core concentration in the dilute phase of the nucleus (or cytoplasm) and C_{sat} is the saturating core concentration[3] (Supplementary

[1]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ, USA. [2]Center for the Physics of Biological Function, Princeton University, Princeton, NJ, USA. [3]Aix Marseille Univ, Université de Toulon, CNRS, CPT, Turing Center for Living Systems, Marseille, France. [4]Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, USA. [5]Princeton Institute for the Science and Technology of Materials, Princeton University, Princeton, NJ, USA. [6]Howard Hughes Medical Institute, Princeton University, Princeton, NJ, USA. ✉e-mail: cbrangwy@princeton.edu

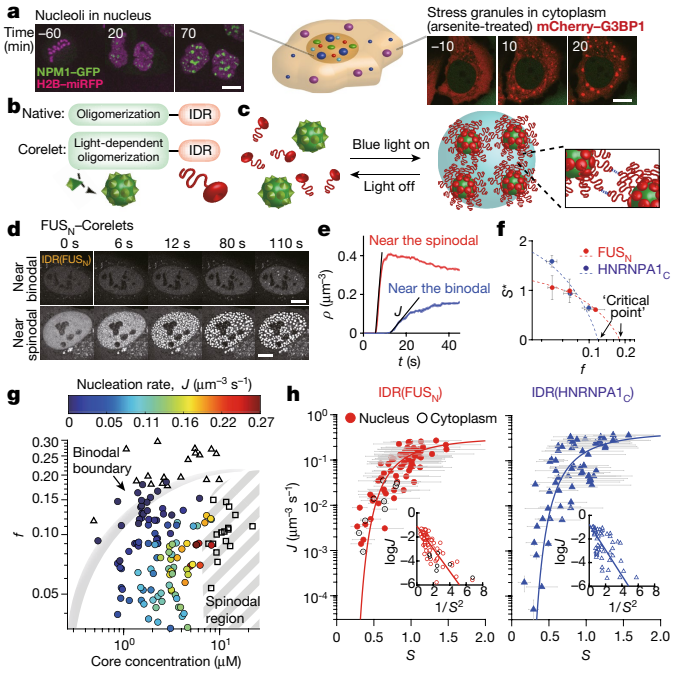


Fig. 1 | Quantifying the nucleation landscape with synthetic condensates. **a**, Time-lapse images of U2OS cells show the nucleation of nucleoli (NPM1–mGFP) and stress granules (mCherry–G3BP1) upon 400 μM arsenite treatment. Scale bar, 10 μm. **b**, Schematic of the Corelet system, which mimics the oligomerization of native condensate-forming proteins. **c**, Schematic of Corelet phase separation. Upon blue-light illumination, up to 24 IDR modules are captured by each core, which may subsequently phase-separate by multivalent IDR interactions. **d**, Time-lapse confocal images of photo-activated U2OS cells with different expression levels of FUS_N–Corelets. The cells display nucleation and growth between the binodal boundary and spinodal region. Scale bars, 10 μm. **e**, Difference in timing of light-induced droplet density ρ, for FUS_N–Corelets. The nucleation rate J is given by the slope as indicated. The maximum droplet density is $0.33 \pm 0.16\ \mu m^{-3}$ (mean ± s.d.) ($n = 124$). **f**, S^* versus core:IDR ratio f, for FUS_N–Corelets ($n = 115$) and $HNRNPA1_C$–Corelets ($n = 68$). The solid circles represent the mean and error bars for S^* and f show standard errors of the fits with equation (1) and standard deviation, respectively. Dashed lines show the best fit to $S^* \propto |f - f_c|^{0.84}$, where f_c is the critical core:IDR ratio. **g**, Phase diagram of FUS_N–Corelets as functions of core concentration and core:IDR ratio ($n = 175$). Solid circles indicate cells where nucleation and growth is observed, and empty triangles and squares indicate cells where no phase separation and spinodal decomposition are observed, respectively. Circle colours correspond to the observed nucleation rate J (see Extended Data Fig. 4 for phase diagram of $HNRNPA1_C$–Corelets). **h**, J versus supersaturation S, for FUS_N–Corelets ($n = 85$) and $HNRNPA1_C$–Corelets ($n = 64$) with $1/24 < f < 1/9$. Solid lines show the best fit to equation (1). n is the number of cells.

Discussion). Notably, we find that S controls J over three orders of magnitude, with a similar dependence for both FUS_N and $HNRNPA1_C$–Corelets (Fig. 1h). Despite the complexity of the intracellular milieu, these data can be fit to a functional form[3] predicted by CNT (see also Extended Data Fig. 5):

$$J(S) = \kappa \cdot \exp\left[-\left(\frac{S^*}{S}\right)^2\right] \quad (1)$$

where κ is a kinetic factor proportional to the condensate-specific density of possible nucleation sites and S^* is a dimensionless parameter that describes the crossover supersaturation between activated and transport-limited regimes of nucleation. At low supersaturation ($S < S^*$), nucleation is an activated process that is strongly dependent on the supersaturation. By contrast, for $S >> S^*$, $J \approx \kappa$, as the nucleation rate is controlled by microscopic transport processes and the limited density of condensate nucleation sites[3] (Supplementary Discussion). Finally, we quantified the dependence of the delay time t_0 on supersaturation and found that $t_0 \approx S^{-\xi}$, with $\xi \approx 1.2$ for FUS_N–Corelets and $\xi \approx 1.7$ for $HNRNPA1_C$–Corelets, probably reflecting surface reaction-limited growth[15] of the critical nucleus, for which $\xi \approx 2$ (Extended Data Fig. 6).

To understand the microscopic origin of S^*, we note that for inanimate systems, CNT predicts $S^* \approx \nu(\gamma/k_BT)^{3/2}$, where ν, γ and k_BT denote molecular volume, surface tension between the droplet and the surrounding fluid and thermal energy (with k_B being the Boltzmann constant), respectively. Furthermore, surface tension in inanimate systems strongly depends on the position of the system in the phase diagram: close to a critical point[16,17], $\gamma \to 0$, implying that $S^* \to 0$. Consistent with S^* being strongly affected by surface tension, we indeed find that S^* decreases monotonically as a function of the core-to-IDR ratio for FUS_N and $HNRNPA1_C$–Corelets (Fig. 1f, Extended Data Figs. 5, 7). Assuming that the core-to-IDR ratio f acts as an effective temperature around the critical point, theory predicts[17] that $S^* \propto |f - f_c|^{0.84}$. The best fit yields a critical core-to-IDR ratio ($f_c^{FUS_N} = 0.18 \pm 0.02$ and $f_c^{HNRNPA1_C} = 0.12 \pm 0.01$), consistent with the phase diagrams in Fig. 1g, Extended Data Fig. 4.

The dependence of J on S in equation (1) corresponds to a homogeneous system in which all possible nucleation sites are equivalent—that is, spatially nonspecific droplet formation, consistent with the roughly flat nucleation probability outside of nucleoli (Extended Data Fig. 2). However, a broad variety of condensates, such as those involved in transcription[18–21] or DNA repair[22], assemble at a specific locus of protein, DNA and/or RNA, suggesting that biomolecules can function as 'seeds'[23,24]. To parse the potential ability of the seeds to spatially modulate S^*—and thus to impart a more complex nucleation landscape—we use a convenient and engineerable chromatin locus, the telomere, which can be modified by fusing telomeric-repeat-binding protein TRF1 to various proteins, including IDRs (Fig. 2a). When FUS_N–TRF1 was co-expressed with FUS_N–Corelets, light-activated FUS_N condensates nucleated and grew at telomeres, as expected for IDR–IDR interaction between the condensates and telomere-associated seeds (Fig. 2b, Extended Data Fig. 8, Supplementary Video 4).

To examine the effects of seeds on S^*, we again measured J as a function of S for FUS_N–Corelets with FUS_N wild-type (WT) seeds. Of note, seeding results in a lower value of κ (indicating fewer nucleation sites) and a lower value of S^* (indicating facilitated nucleation), compared with FUS_N–Corelets without seeds (Fig. 2c). Quantitatively, we find that $S^*_{FUS,seed\,(WT)} = 0.44 \pm 0.04$, whereas $S^*_{FUS} = 0.97 \pm 0.08$ (see also Extended Data Fig. 5). Thus, favourable molecular interactions at specific intracellular seeds can catalyse the nucleation process by locally reducing S^*. Note that at high supersaturation, the driving force for phase separation is so large that the nucleation landscape is effectively 'blurred out', seeding becomes unimportant, and the $J(S)$ curve overlaps with the unseeded one (Fig. 2c, Extended Data Fig. 2b, Supplementary Video 5).

We quantify seeding efficacy by measuring the degree of co-localization of condensates and seeds using the Pearson correlation coefficient (PCC), which measures linear correlation between two images[25], revealing a monotonic decrease of nucleation specificity as the supersaturation S increases (Fig. 2d). This effect can be understood by again examining equation (1), now considering a nucleation landscape consisting of two populations of loci. Nucleation at specific (s) sites is characterized by nucleation parameters $S^{*;s}$ and κ^s, while non-specific (ns) nucleation corresponds to $S^{*;ns}$ and κ^{ns}, with $S^{*;s} < S^{*;ns}$. This allows us to define a specificity index $\Delta(S) = J_{specific}(S)/J_{total}(S)$, which indicates the fraction of droplets that form at specific seeds (Supplementary Discussion). Facilitated nucleation, with $\Delta(S) > 0.5$, occurs for supersaturations $S < S^\dagger$, where $S^\dagger = \sqrt{\frac{(S^{*;ns})^2 - (S^{*;s})^2}{\log(\kappa^{ns}/\kappa^s)}}$ indicates the supersaturation at which the specific and non-specific nucleation curves first overlap (Fig. 2c, $S^\dagger \approx 0.69$ for FUS_N–Corelets); we note that for nucleation to proceed with both rapidity and specificity, $S^{*;s} < S^\dagger$ (Fig. 2c, strong seeding). Using fitted values for the nucleation

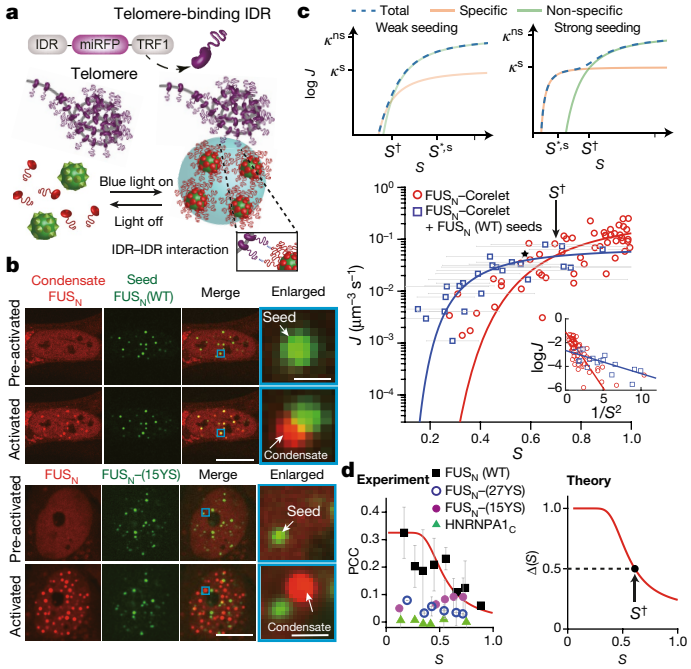


Fig. 2 | Engineering the condensate nucleation landscape. a, Schematic of a synthetic seed, consisting of miRFP-tagged telomere binding protein, TRF1, fused to IDRs, promoting localized Corelet phase separation. **b**, Fluorescence images of U2OS cells expressing FUS_N–Corelets and FUS_N(WT) or FUS_N(15YS) seeds, before and after blue light activation. Enlarged images of indicated regions are shown on the right. Scale bars, 10 μm (main images); 1 μm (enlarged images). **c**, Top, example *J*–S curves for the cases of weak ($S^{*;s} > S^{\dagger}$) versus strong ($S^{*;s} < S^{\dagger}$) seeding. Bottom, nucleation rate (*J*) plotted against the supersaturation (*S*) for FUS_N Corelets with ($n = 23$) or without ($n = 76$) FUS_N(WT) seeds with $1/24 < f < 1/9$. The solid lines show the best fit to equation (1). **d**, Left, PCC between mCherry (Corelets) and miRFP (seeds) signals for FUS_N Corelets and FUS_N(WT) ($n = 57$), FUS_N(15YS) ($n = 23$), FUSN(27YS) ($n = 18$) or $HNRNPA1_C$ ($n = 19$) seeds, against supersaturation (*S*). Right, theoretical prediction of nucleation specificity ΔS versus *S*. *n* is the number of cells.

parameters, we obtain a predicted curve for the specificity index whose variations are consistent with that of the PCC (Fig. 2d). Thus, to nucleate with high specificity ($\Delta(S) \to 1$), cells must not only provide energetically favourable seeding sites (that is, low values of $S^{*;s}$), but also maintain a low degree of supersaturation.

To gain further insight into the biomolecular determinants of $S^{*;s}$, we mutated the telomere-bound FUS_N sequence, changing the number of its tyrosines, a key amino acid mediating IDR–IDR interactions[7–9,11,12]. Specifically, we mutated *n* (5, 15 or 27) out of 27 tyrosines to serine in FUS_N–TRF1, creating FUS_N(nYS) seeds, and co-expressed them with FUS_N–Corelets. Upon blue light activation, FUS_N condensates still nucleated at the seeds in cells expressing FUS_N–Corelets plus FUS_N(5YS) seeds (Extended Data Fig. 9), similar to the behaviour in unmutated FUS_N seeds. However, with either 15 or 27 tyrosines mutated, the FUS_N seeds no longer nucleated FUS_N–Corelets (Fig. 2b, Extended Data Fig. 9), as can be seen in the near-zero PCC values for all supersaturation levels (Fig. 2d). Consistent with this sensitive dependence of the detailed IDR sequence in nucleation sites, telomere-fused $HNRNPA1_C$, which is known to promote phase separation and can interact with FUS[12], was similarly incapable of nucleating FUS condensates (Extended Data Fig. 9, Supplementary Video 6). By plotting *J* against *S* for FUS_N(nYS) seeds for $n = 15$ or 27 or $HNRNPA1_C$ seeds, all yield data roughly consistent with equation (1) with the best-fit S^* and κ for FUS_N–Corelets without seeds (Extended Data Fig. 9d). Collectively, these results suggest that $S^{*;s}$ exhibits a dependence on the biomolecular features of the seed.

Given the biophysical picture that emerges from examining these model condensates, we sought to evaluate its applicability in a system that more closely mimics an endogenous condensate. To this end, we used the Corelet system with the stress granule protein G3BP1, whose native oligomerization domain is deleted and replaced with the light activatable sspB module[6]. When expressed in G3BP-knockout cells, this leads to light-activated synthetic stress granules (opto-SGs), which recapitulate essentially all features of endogenous stress granules, including recruitment of other key stress granule proteins, and a strong dependence of their phase behaviour on RNA concentration[6]. Consistent with this previous work, upon flooding the cell with exposed RNA by treatment with sodium arsenite, opto-SGs form more readily than in untreated cells (Fig. 3a). Quantifying the nucleation rate as a function of supersaturation, we find that cells treated with arsenite as well as untreated cells both appear to fall on the same curve, which is well described by equation (1), with a lower value of S^* and prefactor compared with FUS_N–Corelets (Fig. 3b, Extended Data Fig. 5). The lower S^* probably reflects the effect of heterogeneous stress granule nucleation with endogenous binding partners (Extended Data Fig. 10a), again underscoring the biomolecular specificity of heterogenous condensate nucleation. To probe time-dependent changes in nucleation caused by arsenite treatment, we sequentially activated and then deactivated cells immediately after arsenite addition. Consistent with the degree of supersaturation increasing with each successive quench, owing to the steady increase in exposed-RNA concentration, the nucleation rate increases over the course of roughly 10–60 min (Fig. 3c, d). This corresponds to the system moving along the *J*–*S* curve in time (Fig. 3d), reflecting an interplay of non-equilibrium biological processes with our CNT-like framework.

An intriguing difference between the opto-SGs and endogenous stress granules is that the nucleation rate of endogenous stress granules is $\approx 10^{-4}$ μm^{-3} s^{-1}, which is only comparable to that of the opto-SGs at the earliest times after arsenite treatment (Fig. 3d), when the system is only modestly supersaturated upon light-stimulated quenching. Indeed, the nucleation rates for endogenous stress granules remain low across a range of larger values of nominal supersaturation ($S_{nom} = \log(C_{tot}/C_{sat})$) (Fig. 3b). Nonetheless, opto-SGs recapitulate most other aspects of endogenous stress granules[6], leading us to speculate that the endogenous condensates form at unexpectedly low values of effective supersaturation; the apparent supersaturation for SG nucleation can be determined by the intersection of the *J*–*S* curve for opto-SGs and the dashed horizontal line for endogenous stress granules (Fig. 3b), resulting in $S_{nuc}^{SG} \sim 0.1$. We reasoned that this low apparent supersaturation could potentially reflect a competition between the kinetics of nucleation and the kinetics of arsenite-induced RNA-release. To test this model, we examined the nucleation rates of our opto-SG system under conditions of constant illumination, more closely mimicking endogenous stress granules, where G3BP is oligomerized throughout the duration of arsenite treatment. Remarkably, under these conditions, the opto-SG nucleation rate upon arsenite treatment becomes nearly identical to that of endogenous stress granules (Fig. 3b). We can understand this kinetic effect by noting that any real physical system does not become instantaneously supersaturated, but instead is driven into a supersaturated state at a finite 'quench' rate α, that is, $S_{nom}(t) = \alpha t$. As the supersaturation steadily increases, nucleation commences at a certain supersaturation $S = S_{nuc}$, which is related to α via $\alpha \approx V_{cell} \cdot S_{nuc} \cdot J(S_{nuc})$, where V_{cell} is the cell volume (Fig. 3e, Supplementary Discussion). This causes a decrease in the dilute phase concentration (and hence a decrease in the supersaturation), thereby suppressing further nucleation. Consistent with this picture, in addition to endogenous stress granules, we find that Cajal bodies, DNA repair condensates and nucleoli all exhibit nucleation rates that are independent of the nominal supersaturation (S_{nom}) (Extended Data Fig. 10b), suggesting that the relatively slow kinetics of the governing biological processes control nucleation rates, by maintaining cells in states of low supersaturation. Notably, this effect is intimately coupled

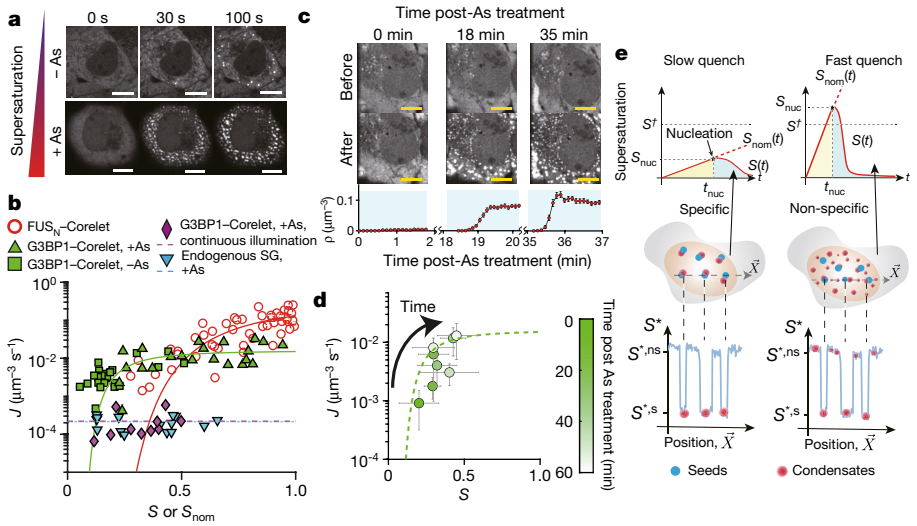


Fig. 3 | Temporal dynamics and the specificity of condensate nucleation. a, Time-lapse confocal images of photo-activated G3BP-knockout U2OS cells expressing G3BP1–Corelets with or without 1 h of treatment with 400 μM arsenite (As). Scale bars, 10 μm. **b**, Nucleation rate J plotted against the supersaturation S, for FUS_N–Corelets ($n = 76$) and G3BP1–Corelets with or without treatment with 400 μM arsenite ($n = 45$) with $1/24 < f < 1/9$, and plotted against nominal supersaturation S_{nom} for G3BP1–Corelets under conditions of constant illumination ($n = 10$) and endogenous stress granules (SG) ($n = 15$), both in aresenite-treated cells. Dashed lines show mean J values for the latter two conditions, and solid lines show best fit to equation (1). **c**, Top, fluorescent images of G3BP-knockout cells expressing G3BP1–Corelets before and after blue light activation at different time points after 400 μM arsenite treatment. Scale bars, 10 μm. Bottom, light-induced droplet number density ρ, at different time points after 400 μM arsenite treatment. **d**, Time change of J and S after treatment with 400 μM arsenite ($n = 56$). **e**, Schematic of the nucleation landscape of intracellular condensates. n is the number of cells.

to the specificity of the nucleation process: under slow quench rates, such that $\alpha \ll V_{cell} \cdot S^{\dagger} \cdot J(S^{\dagger})$, nucleation will only occur at specific nucleation sites, while fast quench rates, for which $\alpha \gg V_{cell} \cdot S^{\dagger} \cdot J(S^{\dagger})$, will lead to rapid and non-specific nucleation. Thus, the timescales of biological processes governing quench rates (for example, protein phosphorylation or transport, RNA transcription and release) are intimately coupled to the subcellular locations at which condensates form.

In conclusion, we have proposed a general framework to quantify and interpret the rate and specificity of droplet nucleation for synthetic and endogenous biomolecular condensates. Despite the complexity of cells, we have shown that the scaling form predicted by CNT (equation (1)) can be used to interpret the nucleation rate and its dependence on supersaturation. Biomolecular and mechanical heterogeneities within the cell result in a spatially varying nucleation landscape (S^*) governing condensate formation. We have here focused on simple nucleation landscapes characterized by only one or two values of S^*—the minimum required to control the specificity of nucleation—but in full generality the spatial distribution of S^* is dictated by a nucleation landscape that is continuous (Fig. 3e, Supplementary Discussion). Nonetheless, this simplification facilitates a mathematical criterion that cells must meet in order to ensure specific nucleation. Such specific nucleation is mediated by favourable interactions between seeds and condensates, which we have shown are strongly dependent on molecular features, including amino acid patterning of IDRs. Nucleation specificity is likely to be crucial for processes such as DNA repair and gene expression, in which condensates must assemble at particular locations. Finally, the framework we have introduced will help guide efforts to design intracellular condensates for various bioengineering applications, where targeting nucleation to specific spatial locations within the cell may strongly impact the engineered functions.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03905-5.

1. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaar4382 (2017).
2. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
3. Kashchiev, D. *Nucleation* (Elsevier, 2000).
4. Auer, S. & Frenkel, D. Prediction of absolute crystal-nucleation rate in hard-sphere colloids. *Nature* **409**, 1020–1023 (2001).
5. Mitrea, D. M. et al. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *eLife* **5**, e13571 (2016).
6. Sanders, D. W. et al. Competing protein-RNA interaction networks control multiphase intracellular organization. *Cell* **181**, 306–324.e28 (2020).
7. Martin, E. W. et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).
8. Wang, J. et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699.e16 (2018).
9. Bracha, D. et al. Mapping local and global liquid phase behavior in living cells using photo-oligomerizable seeds. *Cell* **175**, 1467–1480 (2018).
10. Shin, Y. et al. Liquid nuclear condensates mechanically sense and restructure the genome. *Cell* **175**, 1481–1491.e13 (2018).
11. Wei, M. T. et al. Nucleated transcriptional condensates amplify gene expression. Nat. *Cell Biol.* **22**, 1187–1196 (2020).
12. Kato, M. et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
13. Riback, J. A. et al. Composition-dependent thermodynamics of intracellular phase separation. *Nature* **581**, 209–214 (2020).
14. Rubinstein, M. & Colby, R. H. *Polymer Physics* (Oxford Univ. Press, 2003).
15. Clouet, E. *Modeling of Nucleation Processes* (ASM Handbook, 2009).
16. Honerkamp-Smith, A. R., Veatch, S. L. & Keller, S. L. An introduction to critical points for biophysicists; observations of compositional heterogeneity in lipid membranes. *Biochim. Biophys. Acta* **1788**, 53–63 (2009).
17. Mayoral, E. & Goicochea, A. G. Hyperscaling relationship between the interfacial tension of liquids and their correlation length near the critical point. *Soft Matter* **10**, 9054–9058 (2014).
18. Lafontaine, D. L., Riback, J. A., Bascetin, R. & Brangwynne, C. P. The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.* **22**, 165–182 (2021).
19. Cho, W. K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
20. Chong, S. et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
21. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
22. Kilic, S. et al. Phase separation of 53 BP1 determines liquid-like behavior of DNA repair compartments. *EMBO J.* **38**, e101379 (2019).
23. Shevtsov, S. P. & Dundr, M. Nucleation of nuclear bodies by RNA. *Nat. Cell Biol.* **13**, 167–173 (2011).
24. Berry, J., Weber, S. C., Vaidya, N., Haataja, M. & Brangwynne, C. P. RNA transcription modulates phase transition-driven nuclear body assembly. *Proc. Natl. Acad. Sci.* **112**, E5237–E5245 (2015).
25. Aaron, J. S., Taylor, A. B. & Chew, T. L. Image co-localization—co-occurrence versus correlation. *J. Cell Sci.* **131**, jcs211847 (2018).

Methods

Cell culture

Human cell lines used in this study include U2OS (female) and Lenti-X 293T (female, Takara Bio). Lenti-X 293T cells were used only for virus production and U2OS cells were used for experiments. Cells were cultured in DMEM (GIBCO) supplemented with 10% FBS (Atlanta Biological) and penicillin and streptomycin (GIBCO) at 37 °C with 5% CO_2 in a humidified incubator. All cell lines were validated by STR profiling (ATCC) with 100% match between submitted samples and database profiles and were tested negative for mycoplasma using the Universal Mycoplasma Detection Kit (ATCC, 30-1012K).

Plasmid construction

DNA constructs used for tissue culture were cloned using In-Fusion HD cloning kit (Clonetech) in a standard reaction mixture containing 50 ng PCR-amplified inserts and 20 ng linearized pHR-SFFV backbone in a 5 µl reaction set to 50 °C for 15 min. PCR fragments were produced using a standard PCR reaction using Phusion High-Fidelity DNA Polymerase (NEB). PCR products were verified on an agarose gel and purified using PCR extraction kit (QIAGEN). Plasmids were transformed into Stellar cells (Clontech), from which single colonies were picked, grown in LB supplemented by Ampicillin for 16 h, and miniprepped (QIAGEN) following manufacturer's instructions. All cloning products were confirmed by sequencing (Genewiz). pHR-FUS_N (1–214, WT)-miRFP-TRF1 was first generated by inserting DNA fragments encoding FUS_N and miRFP670-TRF1[10] into a pHR-based vector backbone (pHR- FUSN-mCh-sspb, Addgene 122148) linearized by MluI-HF (NEB) and SbfI-HF (NEB) using following the manufacturer's instruction. For other FUS_N(nYS)–miRFP–TRF1 (n = 5, 15 and 27) and $HNRNPA1_C$–miRFP–TRF1 constructs, FUS_N in FUS_N–miRFP–TRF1 was swapped out for the DNA sequence encoding FUS_N(5YS), FUS_N(15YS), FUS_N(27YS) or $HNRNPA1_C$[9]. FUS_N(5YS) contained tyrosine to serine mutations at Y17, Y75, Y81, Y143 and Y149. FUS_N(15YS) contained tyrosine to serine mutations at Y14, Y17, Y33, Y38, Y41, Y55, Y58, Y91, Y97, Y100, Y130, Y143, Y149, Y155 and Y161. FUS_N(27YS) contained tyrosine to serine mutations at Y6, Y14, Y17, Y25, Y33, Y38, Y41, Y50, Y55, Y58, Y66, Y75, Y81, Y91, Y97, Y100, Y113, Y122, Y130, Y136, Y143, Y149, Y155, Y161, Y177, Y194 and Y208. The other constructs (pHR-FUS_N-mCherry-sspb, pHR-$HNRNPA1_C$-mCherry-sspb, pHR-NLS-iLID-EGFP-FTH1, pHR-iLID-EGFP-FTH1, FM5-NPM1-mGFP, FM5-Coilin-EYFP, FM5-mCherry-G3BP1, pHR-H2B-miRFP670, pHR-mGFP-P2A-mCherry, FM5–miRFP–53BP1, FM5-sspB-mCherry-G3BP1 DelNTF2, FM5-FXR1-miRFP, FM5-UBAP2L-miRFP, FM5-miRFP-LSM14A and FM5-iLID-mGFP-FTH1) were generated in our previous studies[6,9,10,13].

Lentiviral transduction

Lentiviruses were produced by cotransfecting Lenti-X 293T cells grown to approximately 70% confluency in 6-well plates with the transfer plasmids (1.5 µg), pCMV-dR8.91 (1.33 µg) and pMD2.G (0.17 µg) using FuGENE HD Transfection Reagent (Promega) following the manufacturer's instructions. After 2 days, 2 ml of supernatant containing viral particles was collected and filtered with a 0.45-µm filter (VWR). Supernatant was used immediately for transduction or stored at −80 °C in aliquots.

Construction of stable cell lines

U2OS cells were grown to 10–20% confluency on 35-mm glass-bottom dishes (MatTek) and 10–500 µl of filtered viral supernatant was added to the cells. Virus-containing medium was replaced with fresh growth medium 48 h post-infection. Cells infected were typically imaged no earlier than 72 h after infection.

Arsenite treatment on G3BP1-expressing cells

G3BP1-and G3BP2-knockout U2OS cells[6], expressing FM5-mCherry-G3BP1 or G3BP1 Corelet (FM5-iLID-mGFP-FTH1 and FM5-sspB-mCherry-G3BP1 DelNTF2) via lentiviral transfection, were treated with a total concentration of 400 µM sodium arsenite (Sigma-Aldrich).

Bleomycin treatment on 53BP1-expressing cells

Bleomycin (Sigma-Aldrich) was dissolved in PBS at 10 mg ml^{-1} and further diluted in growth medium to a final concentration of 10 µg ml^{-1} before being applied to U2OS cells expressing FM5–miRFP–53BP1 via lentiviral transfection.

Corelets

Corelets (core scaffolds to promote droplets) have a two-module optogenetic system that mimics the endogenous oligomerization of IDR-rich proteins to drive endogenous nucleation of membrane-less condensates, using a light-activatable high valency core[9]. The core comprises 24 human ferritin heavy chain (FTH1) protein subunits, which self-assemble to form a spherical particle of 12 nm diameter (the core), which is fused to a nuclear localization signal (NLS) and an engineered protein iLID. The second module comprises a self-interacting IDR fused to SspB, which upon blue light activation strongly heterodimerize with dissociation constant (K_d) ≈ 130 nM with iLID to form self-interacting particles.

Microscopy

Fluorescence images were taken using a spinning-disk confocal microscope (Yokogawa CSU-X1) with a Nikon 100× oil immersion objective (CFI Apo TIRF, NA 1.49) and an Andor iXon Ultra DU-897 EMCCD camera on a Nikon Eclipse Ti-E body. Samples were maintained at 37 °C and 5% CO_2 with a stage top incubator (Okolab). A laser combiner (Nikon LU-NV) was used to excite mGFP, EGFP and EYFP and for activating iLID with 488 nm laser light at an excitation power of approximately 0.1 W cm^{-2} as measured with a microscope slide photodiode power sensor (PM100D, Thorlabs). 561 and 640 nm lasers were used for imaging mCherry and miRFP, respectively. Phase-separation dynamics for the Corelet system were captured by imaging in the mCherry channel with global activation performed by the 488 nm laser line. Telomere-targeted seeds were captured in the miRFP channel. In the case where a fast frame rate was desirable, a frame interval of 60 ms was used with a 256 × 256 pixel region of interest. All image acquisition was performed using Nikon NIS-Elements AR software. All image analysis was performed using ImageJ or custom-built MATLAB scripts. A laser scanning confocal microscope (Nikon A1) was used only for fluorescence correlation spectroscopy (FCS) experiments.

Absolute concentrations for Corelets components

Absolute concentrations of Corelets components, that is, core and IDR, were estimated using FCS, as reported previously[9]. Data for concentration of proteins were obtained using 30 s FCS measurement time. The measurements were performed on HEK293 cells expressing pHR- FUS_N-mCh-sspb using a Nikon A1 with an oil immersion objective (Plan Apo 60×/1.4). All measurements and data analysis were performed using the SymPhoTime Software (PicoQuant). The autocorrelation function for simple diffusion is:

$$G(\lambda) = G(0)\left(1+\frac{\lambda}{\lambda_D}\right)^{-1}\left(1+\frac{\lambda}{\eta^2\lambda_D}\right)^{-1/2}, \quad (2)$$

here $G(0)$ is the magnitude at short timescales, λ is the lag time, λ_D is the half decay time, and η is the ratio of axial to radial of measurement volume $\eta = \omega_z/\omega_{xy}$. Here, ω_{xy} = 2.3 µm and η = 3.1, as determined by fluophore dye Atto 550 in water. The parameters $G(0)$ and λ_D are optimized in the fit and are used to determine the molecule concentration $C = (\pi^{3/2}\omega_{xy}^2\omega_z G(0))^{-1}$. mCherry fluorescence was converted to absolute concentration using FCS. GFP fluorescence conversion was done by determining the mCherry:GFP fluorescence ratio, which was determined by equimolar expression of mCherry and GFP monomers in HEK293 cell using auto-catalytic P2A containing construct

mGFP–P2A–mCherry. Then, further fluorescence calibration was done between the Nikon A1 microscope and spinning-disk confocal microscope on a Nikon Eclipse Ti-E body (Yokokgawa CSU-X1), using WT U2OS cells expressing mGFP–P2A–mCherry.

Phase diagram construction

The nucleus or cytoplasm boundary in each cell was determined by applying an automated image segmentation MATLAB code for the Gaussian-filtered image of core component (EGFP channel) before phase separation. Then, mean EGFP and mCherry fluorescence within the segmented region were background-subtracted and translated to absolute concentration via the FCS-based concentration estimation, thus enabling estimation of mean core absolute concentration and core-to-IDR ratio. The mean values are shown in Fig. 1g and Extended Data Fig. 4b. The binodal boundary is defined as the boundary to separate whether any visible droplets are observed within our experimental timescale (~5 min) after light activation. The spinodal region is defined as the region in which the interconnected network-like growth is observed, as shown in Extended Data Fig. 3.

Quantification of nucleation rate

To quantify the nucleation rate J, the number of droplets per unit volume—that is, droplet density ρ, was first measured as a function of time t. The nucleus or cytoplasm boundary in each cell was determined by applying an automated image segmentation MATLAB code for the Gaussian-filtered image of core component (EGFP channel) before nucleation. Then, the droplets were automatically segmented by a single intensity threshold based on the bimodal fluorescent histogram of IDR component (mCherry channel) within the segmented nucleus or cytoplasm. The number of droplets were then counted. The depth of focus (~1 μm) of confocal images was used to estimate the analysed volume. For Fig. 1e, Extended Data Fig. 3b, the measures of centre for the shaded error bars are the mean and the error bars show standard deviation. For Fig.3c (bottom), the data points are the mean and the error bars show the standard deviation. The nucleation rate was obtained as

$$J = \frac{\mathrm{d}\rho}{\mathrm{d}t}, \tag{3}$$

at $t = t_0$ from fitting the data with the equation

$$\rho(t) = \rho_0\left(1 - \mathrm{e}^{-\frac{t-t_0}{\tau}}\right) \tag{4}$$

The data points in Figs. 1h, 2c, 3b, Extended Data Figs. 1b (ii), 7, 9d, 10b are the best fit parameters. The error bars show standard errors with 95% confidence intervals of the fits. For Fig. 3d, the data points and error bars show the mean and standard deviation of the best fit parameters, respectively.

Quantification of supersaturation

Supersaturation is defined as $S = \log(C_{\mathrm{dil}}/C_{\mathrm{sat}})$, where C_{dil} is the initial core concentration in the dilute phase of the nucleus (or cytoplasm) and C_{sat} is the saturating core concentration, determined from the dilute phase after nucleation ends, since $C_{\mathrm{sat}} \approx C_{\mathrm{dil}}$ in the steady state. For the quantification of supersaturation, the nucleus or cytoplasm boundary in each cell was first determined by applying an automated image segmentation MATLAB code for the Gaussian-filtered image of core component (EGFP channel) before nucleation. The initial core concentration C_{dil} was estimated from the average fluorescence intensity of the background-subtracted image in the segmented region using the FCS-based calibration. For estimating the core concentration in the diluted phase after nucleation in the steady state, the droplets were automatically segmented by a single intensity threshold based on the bimodal fluorescent histogram of core component within the segmented nucleus or cytoplasm. To accurately determine the concentration within the dilute phase, morphological erosion was performed for the segmented droplets, such that three pixels (≈ 0.48 μm) near the droplet interface were excluded from the analysis. The core concentration in the diluted phase was estimated from the average fluorescence intensity of the background-subtracted image in the segmented diluted region using the FCS-based calibration. The nominal supersaturation $S_{\mathrm{nom}} = \log(C_{\mathrm{tot}}/C_{\mathrm{sat}})$ in Fig. 3b, Extended Data Fig. 10b was estimated from the concentrations of the total composition and diluted phase of a reporter protein (for example, NPM1–mGFP for nucleoli) after the nucleation process ends. The error bars in Figs. 1h, 2c, 3d, Extended Data Figs. 6, 7, 10b are the standard deviation.

Model fit

The data of J against S were fit to a functional form predicted by CNT (equation (1)) by minimizing χ^2. For Figs. 1h, 2c, 3b, Extended Data Fig. 7, the model fit parameters (that is, S^* and κ) and reduced chi-squared (χ^2_{red}) values are shown in Extended Data Fig. 5.

Quantification of relative nucleation probability

To quantify the relative nucleation probability $p^{\mathrm{rel}}_{\mathrm{nuc}}(x)$ in Extended Data Fig. 2, defined by the number of formed droplets per unit volume at position x, U2OS cells expressing $\mathrm{FUS_N}$–Corelets with or without $\mathrm{FUS_N}$ (WT)-miRFP-TRF1 were sequentially light-activated five times. The image where the nucleation is saturated after 9.9 s (Extended Data Fig. 2a), 58.7 s (Extended Data Fig. 2b) and 32.1 s (Extended Data Fig. 2b) of light activation was selected for each activation, and the five images were spatially aligned on the basis of the centre of the nucleus. We then counted the total number of formed condensates in a volume of approximately 5 $\mu\mathrm{m}^3$ with a focus on a position x by iterative activation and averaged it in space to calculate $p^{\mathrm{rel}}_{\mathrm{nuc}}(x)$.

Colocalization analysis

The colocalization coefficient between two images was estimated by calculating PCC defined as

$$\mathrm{PCC} = \frac{\Sigma_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\Sigma_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{5}$$

where x_i and $\bar{x}$ are the ith pixel intensity and mean values in the segmented first image (that is, colour 1), respectively. Similarly, y_i and $\bar{y}$ are the ith pixel intensity and mean values in the second image (that is, colour 2), respectively. The perfect correlation, no correlation and anti-correlation give PCC values of 1, 0 and −1, respectively. The correlation in the whole nucleus was calculated. The data points in Fig. 2d show the mean. The error bars show standard deviation.

Reproducibility

A representative of at least three independent experiments is shown in Figs. 1a, d, 2b, 3a, c (top), Extended Data Fig. 1a, 2, 3, 4a, 8, 9, 10a.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Source data for Figs. 1–3 are provided with the paper. All other data are available from the corresponding authors upon reasonable request.

Code availability

Custom code used to process and analyse the images, as detailed in the Methods, is available from the corresponding authors upon reasonable request.

Acknowledgements We thank members of the Brangwynne laboratory for helpful discussions and comments on this manuscript. We thank laboratory members Y. Kim, M. Walls and D. Bracha for providing plasmids. This work was supported by the Howard Hughes Medical Institute, and grants from the AFOSR MURI (FA9550-20-1-0241), and the Princeton Center for Complex Materials, an MRSEC (NSF DMR-2011750). We also thank Princeton's School of Engineering and Applied Sciences for supporting this work through a Focused Research Team award. S.F.S. is supported by JSPS Postdoctoral Fellowships for Research Abroad. P.R. is supported by the NSF through the Center for the Physics of Biological Function (PHY-1734030), by the French Government programme 'Investissements d'Avenir' managed by the French National Research Agency (ANR-16-CONV-0001) and by the Excellence Initiative of Aix-Marseille University—A*MIDEX.

Author contributions S.F.S., P.R., D.W.S., M.P.H. and C.P.B. designed the research; S.F.S. and D.W.S. performed the experiments; S.F.S. analysed the data; P.R. and M.P.H took the lead in the theoretical formalism; S.F.S., P.R., M.P.H. and C.P.B. wrote, and all authors reviewed and edited, the paper.

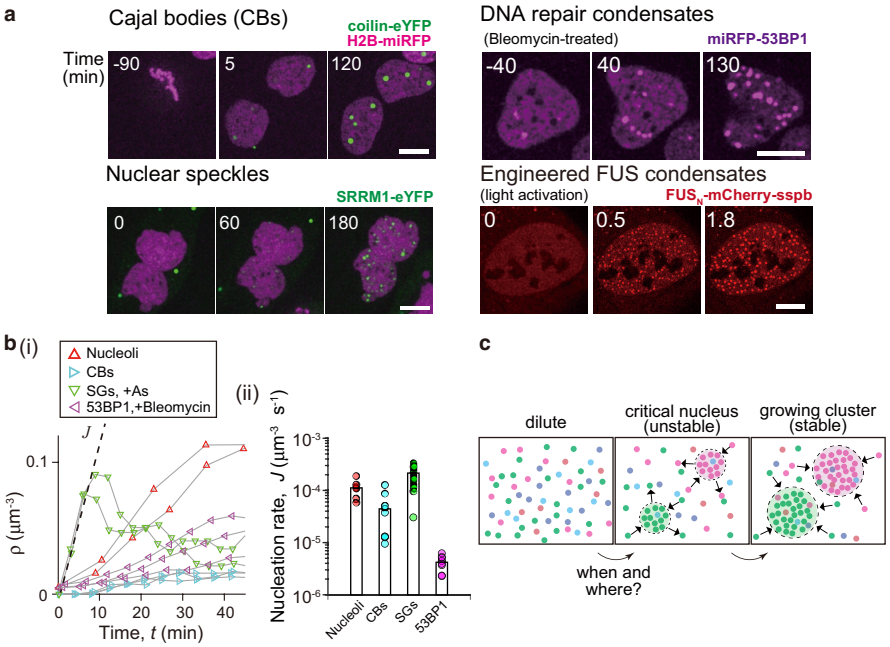
Competing interests C.P.B. is a founder and consultant of Nereid Therapeutics. S.F.S., P.R., D.W.S. and M.P.H. declare no competing interests.

Additional information

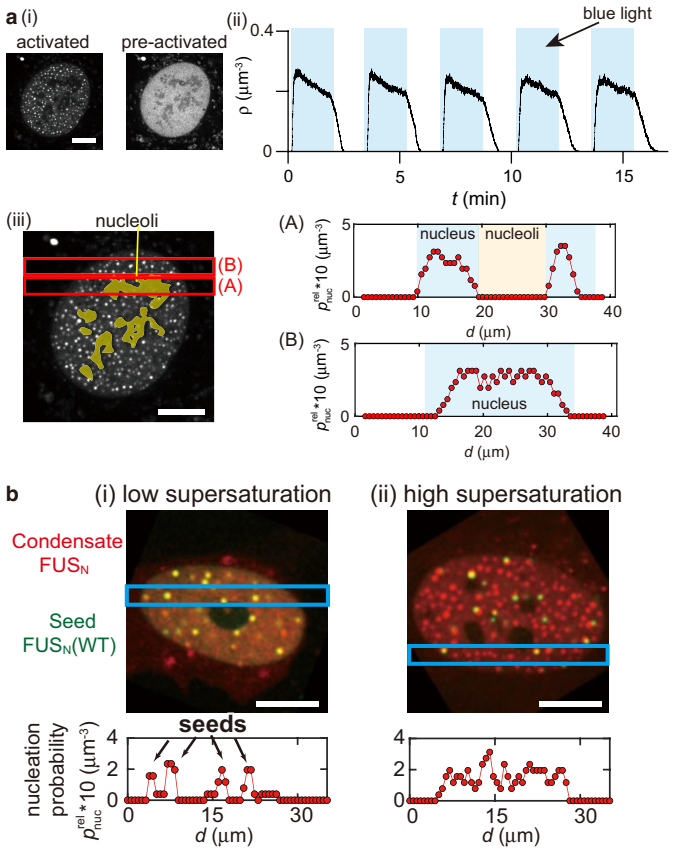
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03905-5.

Correspondence and requests for materials should be addressed to Clifford P. Brangwynne.

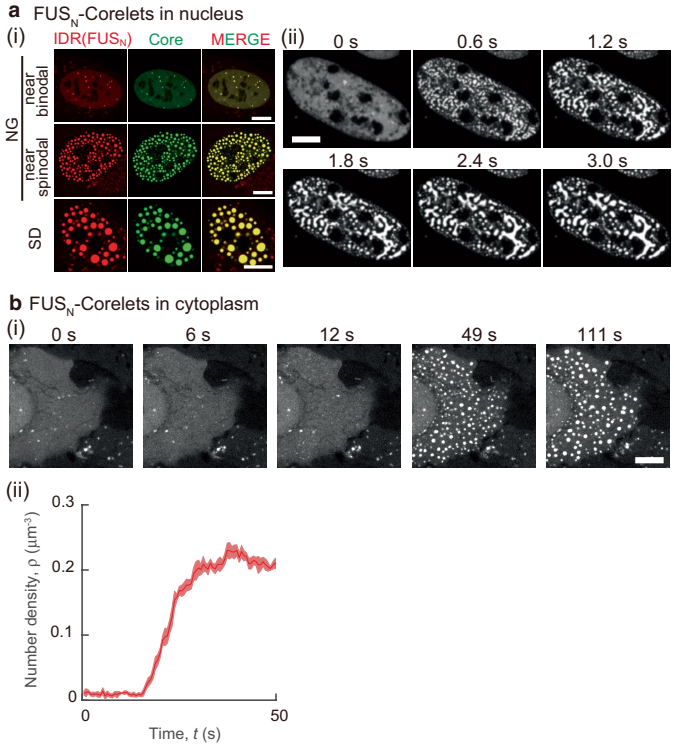
Peer review information *Nature* thanks Mike Sleutel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



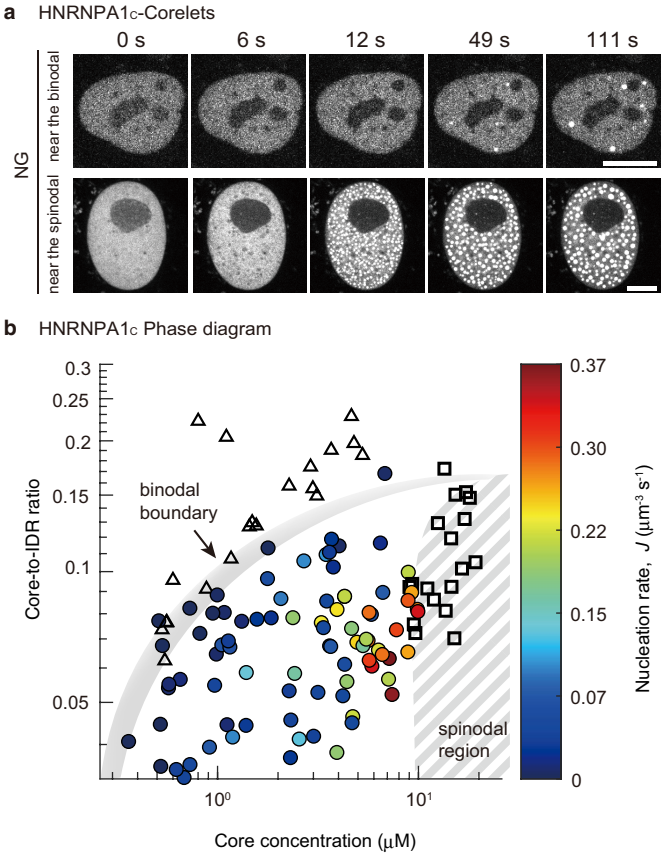
Extended Data Fig. 1 | Nucleation process of various endogenous and biomimetic condensates. a, Time-lapse images of U2OS cells show the nucleation of Cajal bodies (coilin-EYFP) and nuclear speckles (SRRM1-EYFP) in mitosis, DNA repair condensates (miRFP-53BP1) upon 10 μg/ml bleomycin treatment, and engineered FUS condensates. Scale bars, 10 μm. **b**, (i) Number density of condensates, ρ, as a function of time, t, for nucleoli and Cajal bodies in mitosis, stress granules (SGs) upon 400 μM As treatment, and 53BP1 condensates upon 10 μg/ml bleomycin treatment. The nucleation rate, J, is quantified by the slope. Here, $t = 0$ is $t_0 - 5$ (min). (ii) Mean nucleation rate, J, and its standard deviation for nucleoli (n = 6), Cajal bodies ($n = 9$), SGs ($n = 24$) and DNA repair condensates ($n = 9$); n = number of cells. **c**, Schematic diagram of the nucleation process of intracellular condensates.



Extended Data Fig. 2 | Quasi-1D nucleation probability. a, (i) Fluorescence images of U2OS cells expressing HNRNPA1$_C$-Corelets before and after blue light activation. Scale bar, 10 µm. (ii) Droplet number density as a function of time upon iterative blue-light activation and deactivation. (iii) Quasi-1D nucleation probability p_{nuc}^{rel} in the indicated region which include nucleoli (A) or not (B). **b**, Quasi-1D nucleation probability p_{nuc}^{rel} of U2OS cells expressing FUS$_N$-Corelets and FUS$_N$-miRFP-TRF1 in the indicated region ((i) low and (ii) high supersaturation), calculated from five successive activation cycles. Scale bar, 10 µm.



Extended Data Fig. 3 | Photo-activated phase separation in nucleus and cytoplasm. a, Photo-activated phase separation in nucleus. (i) Confocal images of U2OS cells with different expression levels of FUS_N Corelets (FUS_N IDRs (red) and Cores (green)) after light-activation. The cells display nucleation and growth (NG) regimes between the binodal boundary and spinodal region, and spinodal decomposition (SD). Scale bars, 10 μm. (ii) Connected network-like growth and coarsening akin to spinodal decomposition. Scale bar, 10 μm. **b**, Photo-activated phase separation in cytoplasm. (i) Time-lapse confocal images of photo-activated U2OS cells expressing FUS_N-Corelets composed of Core without NLS and FUS_N-mCh-sspb. Scale bars, 10 μm. (ii) Time change of light-induced droplet number density, ρ. The shaded error bars show standard deviation.



Extended Data Fig. 4 | Phase separation behaviors of HNRNPA1$_C$-Corelets. **a**, Time-lapse confocal images of photo-activated U2OS cells with different expression levels of HNRNPA1$_C$-Corelets. Nucleation growth (NG) regime near the binodal boundary (top) and spinodal region (bottom). Scale bars, 10 μm. **b**, Phase diagram of HNRNPA1$_C$-Corelets as functions of Core concentration and Core-to-IDR ratio (n = 161). Solid circles exhibit cells where nucleation growth is observed, while empty triangles and squares show cells where no phase separation and spinodal decomposition are observed, respectively. The colours of solid circles indicate the observed nucleation rate, J. n: number of cells.

a

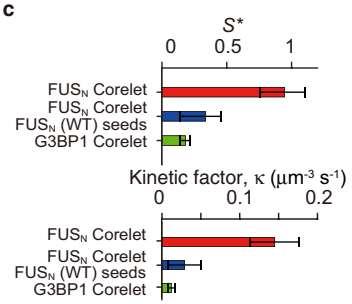
	FUS_N-Corelets	FUS_N Corelets plus FUS_N (WT) seeds	$HNRNPA1_C$-Corelets	G3BP1 corelets
$\kappa\ [\mu m^{-3} \cdot s^{-1}]$	0.34 ± 0.03	0.07 ± 0.01	0.47 ± 0.04	0.02 ± 0.00
S^*	0.97 ± 0.08	0.44 ± 0.04	1.04 ± 0.02	0.26 ± 0.01
χ^2_{red}	0.55	1.62	2.70	39.16

b

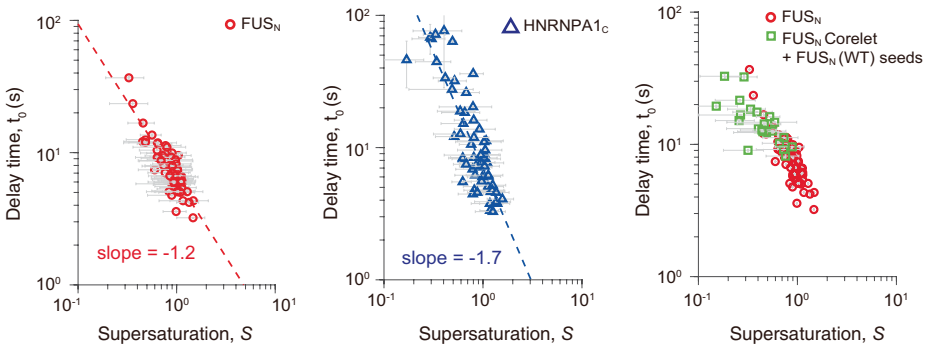
	FUS_N-Corelets, 1/24 < f < 1/18	FUS_N-Corelets, 1/18 ≤f< 1/12	FUS_N-Corelets, 1/12 ≤f< 1/6
$\kappa\ [\mu m^{-3} \cdot s^{-1}]$	0.28 ± 0.07	0.33 ± 0.01	0.26 ± 0.01
S^*	1.05 ± 0.24	0.99 ± 0.24	0.61 ± 0.01
χ^2_{red}	0.34	0.29	6.83

	$HNRNPA1_C$-Corelets, 1/24 < f < 1/18	$HNRNPA1_C$-Corelets, 1/18 ≤f< 1/12	$HNRNPA1_C$, 1/12 ≤f< 1/6
$\kappa\ [\mu m^{-3} \cdot s^{-1}]$	0.76 ± 0.11	0.46 ± 0.03	0.29 ± 0.03
S^*	1.59 ± 0.12	0.93 ± 0.08	0.64 ± 0.06
χ^2_{red}	1.39	1.25	2.85

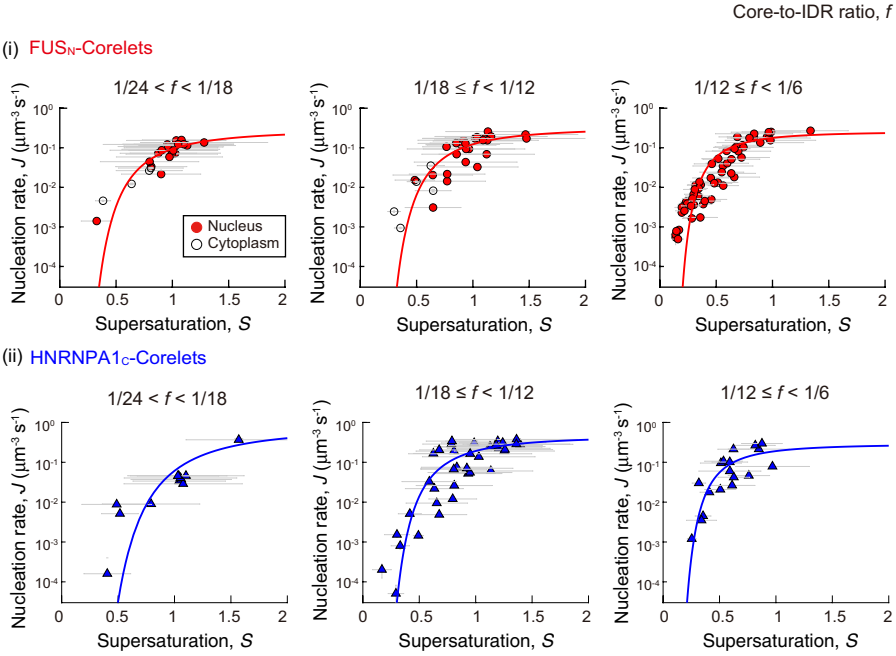
c



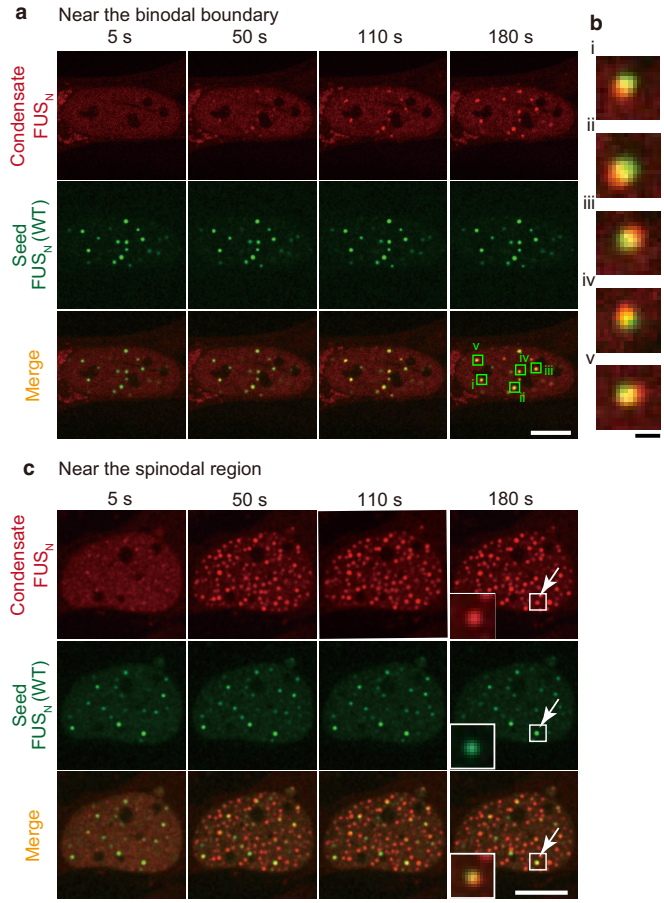
Extended Data Fig. 5 | Model fit parameters and chi-square values. **a**, **b**, Model fit parameters of S^*, κ, χ^2_{red} for various biomimetic condensates. The errors show the standard errors with 68% confidence intervals of the fits by equation (1). The parameter $\chi^2_{red} \gg 1$ indicates a poor model fit while $\chi^2_{red} \lesssim 1$ indicates a good model fit. **c**, S^* (top) and κ (bottom) obtained from fits to equation (1) for FUS_N Corelets with (n = 23) or without (n = 76) FUS_N (WT) seeds and G3BP1 corelets ± 400 μM As (n=45). Error bars represent standard fit errors.



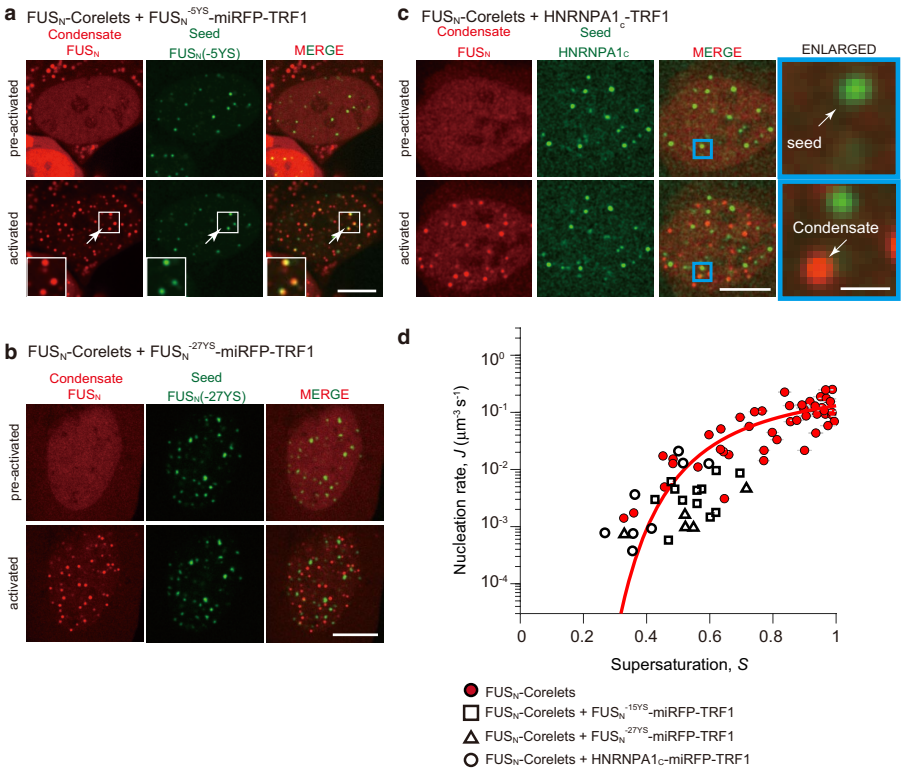
Extended Data Fig. 6 | Delay time against supersaturation. Delay time, t_0, against supersaturation, S, for FUS$_N$-Corelets ($n = 70$), HNRNPA1$_C$-Corelets ($n = 60$), and FUS$_N$ Corelets plus FUS$_N$ (WT) seeds ($n = 23$) with $1/24 < f < 1/10$. The dashed lines show the best power-law fits. Error bars for t_0 and S show the standard errors of the fits and standard deviation, respectively. n: number of cells.



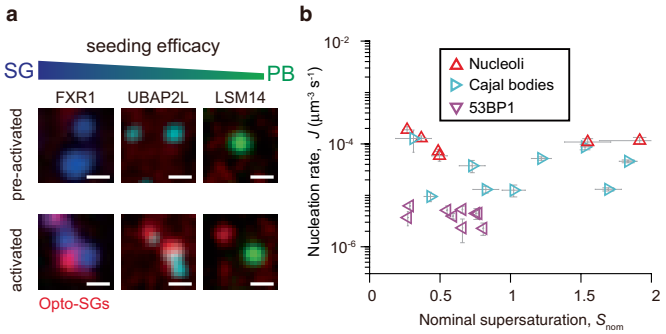
Extended Data Fig. 7 | J against S with various ranges of Core-to-IDR ratio. Nucleation rate, J, plotted against the degree of supersaturation, S, for (i) FUS$_N$-Corelets ($n = 26$ for $1/24 < f < 1/18$, $n = 33$ for $1/18 \le f < 1/12$ and $n = 56$ for $1/12 \le f < 1/6$) and (ii) HNRNPA1$_C$-Corelets ($n = 12$ for $1/24 < f < 1/18$, $n = 38$ for $1/18 \le f < 1/12$ and $n = 18$ for $1/12 \le f < 1/6$). Error bars for J and S show the standard errors of the fits and standard deviation, respectively. See Extended Data Fig. 5 for model fit parameters and χ^2_{red}. n: number of cells.



Extended Data Fig. 8 | Seeded nucleation near the binodal boundary and spinodal region. Time-lapse confocal images of U2OS cells expressing FUS_N-Corelets and FUS_N-miRFP-TRF1 near the binodal boundary (**a**) and spinodal region (**c**). The insets show the regions indicated by the arrows. Scale bars, 10 μm. **b**, The seeded condensates, which are indicated in **a**, are enlarged. Scale bar, 1 μm.



Extended Data Fig. 9 | Nucleation of FUS_N-Corelets with tyrosine-mutated FUS_N-seeds. a–**c**, Fluorescent images of U2OS cells expressing FUS_N-Corelets and FUS_N (−5YS, −27YS)-miRFP-TRF1 or $HNRNPA1_C$-miRFP-TRF1 before and after blue light activation. The insets and enlarged images show the regions indicated by the squares. Scale bars, 10 µm. Scale bar of the enlarged image, 1 µm. **d**, Nucleation rate, J, plotted against the degree of supersaturation, S, for FUS_N-Corelets ($1/24 < f < 10$) and FUS_N-Corelets ($1/30 < f < 10$) with FUS_N (−15YS, −27YS)-miRFP-TRF1 ($n_{15YS}, n_{27YS} = 12,5$) or $HNRNPA1_C$-miRFP-TRF1 ($n = 8$). n: number of cells.



Extended Data Fig. 10 | Specificity of the nucleation of endogenous condensates. a, Confocal images of G3BP KO cells expressing G3BP1 corelets and iRFP-tagged FXR1/UBAP2L/LSM14 before and after blue light activation in the absence of As. Scale bars, 1 µm. **b**, Nucleation rate of endogenous condensates against nominal supersaturation. Nucleation rate, J, plotted against the nominal supersaturation, S_{nom}, for nucleoli ($n = 6$) and Cajal bodies ($n = 9$) in mitosis and 53BP1 condensates ($n = 9$) upon 10 µg/ml bleomycin treatment. n: number of cells.

nature research

Corresponding author(s): Clifford P. Brangwynne

Last updated by author(s): Jul 22, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Nikon Elements AR and SymPhoTime Software were used to collect data.
Data analysis	All data analysis was done with ImageJ 1.50e or custom code written in MATLAB R2019a.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Source data for Figs.1-3 are provided with the paper. All other data are available from the corresponding authors upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical method was used to predetermine sample size. Sample size was adequate as the differences between experimental groups were significant and reproducible.
Data exclusions	No data were excluded from the analysis.
Replication	All the data were collected from at least three independent replicates. All attempts at replication was successful.
Randomization	No randomization was preformed as the study has been highly reproducible e (e.g. day to day, locations used on multi-well plate).
Blinding	Investigators were not blinded for data collection as analysis is required to assure expression levels span required dynamic range as is typical for biophysical experiments. Also, dead cells and dividing cells were not collected. Investigators were blinded for data analysis as the analysis was performed in an automated way.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	U2OS cells (Tom Muir Lab, Princeton Univ., RRID:CVCL_0042), 293T-LentiX cells (Takara Bio USA, Cat# 632180), HEK293 cells (Marc Diamond Lab,UT Southwestern, RRID:CVCL_0045) and G3BP1/2 KO cells (Kedersha et al., 2016, Nancy kedersha and Paul Anderson, Harvard Univ.) were used.
Authentication	All cell lines were validated by STR profiling (ATCC) with 100% match between submitted samples and database profiles.
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	None.