Published in final edited form as:
Cell Rep. 2023 August 29; 42(8): 112955. doi:10.1016/j.celrep.2023.112955.

An evolutionarily nascent architecture underlying the formation and emergence of biomolecular condensates

Nima Jaberi-Lashkari[1,3], **Byron Lee**[1,3], **Fardin Aryan**[1], **Eliezer Calo**[1,2,4,*]

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[3]These authors contributed equally

[4]Lead contact

SUMMARY

Biomolecular condensates are implicated in core cellular processes such as gene regulation and ribosome biogenesis. Although the architecture of biomolecular condensates is thought to rely on collective interactions between many components, it is unclear how the collective interactions required for their formation emerge during evolution. Here, we show that the structure and evolution of a recently emerged biomolecular condensate, the nucleolar fibrillar center (FC), is explained by a single self-assembling scaffold, TCOF1. TCOF1 is necessary to form the FC, and it structurally defines the FC through self-assembly mediated by homotypic interactions of serine/glutamate-rich low-complexity regions (LCRs). Finally, introduction of TCOF1 into a species lacking the FC is sufficient to form an FC-like biomolecular condensate. By demonstrating that a recently emerged biomolecular condensate is built on a simple architecture determined by a single self-assembling protein, our work provides a compelling mechanism by which biomolecular condensates can emerge in the tree of life.

In brief

Jaberi-Lashkari et al. present a model for the evolution of biomolecular condensates by showing that a recently emerged condensate relies on a self-assembling scaffold rather than on many co-interacting scaffolds and that emergence of this single scaffold can explain the emergence of this condensate in the tree of life.

---

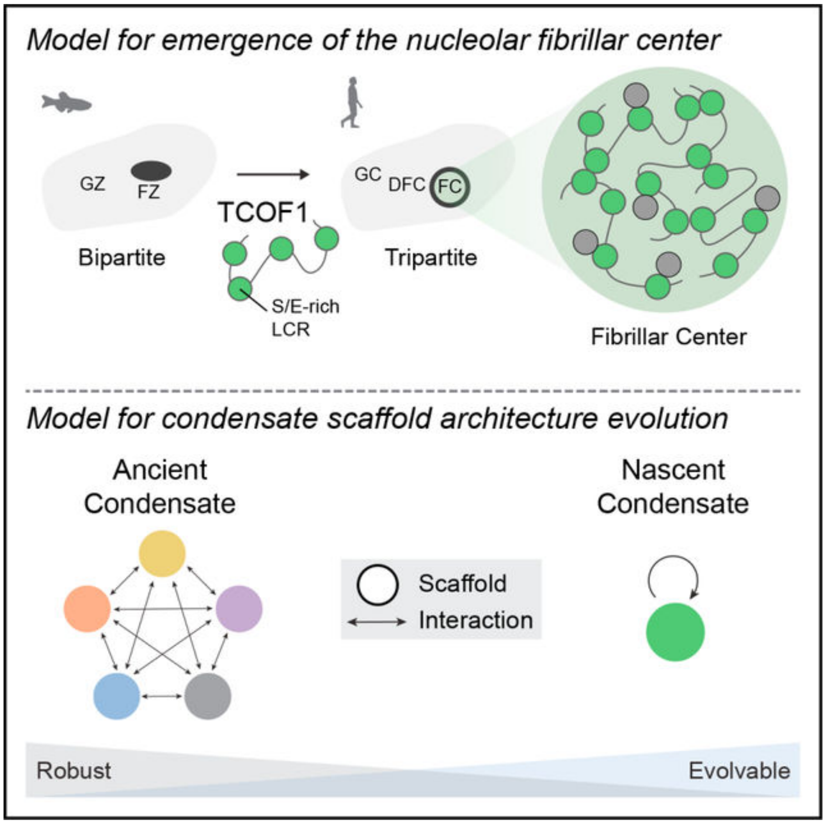
[*]Correspondence: calo@mit.edu.

AUTHOR CONTRIBUTIONS
Conceptualization, N.J.-L. and B.L.; methodology, N.J.-L., B.L., and F.A.; software, N.J.-L. and B.L.; validation, N.J.-L. and B.L.; formal analysis, N.J.-L. and B.L.;investigation,N.J.-L.,B.L.,andF.A.;datacuration,N.J.-L.andB.L.;writing – original draft preparation, N.J.-L., B.L., and F.A.; writing – review & editing, N.J.-L., B.L., and E.C.; visualization, N.J.-L. and B.L.; supervision, N.J.-L. and B.L.; resources, E.C.; funding acquisition, E.C.

DECLARATION OF INTERESTS
B.L. is currently affiliated with Sequome, Inc.

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112955.

Graphical Abstract:



INTRODUCTION

Biomolecular condensates, or condensates, such as nucleoli and stress granules, are membraneless assemblies of biomolecules, such as proteins and RNAs, implicated in key cellular processes from ribosome biogenesis to stress response.[1–6] The architecture of condensates is thought to rely on collective, interdependent interactions between many components.[7,8] However, it remains unclear how such condensates emerge during evolution. This question highlights an unexplored link between the underlying architecture of condensates and their evolution.

One of the most recently emerged condensates exists within the nucleolus, a membraneless organelle composed of distinct, co-existing condensates that play sequential roles in ribosome biogenesis.[2,9] In humans, nucleoli are tripartite, consisting of the fibrillar center (FC), nested within the dense fibrillar component (DFC), which is nested within the granular component (GC).[2,9] Although nucleoli coordinate ribosome biogenesis across eukaryotes, their tripartite architecture is not ubiquitous. Many other eukaryotes possess bipartite nucleoli, which lack an FC-like subcompartment,[10,11] and only contain GC-like and DFC-like subcompartments called granular and fibrillar zones (GZs and FZs), respectively. How the underlying structure of the FC relates to its emergence in the bipartite-to-tripartite transition is a long-standing question, subject to speculation.[10]

To understand the relationship between the architecture and emergence of condensates, we used the FC, a recently emerged condensate with an unknown underlying structure, as a case study. Thought to have emerged in amniotes,[11] the FC may have a more evolutionarily “nascent” architecture than more highly studied condensates such as stress granules and the nucleolar GC, which exist in species that diverged from humans over a billion years ago. While the scaffold(s) of the FC are unknown, early electron microscopy (EM) studies demonstrated that the FC structure is sensitive to protease but not RNase, suggesting that protein, and not RNA, is necessary for the underlying structure of the FC.[12]

Consistently, we previously observed that the FC protein TCOF1 formed a condensate that recruited other FC proteins when targeted outside of the nucleolus.[13] Furthermore, these extra-nucleolar TCOF1 assemblies do not accumulate nascent RNA (Figure S1). These results, along with observations that protein is required for FC structure,[12] led to our hypothesis that TCOF1 may scaffold the FC and may thus provide insight into its emergence.

RESULTS

The FC, DFC, and GC can be discerned by their appearance by EM (Figure 1A). To test the importance of TCOF1 for FC structure, we made a degron-based loss-of-function system for TCOF1 in HeLa cells to assay the presence of the FC using EM (Figures 1B, S2A, and S2B).[14] TCOF1 expression in TCOF1-degron cell lines was reduced to undetectable levels within 4 days of dTAG-13 treatment (Figures S2C–S2E). Moreover, TCOF1 degradation reduced nascent nucleolar RNA levels (Figure S2F), consistent with previous studies.[15]

TCOF1 degradation resulted in ablation of the FC in two separate TCOF1-degron cell lines (Figures 1C and S3A). These results were not due to dTAG-13 off-target effects, as we observed no effect on wild-type (WT) cells, which lack the degron (Figures 1C and S3A). Importantly, TCOF1 degradation did not broadly compromise nucleolar integrity, as we still detected the GC and DFC as seen in WT and DMSO-treated TCOF1-degron cells (Figures 1C and S3A). Moreover, the phenotype of TCOF1 degradation was distinct from known effects of rRNA transcription inhibition with actinomycin D or CX-5461, neither of which abolishes the FC.[16–19] The distinction between TCOF1 loss and rRNA transcriptional inhibition confirms that TCOF1, rather than its role in rRNA transcription, is important for FC formation. Thus, these results demonstrate that TCOF1 is specifically required for the assembly of the FC.

When expressed in HeLa cells, mEGFP-TCOF1 formed nucleolar condensates, many of which were larger than normal FCs (Figure S3B). Strikingly, FC components maintain co-localization with TCOF1 condensates regardless of their size and dilute across larger TCOF1 condensates (Figures 1D and S3B). In contrast, DFC, GC, and nucleolar rim components remain excluded from TCOF1 condensates (Figure 1D). These results suggest that in cells, TCOF1 condensates maintain an FC identity irrespective of their size (Figure 1E). This behavior is not observed for another FC protein, RPA43 (Figure S3C). Moreover, in TCOF1-expressing cells, nascent RNA is synthesized on the periphery of TCOF1 condensates and progresses into the GC (Figures S3D and S3E), consistent with normal rRNA transcription around the FC.[20,21] Collectively, the observations that TCOF1 is necessary for FC formation and that, specifically, TCOF1 dictates the size, shape, and form of FCs in cells indicate that TCOF1 structurally defines the FC.

Given this, we wanted to understand how this single protein mediates assembly of this recently emerged condensate. The observations that TCOF1 forms condensates when targeted outside of the nucleolus[13] and that varying TCOF1 levels cause concomitant changes in FC size led us to hypothesize that TCOF1 may be capable of self-assembly rather than assembly dependent on collective interactions of several components. Notably, TCOF1 self-assembly would be in contrast to assembly of other condensates, which rely on multiple interacting components,[7] and may thus represent a more evolutionarily nascent architecture for condensates in cells.

To determine if TCOF1 self-assembles inside human cells, we used a biophysical approach that determines whether a protein relies on other cellular components for its assembly.[7] The relative concentrations of TCOF1 in the dense phase (within the assembly) and the dilute phase (outside of the assembly) determine its free energy of transfer ($\Delta G^{transfer}$), a measure inversely related to the favorability for entering the assembly. Upon increased expression, if TCOF1 assembly is dependent on other component(s), it will reach a point where it saturates the dense phase and accumulates in the dilute phase. This behavior is referred to as multi-component assembly and is reflected in an increase in $\Delta G^{transfer}$ and little/no change in dense phase size (Figure 2A). On the other hand, if TCOF1 is capable of self-assembly, it will partition into the assembly across all expression levels. This behavior is referred to as single-component assembly, in which $\Delta G^{transfer}$ remains constant and dense phase size increases with expression (Figure 2A).

Using this framework, we observed that although DFC and GC scaffolds Fibrillarin and Nucleophosmin (NPM1)[2,20,22] displayed multi-component assembly in cells, TCOF1 displayed single-component assembly. Partitioning of both Fibrillarin and NPM1 became unfavorable upon increased expression, as seen by increased dilute phase in cells (Figures 2B and S4), and both proteins exhibited little/no increase in their dense phase size (Figures 2B and S4). This observation is consistent with assembly of Fibrillarin and NPM1 depending on other nucleolar components.[2,7,22] Furthermore, this assembly behavior is consistent with that of scaffolds of p-bodies, stress granules, and Cajal bodies.[7] In contrast, TCOF1 maintained its partitioning, and increased in dense phase size, across ~3 orders of magnitude of expression (Figure 2C).

That TCOF1 assembles on its own inside cells is a unique example of a natural protein capable of single-component assembly in an intracellular condensate. We reasoned that if TCOF1 self-assembly is necessary for its ability to structurally define the FC, then TCOF1 self-assembly underlies FC formation. Such a direct relationship between self-assembly of a single protein and the structure of a condensate could provide a simple explanation for the emergence of this condensate.

To test if TCOF1 self-assembly underlies FC structure, we sought to determine the molecular features of TCOF1 that mediate its self-assembly and to test if specific perturbations to its self-assembly affect its ability to structurally define the FC in cells. A majority of TCOF1's ~1,500 amino acids are in a repeat domain consisting of 16 evenly spaced S/E-rich low-complexity regions (LCRs) (Figure S5A). Given the importance of valency for condensate formation,[23] we hypothesized that this repeat domain may be involved in its assembly. A TCOF1 mutant with this region deleted, TCOF1Δ83–1170, failed to assemble in a single-component manner (Figure 3A), showing that the repeat domain is essential for single-component assembly. Moreover, serial truncations of the repeat domain resulted in eventual loss of single-component assembly (Figures 3A, S5B, and S5E). This could not be explained by the direction of truncation (Figures 3A, S5B, and S5E) or by any repeat domain segment shared between constructs capable of single-component assembly (Figures S5C and S5E). Given that the region necessary for single-component assembly could not be mapped to one location within this repeat domain (Figure S5D), we reasoned that it is dispersed along the sequence.

We hypothesized that the S/E-rich LCRs are the dispersed feature necessary for single-component assembly of TCOF1. Consistently, we observed a strong relationship between the number of S/E-rich LCRs present in each mutant and whether it exhibited single-component or multi-component assembly (Figure S5D). Mutants with ≥9 copies of S/E-rich LCRs maintained single-component assembly indistinguishable from WT TCOF1, while those with ≤5 copies failed to assemble in a single-component manner (Figure S5D). To test the role of the S/E-rich LCRs in TCOF1 assembly, we generated a TCOF1 mutant with all but 5 S/E-rich LCRs specifically deleted (TCOF1ΔSE83–785). We observed that TCOF1ΔSE83–785 failed to assemble in a single-component manner (Figures 3B and S5D), showing that TCOF1 self-assembly requires its S/E-rich LCRs.

In addition to mediating self-assembly, the S/E-rich LCRs of TCOF1 are required for its ability to structurally define the FC. When assessing the ability of the truncation mutants to expand the FC, mutants with more S/E-rich LCRs excluded the GC (assessed by MPP10), while those with fewer S/E-rich LCRs did not (Figure S5F), suggesting that the ability of TCOF1 to assemble is intertwined with its ability to structurally define the FC. If so, specific perturbations to TCOF1 assembly by deletion of the S/E-rich LCRs should disrupt its ability to structurally define the FC. Indeed, a mutant with all 16 S/E-rich LCRs deleted (TCOF1ΔSE) failed to exclude the GC marker MPP10 (Figure 3C; see STAR Methods), similar to TCOF1 mutants that failed to assemble (Figure S5F), showing that the S/E-rich LCRs are necessary for TCOF1 to structurally define the FC. Together, these results argue that TCOF1 self-assembly structurally defines the FC, showing that self-assembly of a single scaffold can underlie the structure of an endogenous condensate.

The coupling between TCOF1 assembly and FC formation raises the question of how S/E-rich LCRs facilitate TCOF1 assembly. We reasoned that if TCOF1 self-assembles via heterotypic interactions between S/E-rich LCRs and other element(s) in TCOF1, then a version of TCOF1 lacking the S/E-rich LCRs will co-assemble with a version of TCOF1 containing the S/E-interacting element(s), as interactions are possible between them. However, if TCOF1 self-assembles via homotypic interactions between the S/E-rich LCRs, then two versions of TCOF1 will not co-assemble in cells unless they both possess S/E-rich LCRs, as they cannot interact when either lacks all of its S/E-rich LCRs.

To determine how the S/E-rich LCRs mediate TCOF1 assembly, we co-expressed mEGFP- and mCherry-tagged versions of TCOF1 and assessed their co-assembly. In cells, WT TCOF1 co-assembles with itself, demonstrated by co-localization of mEGFP and mCherry (Figure 3D). However, WT TCOF1 and TCOF1ΔSE failed to co-assemble. Rather, WT TCOF1 excluded TCOF1ΔSE (Figure 3D), suggesting that TCOF1's S/E-rich LCRs mediate assembly through self-interaction and not interaction with another element within TCOF1. This was not dependent on WT TCOF1 nucleolar localization, as co-assembly assays between TCOF1ΔSE and a non-nucleolar TCOF1 mutant[13] containing the repeat domain yielded similar results (Figure S6A). This suggests that homotypic interactions between S/E-rich LCRs mediate TCOF1 self-assembly, though it remains unclear how negatively charged sequences can interact in cells. We speculate that an abundant divalent cation in cells, like Mg^{2+}, may mediate the self-interactions of TCOF1 through salt bridges. Mg^{2+} mediates interphosphate interactions in nucleic acids and is an appealing candidate because serines in the TCOF1 S/E-rich region can be phosphorylated by the kinase CK2.[24]

Collectively, these results provide a molecular mechanism by which a cellular condensate can be scaffolded by a self-interacting component. TCOF1 scaffolds the nucleolar FC through self-assembly mediated by homotypic interactions of its repeated S/E-rich LCRs. These findings show that the size, shape, and even presence of a cellular condensate can be dictated through the precise expression, or lack of expression, of a single, naturally existing scaffold.

The ability of TCOF1 to scaffold the FC through self-assembly is a fundamentally different scaffold architecture for an intracellular condensate from those that depend on collective, interdependent interactions. This simple architecture presents a compelling mechanism by which condensates may evolve: the emergence of a condensate scaffolded by a single self-assembling component could be entirely linked to the emergence of this single component. A condensate scaffolded by a single component may represent an evolutionary intermediate for condensates that rely on interdependent interaction networks, such as the GC and stress granules. We therefore sought to test whether our understanding of TCOF1 can provide a molecular explanation for the evolutionary origins of the FC.

Given that the FC is thought to have emerged in amniotes,[11] we asked if TCOF1 emerged at a similar point in evolution. We first broadly assessed whether the presence of TCOF1 orthologs correlates with the presence of an FC (i.e., tripartite nucleolus). We looked for species where we could assess both nucleolar structure by existing EM data and the presence or the absence of TCOF1 orthologs in the proteome. Of the eight species we looked at, we observed a relationship between the existence of an FC and the presence of TCOF1 orthologs (Figures 4A and S7B–S7F; Table S1), with most species either having both a TCOF1 ortholog and an FC or lacking both (Figure 4A).

We next focused on species spanning the amniote transition: the bony fish *D. rerio*, the amphibian *X. laevis,* the turtle *P. scripta,* and the bird *G. gallus* (Figure 4A). Consistent with co-emergence of FC and TCOF1, *D. rerio* lacks both TCOF1 and an FC, while *G. gallus* has both. Although *P. scripta* possesses TCOF1 but lacks an FC, the existence of a TCOF1 ortholog and an FC in *X. laevis*[25] suggests that the FC may have been lost in turtles for other reasons. This association between the presence of TCOF1 and the FC led us to hypothesize that the emergence of TCOF1 may have played a role in FC emergence during evolution. If true, this hypothesis would suggest that introduction of TCOF1 into a bipartite system may be sufficient to form an FC-like nucleolar compartment.

We introduced human TCOF1 into zebrafish, which lack both TCOF1 and an FC (Figures 4A and S7A), to see if it forms an FC-like condensate that reorganizes the bipartite structure of the zebrafish nucleolus. We injected mEGFP-TCOF1 mRNA into zebrafish embryos, derived cells from these embryos, and used the FZ marker Fibrillarin to assess whether TCOF1 forms an FC-like condensate (Figures 4B and S7G). In uninjected embryos, Fibrillarin staining appears as small nuclear bodies, consistent with its FZ localization in bipartite nucleoli (Figure 4C). Strikingly, in cells from embryos injected with WT TCOF1 mRNA, we observed nuclear TCOF1 condensates, nested within Fibrillarin (Figure 4C). Fibrillarin organization conforms to the shape of the TCOF1 condensate, surrounding it as a DFC surrounds an FC in tripartite nucleoli. Strikingly, Fibrillarin exclusion is maintained irrespective of TCOF1 condensate size, like the expanded FCs we observed in human cells, which have tripartite nucleoli (Figures 1D and S7H). Importantly, an assembly deficient TCOF1 mutant mixes with Fibrillarin (Figure 4C), suggesting that the ability of human TCOF1 to form an FC-like condensate in zebrafish requires its self-assembly.

We sought to understand the degree to which these FC-like condensates recapitulate FC properties, namely formation of nucleoli with three distinct subcompartments and spatial organization of transcription surrounding the FC. To test if TCOF1 can induce a bipartite-to-tripartite reorganization in zebrafish nucleoli, we co-injected mCherry-zfMpp10 (zebrafish homolog of GC marker MPP10) and mEGFP-TCOF1 mRNAs into zebrafish embryos. We observed that TCOF1 FC-like condensates exclude Mpp10. Strikingly, with TCOF1, the nucleolus forms three distinct nested layers reminiscent of tripartite nucleoli in which Mpp10 surrounds fibrillarin, which in turn surrounds the TCOF1 FC-like condensates (Figures 4D and S7I). Moreover, TCOF1 FC-like condensates excluded nascent nucleolar RNA (Figure S7J), much like in tripartite systems[20,21] (Figures S3D and S3E). Collectively, these results indicate that TCOF1-induced FC-like condensates in zebrafish reorganize bipartite nucleoli to exhibit key structural features defining tripartite nucleoli.

The discovery that TCOF1 is sufficient to form an FC-like compartment in bipartite nucleoli provides direct experimental evidence for a molecular explanation of the bipartite-to-tripartite transition, which can be explained through the emergence of TCOF1. More broadly, these findings demonstrate that understanding the underlying architecture of biomolecular condensates can provide insight into their evolution.

DISCUSSION

Together, this work presents a model for the structure and molecular origins of the nucleolar FC in which emergence of TCOF1 is coupled to the emergence of the FC, the underlying structure of which is defined by the ability of TCOF1 to self-assemble through homotypic interactions between its S/E-rich LCRs (Figure 4E). More generally, this work not only suggests that a condensate may be formed by a protein acting as a single-component scaffold but also presents a general model for the formation and origin of these complex structures.

The idea that condensates may be scaffolded by a single protein is consistent with an emerging view that the underlying structural principles of membraneless organelles may relate to those in other biological structures,[13] such as the siliceous skeletons of certain sponges[26] or spider silk. These assemblies often form from very few scaffolding components, allowing the properties of the assembly to be largely controlled by the expression of just a few components. For example, while other components influence the specific physical properties of the silk fibers,[27] the formation of spider silk is dependent on the expression of two spidroin proteins,[28–30] which are sufficient to form silk fibers.[31,32] Likewise, we find that the size of the nucleolar FC is tuned by TCOF1 levels. These findings raise the possibility that other condensates, particularly recently emerged ones, may be similarly tuned and organized. Thus, understanding the architecture of biological structures is key to understanding their formation and evolution.

Given that emergence of an assembly based on single-component scaffolds requires only a single self-assembling component for its underlying architecture, these assemblies may arise in fewer evolutionary steps compared with ones with many co-interacting molecules. Following emergence of a single-component scaffold, molecules that contribute to the structure or function of the assembly can be gained gradually. In fact, assemblies thought to be made up of highly interconnected components, such as stress granules, p-bodies, and the nucleolar GC,[7,8,22,33] are much more evolutionarily ancient than the FC, which relies on TCOF1. Similarly, several species-specific (and potentially more recently emerged) assemblies, such as spider silk, are scaffolded by just a few components.[28] These distinct assembly architectures may be explained by a more general model for the emergence of higher-order biological assemblies with complex architectures, in which assemblies first arise from single self-interacting scaffolds, gain interaction partners over time, and eventually rely on these interaction partners for the structure of the assembly itself.

While we are only beginning to gain insight into these diverse complex biological assemblies, our understanding of TCOF1 as an archetypal self-assembling scaffold underlying the architecture of a nascent biomolecular condensate may inform future studies of the assembly and evolution of biomolecular condensates across the tree of life.

Limitations of the study

We use a thermodynamic framework to demonstrate that TCOF1 exhibits single-component self-assembly. While we observe this across orders of magnitude of TCOF1 concentrations, it is formally possible that an unknown co-assembling component remains in excess of TCOF1 across our conditions or increases in concentration upon TCOF1 expression. However, given our results showing that TCOF1 can occupy nearly an entire nucleus, it would need to not be limiting for this level of TCOF1 in cells. While we have not directly excluded the possibility that TCOF1 may promote the expression of a co-assembling scaffold, our results in zebrafish imply that if this hypothetical co-scaffold exists in humans, then it (and the ability of TCOF1 to increase its concentration) would also have to exist in zebrafish, which lack an FC. Another limitation is the extent to which we can “demonstrate” causality in evolution. We attempt to make clear that our study presents a possible, but compelling, mechanism for the emergence of biomolecular condensates.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Eliezer Calo (calo@mit.edu).

Materials availability—Plasmids and cell lines generated in this study are available upon request to the lead contact.

Data and code availability

- All data and materials are available upon request to the lead contact for the purposes of reproducing or extending the analysis.
- Code used for image analysis can be found on zenodo at: https://doi.org/10.5281/zenodo.7236056.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human cell culture—HeLa S3 cells were obtained from ATCC (CCL-2.2). Cells tested negative for mycoplasma. HeLa cells were cultured in 5% CO2 on cell culture-treated 10 cm plates (Genesee Scientific, 25–202) in Dulbecco’s Modified Eagle Medium (DMEM, Genesee Scientific, 25–500) supplemented with 10% Fetal bovine serum (FBS, Gemini Bio-products, 100–106) and 1% Penicillin/Streptomycin (Gibco, 10378–016). Cells were split 1:10 every 3 days by using trypsin (Gibco, 25200072).

Zebrafish husbandry—Zebrafish were housed in AAALAC-approved facilities and maintained according to protocols approved by the Massachusetts Institute of Technology Committee on Animal Care. AB/Tübingen (TAB5/14) genetic background strains (ZFIN: ZDB-GENO-010924-10) were used. Zebrafish embryos are of indeterminate sex at 27 hpf, and are neither male nor female at this stage.

METHOD DETAILS

Generation of plasmids and constructs—Coding sequences of wildtype proteins were cloned from human or zebrafish cDNA, from which deletions were made, if applicable. The only exception was pcDNA3.1(+) mEGFP - TCOF1 ΔSE and DSE 83–785, which were cloned from a synthesized gene block (Twist Bioscience) which was codon optimized to reduce repetitiveness of the DNA sequence. All constructs contain a GSAAGGSG peptide linker between GFP/mCherry and the protein of interest. SnapGene software (https://www.snapgene.com/) was used to design and view plasmids and constructs.

Generation of degron cell lines—Degron lines were generated in HeLa cells using sequential rounds of the following approach of CRISPR-Cas9-mediated cutting, homologous recombination (HR)-mediated tagging with the FKBP degron[14] fused to different selection markers, and selection, to achieve complete tagging of the TCOF1 locus. See Figure S2A for flowchart, and Figure S2B for HR templates. An sgRNA targeting the 3′UTR of TCOF1 proximal to the stop codon was cloned into the Cas9-containing plasmid PX458 (Addgene, 48138) by golden gate cloning, and subsequently sequence verified. The sgRNA used (5′-GTATGACGAGCACCAGCACC-3′) had on-target and off-target scores of 58.8 and 57.8, respectively, and was designed using Benchling. Off-target scores calculation included masked and low-complexity genomic regions.

HR templates depicted in Figure S2B are described here for completeness. HR templates contained an upstream homology arm 900 bp in length, which corresponds to the 900 bp of the TCOF1 genomic locus immediately upstream of the stop codon, and was cloned from HeLa genomic DNA. downstream of this homology arm, we cloned the FKBP degron cassette consisting of (GSG linker)-(FKBP-V)-(2xHA)-(GSG linker)-(P2A)-(Selection marker)-(STOP). The stop codon was followed by a downstream homology arm, which corresponded to the 821 bp of the TCOF1 genomic locus immediately downstream of the stop codon, and was cloned from HeLa genomic DNA. In the downstream homology arm, the PAM site at which the selected sgRNA cuts was mutated from AGG to AGC to prevent cutting of the locus after successful tagging. The entire HR template was cloned into a plasmid.

The process for generating the lines is depicted in Figure S2A, and described here. First, HeLa cells were co-transfected with the TCOF1 sgRNA plasmid and the HR template carrying Hygromycin resistance. Two days after transfection, cells were selected with 375 μg/mL of Hygromycin for 1 week, generating a Hygromycin resistant (HygR) population. This HygR population was then used in two independent tagging processes, with either Blasticidin- or Puromycin-containing versions of the HR-template, which led to the generation of the E6 and E12 clones used in the study.

For generation of the E6 clone, the HygR population was co-transfected with the TCOF1 sgRNA population and the HR template carrying Blasticidin resistance. Two days after transfection, cells were selected with 4 μg/mL of Blasticidin, for 15 days, generating a HygR, BlastR resistant population. This HygR BlastR population was then single-cell sorted into 96-well plates using a BD FACSAria cell sorter and grown up in the absence of selection. Clones were grown up and screened by PCR, followed by western blotting and immunofluorescence, as described in ‘validation of degron cell lines’ below, and Clone E6 was selected for use in subsequent experiments.

For generation of the E12 clone, the HygR population was co-transfected with the TCOF1 sgRNA population and the HR template carrying Puromycin resistance. Two days after transfection, cells were selected with 2 μg/mL of Puromycin, for 15 days, generating a HygR, PuroR resistant population. This HygR PuroR population was then single-cell sorted into 96-well plates using a BD FACSAria cell sorter and grown up in the absence of selection. Clones were grown up and screened by PCR, followed by western blotting and immunofluorescence, as described in ‘validation of degron cell lines’ below, and Clone E12 was selected for use in subsequent experiments.

Validation of degron cell lines—Degron cell lines were first screened by PCR, and then further validated by Western blot and immunofluorescence (Figures S2C–S2E). Putative clones, including clones E6 and E12, were identified in the PCR screening, looking for lack of a wildtype PCR product (1268 bp) using genotyping primers spanning the genomic locus outside of the left homology arm and inside the right homology arm (forward: 5′-GCTGGCCTCCAGGGGGCAGGTGAA-3’; reverse primer: 5′-ACAGGGGACACCAGAGCTGT-3′). Following PCR screening, clones were tested for TCOF1 degradation upon dTAG-13 treatment (24hrs, 100 nM) by Western blot as well as immunofluorescence. Two independent clones, E6 and E12, were selected on the basis of 1) a lack of a wild-type band by PCR screening, 2) loss of TCOF1 protein after dTAG treatment, and 3) nucleolar localization which is lost after dTAG treatment by immunofluorescence.

Inducible degradation using the degron system and dTAG-13 treatment—In experiments using dTAG-13 treatment to induce TCOF1 degradation, 24 h after plating, cells were treated with DMSO or 100 nM dTAG-13 for 4 days, replacing media once 2 days after the initial treatment. For these experiments, cells were plated at low seed densities to account for the 4 days of treatment. For immunofluorescence assays in 24-well plates, cells were seeded at 12,500 cells per well. For western blots in 6-well plates, cells were seeded at 75,000 cells per well.

Western blotting—Cells were collected for whole cell protein lysate using RIPA buffer (150 mM NaCl, 5 mM EDTA, 50 mM TRIS pH 8.0, 0.5% Sodium deoxycholate, 1% NP-40, 0.1% SDS) with 1x cOmplete protease inhibitor cocktail (Sigma Aldrich, 11697498001) and 1 mM PMSF (ThermoFisher Scientific, 36978). Lysate was sonicated using a BioRuptor 300 for 5 min (5 cycles of 30 s on, 30 s off) and cleared by taking the supernatant after spinning at 15,000 × g. Total protein concentrations were measured by Bradford Assay (Thermo Fisher Scientific, 23246), and 50 μg of total protein was run on a Tris-glycine polyacrylamide gel. Transfer to a methanol-activated PVDF membrane was run in Tris-glycine buffer (1x Tris glycine, 10% methanol, 0.05% SDS) at 40 V for 16 h at 4°C. After transfer, membranes were blocked in 5% milk in PBST (1x PBS, 0.1% Tween 20) for 1 h and blotted with primary antibody (anti-TCOF1, Protein-tech, 11003–1-AP; anti-HA, Cell Signaling (C29F4), 3724S; anti-Tubulin, Thermo Scientific, MA5-16308) overnight at 4°C. Membranes were then blotted with secondary antibodies (anti-rabbit, Invitrogen, 32260; anti-mouse, Invitrogen, 32230) for 1 h at room temperature, developed using SuperSignal West Femto (Thermo Scientific, 34096), and imaged on a Bioanalytical Imaging System model c500 (Azure Biosciences).

Fluorescence microscopy—Glass coverslips (Fisherbrand, 12-545-80 or 12-541-001) were placed in 24-well plates (Genesee Scientific, 25–107) and coated in 3 μg/mL of fibronectin (EMD Millipore, FC010) for 30 min at room temperature. HeLa cells were seeded in each well at 50,000 cells per well. 24 h after seeding, the cells were transfected with GFP-tagged protein plasmids using Lipofectamine 2000 (Invitrogen, 11668027). Each well was transfected using 100 ng of plasmid and 1 μL of Lipofectamine 2000 in a total of 50 μL of OptiMEM (Gibco, 31985070) according to the Lipofectamine 2000 instructions. For co-assembly experiments, 100 ng of each plasmid (200 ng total) was transfected using the same mix. For Figures S1 and S3D–S3E, 200 ng of plasmid was transfected.

Cells on glass coverslips were collected for immunofluorescence 48 h after transfection. Cells were collected by washing with 1x PBS (Genesee Scientific, 25–508) and fixation in 4% paraformaldehyde (PFA) for 15 min at room temperature, followed by another 3 washes with 1x PBS. In experiments with costaining for markers by immunofluorescence, cells were permeabilized and blocked by incubation in blocking buffer (1% BSA (w/v), 0.1% Triton X-100 (v/v), 1x PBS) for 1 h at room temperature. Coverslips were then incubated overnight at 4°C in a 1:100 dilution of primary antibody (anti-UBTF, Novus Biologicals, NBP1-82545; anti-POLR1A, Novus Biologicals, NBP2-56122; anti-Fibrillarin, EMD Millipore, MABE1154; anti-MPP10, Novus Biologicals, NBP1-84341; anti-Nucleophosmin, Abcam, ab86712; anti-TCOF1, Novus Biologicals, NBP1-86909; anti-HA, Invitrogen, 26183) in blocking buffer. After 3 washes with blocking buffer, coverslips were incubated for 2 h in a 1:1000 dilution of secondary antibody (anti-rabbit Alexa 647, Invitrogen, A-27040; anti-mouse Alexa 647, Invitrogen, A-32728, anti-mouse Alexa 488, Invitrogen, A-32723). Coverslips were washed 3 times with blocking buffer, and then once with 1x PBS. Nuclei were then stained with a 1:1000 dilution of Hoechst 33342 (Thermo Scientific, 62249) for 15 min, then washed twice with 1x PBS, and mounted on glass slides using ProLong Diamond antifade mountant (Invitrogen, P36961). Slides were sealed using clear nail polish, allowed to dry, and stored at 4°C.

Slides were imaged on either an Olympus FV1200 Laser Scanning confocal microscope or a DeltaVision 2 TIRF microscope. All partition and size analyses were performed using images from the Olympus FV1200 confocal microscope. Quantification of 5-EU incorporation in Figure S2F was performed using images from the Olympus FV1200 confocal microscope; all other images of 5-EU incorporation were performed on the DeltaVision 2 TIRF microscope. Costains of transfected human cells were imaged on either microscope. Co-assembly experiments were imaged on the DeltaVision 2 TIRF microscope. Zebrafish cells derived from injected embryos were imaged on the DeltaVision 2 TIRF microscope. The same set of exposure conditions (one exposure per channel) was used across all slides within the same experiment. For images taken on Olympus FV1200, maximum projections across the z axis of acquired images were made using Fiji (https://imagej.net/software/fiji/)[35] and used for any further analyses. For images taken on DeltaVision 2 TIRF microscope, deconvolution and max projection was performed using softWoRx software. Representative images were shown.

Transmission electron microscopy (TEM)—Cells were seeded in 6-well cell culture plates. Upon collection, media was removed from the plates and immediately fixed using 2.5% Glutaraldehyde 2.5% Paraformaldehyde in 0.1 M sodium cacodylate buffer at pH 7.4 (Electron Microscopy Sciences, 15949) for at least 1 h at room temperature before embedding.

For embedding, fixative was removed and the cells were washed twice in 0.1 M Sodium Cacodylate buffer, pH 7.4. Then, the cells were incubated in 1% Osmium tetroxide (OsO4)/ 1.5% Potassium ferrocyanide (KFeCN6) for 30 min, washed twice in water, and once in 50 mM Maleate buffer, pH 5.5 (MB). The plate was then incubated in 1% uranyl acetate in MB for 30 min followed by one wash in MB and two washes in water and subsequent dehydration in grades of alcohol (5 min each; 50%, 70%, 95%, 2 × 100%). Following the washes, cells were removed from the dish in propyleneoxide, pelleted at 3000 rpm for 3 min and incubated overnight in a 1:1 mixture of propyleneoxide and TAAB Epon (TAAB Laboratories Equipment Ltd, https://taab.co.uk). The following day the samples were embedded in TAAB Epon and polymerized at 60°C for 48 h.

Ultrathin sections (about 60–80 nm) were then cut on a Reichert Ultracut-S microtome, picked up on copper grids, stained with lead citrate and examined in a JEOL 1200EX Transmission electron microscope at 4000X. Images were recorded with an AMT 2k CCD camera and representative images were shown.

***In vitro* transcription of RNA—**To generate the RNA for zebrafish microinjections in Figure S7H, a construct encoding mEGFP-TCOF1 (RP133) was amplified using Q5 DNA polymerase (New England BioLabs, M0491L) reaction mix with primers RPO378 (GCGTAATACGACTCACTATAGGGAGACCCAA) and RPO379 (GTGGATCCGAGCTCGGTACCAA). Products were then run on a 1% agarose gel and the gel was extracted using NucleoSpin (Macherey-Nagel, 740588.50) clean up kit according to manufacturer specifications. For Figures 4C, 4D, and S7I–S7J, RNA was generated from digesting RP133, RP168, or RP226 with BamHI-HF (New England BioLabs, R3136S). The product was then purified from an agarose gel and extracted using the NucleoSpin clean up kit (Macherey-Nagel, 740588.50).

PCR DNA or digested linearized DNA were used to set up an overnight IVT reaction using mMESSAGE mMACHINE T7 ULTRA Transcription Kit (Invitrogen, AM1345). The overnight product was then treated with TURBO DNase, from the same mMESSAGE mMACHINE Kit. After the DNAse treatment the samples were poly-A tailed using the tailing reaction provided in the mMESSAGE mMACHINE Kit. An aliquot of the DNA was removed prior to adding E-PAP enzyme, to be run at a later time point on RNA agarose gel to confirm band size and act as control for post poly-A tailing reaction. Samples were then purified using Zymo RNA Clean and Concentrator kit (Zymo Research, R1016) or Qiagen RNeasy Mini kit (Qiagen, 74104).

TCOF1 ortholog search approach and definition—To search for the existence of orthologs of TCOF1 in other species, we used the BLAST webtool from UniProt (https://www.uniprot.org/blast) and NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi), allowing for maximal returned hits and no filtering based on low-complexity sequences. Because of the low-complexity of human TCOF1, orthologs of TCOF1 were determined based on manual assessment of the sequence rather than by alignment coverage. Species which contained TCOF1 were defined as species where a sequence from the BLAST results was similar to human TCOF1 based on manual dotplot analysis of the LCR relationships within each given protein sequence. In particular, TCOF1 orthologs were defined based on the presence of repeat domains containing multiple S/E-rich LCRs and a C-terminal region containing K-rich LCR(s).

Dotplot analysis of protein sequence—Self-comparison dotplots of TCOF1 protein sequences were generated such that the protein sequence is compared to itself position-by-position in a matrix, where any identically-matching residues are colored by a pixel. For each dotplot, protein sequences were integer-encoded such that each of the 20 amino acids corresponds to a unique integer from 1 to 20, inclusive. Two arrays of this sequence x N, row-wise and column-wise, were generated, such that each array was a matrix of size N × N, where N is the protein sequence length. The two matrices were subtracted such that any identical amino acid matches equaled 0 and non-matches were non-zero. The final dotplot matrix was generated by replacing any 0 values with 1 and replacing any non-zero values with 0. Dotplot matrices were plotted such that values of 1 were colored with a purple pixel.

Injection of zebrafish embryos—Zebrafish were crossed by placing a male and female fish in a shared container, but kept physically separate until combined the next morning. Embryos were collected for injection 20 min after males and females were combined for fertilization. Injection mix consisting of 0.27 μM *In vitro* transcribed RNA and 0.07% Phenol Red solution (Sigma, P0290) was prepared and loaded into needles pulled from glass capillary tubing (64–0766). Embryos were injected 1 nL of injection mix using a pico-liter injector (Warner Instruments, PLI90A). For co-injection experiments, injection mix consisted of 0.27 μM *In vitro* transcribed mEGFP-TCOF1 RNA, 0.12 μM *In vitro* transcribed mCherry-zfMpp10 RNA, and Phenol Red solution. Single injection control mixes for co-injection experiments contained the same amount of individual RNAs as in the co-injection mix. Injected embryos and uninjected controls were then incubated in fish water at 28°C until collection at approximately 27 h postfertilization (hpf).

Whole-mount imaging of zebrafish embryos—Zebrafish embryos at approximately 27 hpf were collected, dechorionated using ethanol-sterilized forceps, and moved to fresh fish water. Using a transfer pipette, individual embryos were anesthetized by transfer into 1X Tricane solution, prepared by diluting 25 X Tricane solution (pH 7 solution of 4 mg/mL Tricane-S (Syndel) and 20 mM TRIS) to 1 X in fish water. Fish were subsequently transferred into 1.5% low-melt agarose (Lonza, 50100), and transferred to a clear dish suspended in low-melt agarose for imaging. Bright-field and GFP fluorescence images were taken on a Leica M205 FCA fluorescence stereo microscope. Images were taken with the same settings for all embryos, at 5.5× magnification, and representative embryos were shown.

Deriving cells from zebrafish embryos—Cells for immunofluorescence were derived from ~27 hpf embryos grown at 28°C. Embryos expressing high GFP signal were selected prior to dechorionation. For embryos injected with only mCherry encoding RNA, no selection was performed. Embryos were dechorionated, using ethanol-sterilized forceps for immunofluorescence or pronase (1 mg/mL) for electron microscopy, and then transferred to 1x PenStrep Glutamine (Gibco, 10378016) in 1x PBS (PBS+PSG) for 30 min. Once the embryos were dechorionated sterile tubes and pipettes tips were used to minimize any possible contamination to the embryos/cells. Embryos were then transferred to new tubes with fresh PBS+PSG. Yolk-sacs were removed through pipetting several times with P200 pipette. Embryos were then spun at 1200 × g for 2 min at room temperature and the supernatant was then discarded. The embryos were then incubated in 300 μL of prewarmed .25% Trypsin (Gibco, 25200072) for 30 min at 1000 rpm on a thermomixer (Eppendorf, Thermomixer R). The embryos were then pipetted with P200 to further resuspend the embryos into single cell suspension and then were spun at 1200 × g for 4 min at room temperature. Supernatant was disposed of and the embryos were serially washed with sterile PBS (Genesee Scientific, 25–508). Cells were then resuspended in 400 μL of DMEM, supplied with 15% FBS (Gemini Bio, 100–106), 1X PSG (Gibco, 10378016) and 1 mg/mL gentamicin (Gibco, 15750060) in 24-well cell cultures plates (Genesee Scientific, 25–107) with coverslips (Fisher Scientific, 1254580) that were coated with 3 mg/mL fibronectin in PBS for at least 30 min prior to plating cells. Cells were allowed to attach overnight at 28°C supplied with 5% carbon dioxide prior to 5-EU treatment (3 h at 28°C) and/or immunofluorescence staining.

QUANTIFICATION AND STATISTICAL ANALYSIS

Partition and size analysis—Maximum projection images were analyzed to calculate average nuclear intensity, the free energy of transfer ($\Delta G^{\text{transfer}}$), and dense phase size of GFP-tagged proteins using custom python code (found on zenodo at: https://doi.org/10.5281/zenodo.7236056). Briefly, nuclei were segmented in the images by applying Gaussian blur, followed by thresholding at the mean intensity in the Hoecsht 33342 channel.

The average nuclear intensity for each nucleus was defined as the average nuclear GFP signal following background correction, and was used as a proxy for nuclear concentration of the GFP-tagged protein. Background signal was defined as the 75th percentile of GFP signal in nuclei of an untransfected sample. This background value was subtracted from the GFP intensity of all images of transfected samples within the set of images taken with identical settings. GFP signal beyond this point refers to the background corrected GFP signal.

To calculate $\Delta G^{\text{transfer}}$, which is a measurement of the energy of partitioning into the dense phase, the dense and dilute phase were determined by segmenting the GFP channel using Otsu thresholding. Nuclei in which the Otsu threshold was below an intensity value of 100 were considered untransfected cells and excluded from analysis. This threshold was determined by manually checking that untransfected cells were excluded based on the value of the threshold. Any nuclei with overexposed pixels in the GFP channel were also excluded from analysis. For partition analysis, dense phase concentrations were defined as the 90th percentile GFP intensity within the segmented region, and dilute phase concentrations were defined as the median GFP intensity outside of the segmented region. $\Delta G^{\text{transfer}}$ was then calculated using the following equation:

$$\Delta G^{transfer} = RTIn(K_{partition}); K_{partition} = \frac{dense\ phase\ concentration}{dilute\ phase\ concentration}$$

where R is the gas constant 1.987×10^{-3} kcal K^{-1} mol^{-1} and T = 310.15 K. Dense phase size was defined as the number of pixels in the dense phase as segmented above.

In some nuclei, the low intensity of dilute phase after background correction resulted in noisy partition coefficient calculations, so we used a minimum dilute phase intensity cutoff intensity of 6 to mitigate this. $\Delta G^{transfer}$ was not calculated for nuclei which did not pass this cutoff, as this calculation requires a dilute phase signal. However the dense phase size calculation does not require a dilute phase signal, and as such this calculation was still performed on these nuclei. As a consequence, the number of nuclei used for the $\Delta G^{transfer}$ calculation for some experiments is smaller than the number of nuclei used for the dense phase size calculation.

Fits for partition analyses were performed using the regplot function in seaborn, with “logx = True” for $\Delta G^{transfer}$ plots, and “logx = False” for dense phase size plots. For all fits, 95% confidence intervals are shown.

For the TCOF1ΔSE construct, partition analysis could not be performed due to colocalization with microtubules in cells with high expression of the construct (Figure S6B). This microtubule co-localization is reflected in the images following max projection, confounding quantification of dilute and dense phases.

Image quantification of nascent transcription in nucleoli—Nascent transcription in nucleoli was measured by incorporation of 5-ethynyl uridine (5-EU) in cells using a Click-iT RNA Alexa Fluor 594 Imaging Kit (Fisher Scientific, C10330), according to the manufacturer’s recommended protocol except where noted in the following. Briefly, cells were plated on fibronectin-coated glass coverslips in 24-well plates, as described in the respective sections above. On the day of collection, cells were incubated in 1 mM 5-EU for 3 h (unless noted otherwise) at 37°C. After 5-EU incorporation, cells were washed once with 1x PBS, and then fixed in 4% paraformaldehyde for 15 min at room temperature. Permeabilization, Click reaction to covalently attach Alexa 594, and washes were all done as recommended by the kit. After washing the Alexa 594 Click reaction, but before Hoechst 33342 nuclear staining, cells were stained for nucleolar marker MPP10 using the immunofluorescence and imaging methods as described above.

To quantify nascent transcription in cells, nuclei of cells were segmented, followed by segmentation of nucleoli, and measurement of 5-EU Alexa 594 signal using custom python code (found on zenodo at: https://doi.org/10.5281/zenodo.7236056). Using max projection images, nuclei were segmented using the Hoechst 33342 channel as previously described above, and nucleoli were segmented using the MPP10 channel by performing Otsu thresholding on nuclear MPP10 signal. Then, for each nucleus, the amount of nascent transcription in nucleoli was calculated as the average 5-EU intensity within MPP10 segmented region(s). Background correction was performed by subtracting the mean nucleolar 5-EU intensity of a sample that was not treated with 5-EU. These data were then plotted to compare different cell lines and conditions. A two-tailed t test with no assumption of equal variance was used to test statistical significance, defined by a p value of less than 10^{-3}.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank all Calo Lab members for helpful discussions and feedback on the manuscript, the Swanson Biotechnology Center Microscopy and Flow Cytometry Facilities, and the Harvard Medical School EM Facility. We thank the Burge and Boyer Labs at MIT for providing FKBP degron plasmids and dTAG-13. This work was supported by the National Institutes of Health (B.L., T32GM007287), the National Institute of General Medical Sciences (N.J.-L. and E.C., R35GM142634), The Pew Charitable Trusts (E.C.), and the National Cancer Institute (E.C., P30-CA14051).

REFERENCES

1. Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM, et al. (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell 175, 1842–1855.e16. 10.1016/j.cell.2018.10.042. [PubMed: 30449618]
2. Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, Kriwacki RW, Pappu RV, and Brangwynne CP (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell 165, 1686–1697. 10.1016/j.cell.2016.04.047. [PubMed: 27212236]
3. Shin Y, and Brangwynne CP (2017). Liquid phase condensation in cell physiology and disease. Science 357, eaaf4382. 10.1126/science.aaf4382. [PubMed: 28935776]
4. Banani SF, Lee HO, Hyman AA, and Rosen MK (2017). Biomolecular condensates: organizers of cellular biochemistry. Nat. Rev. Mol. Cell Biol. 18, 285–298. 10.1038/nrm.2017.7. [PubMed: 28225081]
5. Brangwynne CP, Mitchison TJ, and Hyman AA (2011). Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proc. Natl. Acad. Sci. USA 108, 4334–4339. 10.1073/pnas.1017150108. [PubMed: 21368180]
6. Kedersha NL, Gupta M, Li W, Miller I, and Anderson P (1999). RNA-Binding Proteins Tia-1 and Tiar Link the Phosphorylation of Eif-2α to the Assembly of Mammalian Stress Granules. J. Cell Biol. 147, 1431–1442. 10.1083/jcb.147.7.1431. [PubMed: 10613902]
7. Riback JA, Zhu L, Ferrolino MC, Tolbert M, Mitrea DM, Sanders DW, Wei M-T, Kriwacki RW, and Brangwynne CP (2020). Composition-dependent thermodynamics of intracellular phase separation. Nature 581, 209–214. 10.1038/s41586-020-2256-2. [PubMed: 32405004]
8. Sanders DW, Kedersha N, Lee DSW, Strom AR, Drake V, Riback JA, Bracha D, Eeftens JM, Iwanicki A, Wang A, et al. (2020). Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. Cell 181, 306–324.e28. 10.1016/j.cell.2020.03.050. [PubMed: 32302570]
9. Lafontaine DLJ, Riback JA, Bascetin R, and Brangwynne CP (2021). The nucleolus as a multiphase liquid condensate. Nat. Rev. Mol. Cell Biol. 22, 165–182. 10.1038/s41580-020-0272-6. [PubMed: 32873929]
10. Thiry M, and Lafontaine DLJ (2005). Birth of a nucleolus: the evolution of nucleolar compartments. Trends Cell Biol. 15, 194–199. 10.1016/j.tcb.2005.02.007. [PubMed: 15817375]
11. Lamaye F, Galliot S, Alibardi L, Lafontaine DLJ, and Thiry M (2011). Nucleolar structure across evolution: the transition between bi- and tri-compartmentalized nucleoli lies within the class Reptilia. J. Struct. Biol. 174, 352–359. 10.1016/j.jsb.2011.02.003. [PubMed: 21335089]
12. Recher L, Whitescarver J, and Briggs L (1969). The fine structure of a nucleolar constituent. J. Ultrastruct. Res. 29, 1–14. 10.1016/S0022-5320(69)80052-3. [PubMed: 4900225]
13. Lee B, Jaberi-Lashkari N, and Calo E (2022). A unified view of low complexity regions (LCRs) across species. eLife 11, e77058. 10.7554/eLife.77058. [PubMed: 36098382]
14. Nabet B, Roberts JM, Buckley DL, Paulk J, Dastjerdi S, Yang A, Leggett AL, Erb MA, Lawlor MA, Souza A, et al. (2018). The dTAG system for immediate and target-specific protein degradation. Nat. Chem. Biol. 14, 431–441. 10.1038/s41589-018-0021-8. [PubMed: 29581585]

15. Valdez BC, Henning D, So RB, Dixon J, and Dixon MJ (2004). The Treacher Collins syndrome (TCOF1) gene product is involved in ribosomal DNA gene transcription by interacting with upstream binding factor. Proc. Natl. Acad. Sci. USA 101, 10709–10714. 10.1073/pnas.0402492101. [PubMed: 15249688]
16. Reynolds RC, Montgomery PO, and Hughes B (1964). NUCLEOLAR “CAPS” PRODUCED BY ACTINOMYCIN D. Cancer Res. 24, 1269–1277. [PubMed: 14216161]
17. Shav-Tal Y, Blechman J, Darzacq X, Montagna C, Dye BT, Patton JG, Singer RH, and Zipori D (2005). Dynamic Sorting of Nuclear Components into Distinct Nucleolar Caps during Transcriptional Inhibition. Mol. Biol. Cell 16, 2395–2413. 10.1091/mbc.E04-11-0992. [PubMed: 15758027]
18. Snyers L, Laffer S, Löhnert R, Weipoltshammer K, and Schöfer C (2022). CX-5461 causes nucleolar compaction, alteration of peri- and intranucleolar chromatin arrangement, an increase in both heterochromatin and DNA damage response. Sci. Rep. 12, 13972. 10.1038/s41598-022-17923-4. [PubMed: 35978024]
19. Tchelidze P, Benassarou A, Kaplan H, O’Donohue M-F, Lucas L, Terryn C, Rusishvili L, Mosidze G, Lalun N, and Ploton D (2017). Nucleolar sub-compartments in motion during rRNA synthesis inhibition: Contraction of nucleolar condensed chromatin and gathering of fibrillar centers are concomitant. PLOS ONE 12, e0187977. 10.1371/journal.pone.0187977. [PubMed: 29190286]
20. Yao R-W, Xu G, Wang Y, Shan L, Luan P-F, Wang Y, Wu M, Yang L-Z, Xing Y-H, Yang L, and Chen LL (2019). Nascent Pre-rRNA Sorting via Phase Separation Drives the Assembly of Dense Fibrillar Components in the Human Nucleolus. Mol. Cell 76, 767–783.e11. 10.1016/j.molcel.2019.08.014. [PubMed: 31540874]
21. Koberna K, Malínský J, Pliss A, Mašata M, Večeřová J, Fialová M, Bednár J, and Raška I (2002). Ribosomal genes in focus : new transcripts label the dense fibrillar components and form clusters indicative of “Christmas trees” in situ. J. Cell Biol. 157, 743–748. 10.1083/jcb.200202007. [PubMed: 12034768]
22. Mitrea DM, Cika JA, Guy CS, Ban D, Banerjee PR, Stanley CB, Nourse A, Deniz AA, and Kriwacki RW (2016). Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. eLife 5, e13571. 10.7554/eLife.13571. [PubMed: 26836305]
23. Banani SF, Rice AM, Peeples WB, Lin Y, Jain S, Parker R, and Rosen MK (2016). Compositional Control of Phase-Separated Cellular Bodies. Cell 166, 651–663. 10.1016/j.cell.2016.06.010. [PubMed: 27374333]
24. Werner A, Baur R, Teerikorpi N, Kaya DU, and Rape M (2018). Multisite dependency of an E3 ligase controls monoubiquitylation-dependent cell fate decisions. eLife 7, e35407. 10.7554/eLife.35407. [PubMed: 29999490]
25. Mais C, and Scheer U (2001). Molecular architecture of the amplified nucleoli of Xenopus oocytes. J. Cell Sci. 114, 709–718. 10.1242/jcs.114.4.709. [PubMed: 11171376]
26. Shimizu K, Amano T, Bari MR, Weaver JC, Arima J, and Mori N (2015). Glassin, a histidine-rich protein from the siliceous skeletal system of the marine sponge Euplectella, directs silica polycondensation. Proc. Natl. Acad. Sci. USA 112, 11449–11454. 10.1073/pnas.1506968112. [PubMed: 26261346]
27. Kono N, Nakamura H, Mori M, Yoshida Y, Ohtoshi R, Malay AD, Pedrazzoli Moran DA, Tomita M, Numata K, and Arakawa K (2021). Multicomponent nature underlies the extraordinary mechanical properties of spider dragline silk. Proc. Natl. Acad. Sci.USA 118, e2107065118. 10.1073/pnas.2107065118. [PubMed: 34312234]
28. Sponner A, Vater W, Monajembashi S, Unger E, Grosse F, and Weisshart K (2007). Composition and Hierarchical Organisation of a Spider Silk. PLOS ONE 2, e998. 10.1371/journal.pone.0000998. [PubMed: 17912375]
29. Xu M, and Lewis RV (1990). Structure of a protein superfiber: spider dragline silk. Proc. Natl. Acad. Sci. USA 87, 7120–7124. 10.1073/pnas.87.18.7120. [PubMed: 2402494]
30. Hinman MB, and Lewis RV (1992). Isolation of a clone encoding a second dragline silk fibroin. Nephila clavipes dragline silk is a two-protein fiber. J. Biol. Chem. 267, 19320–19324. 10.1016/S0021-9258(18)41777-2. [PubMed: 1527052]

31. An B, Hinman MB, Holland GP, Yarger JL, and Lewis RV (2011). Inducing β-sheets formation in synthetic spider silk fibers by aqueous post-spin stretching. Biomacromolecules 12, 2375–2381. 10.1021/bm200463e. [PubMed: 21574576]

32. Malay AD, Suzuki T, Katashima T, Kono N, Arakawa K, and Numata K (2020). Spider silk self-assembly via modular liquid-liquid phase separation and nanofibrillation. Sci. Adv. 6, eabb6030. 10.1126/sciadv.abb6030. [PubMed: 33148640]

33. Youn J-Y, Dunham WH, Hong SJ, Knight JDR, Bashkurov M, Chen GI, Bagci H, Rathod B, MacLeod G, Eng SWM, et al. (2018). High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. Mol. Cell 69, 517–532.e11. 10.1016/j.molcel.2017.12.020. [PubMed: 29395067]

34. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, and Zhang F (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8, 2281–2308. 10.1038/nprot.2013.143. [PubMed: 24157548]

35. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682. 10.1038/nmeth.2019. [PubMed: 22743772]

Highlights

- TCOF1 is a self-assembling scaffold necessary for the nucleolar fibrillar center (FC)
- Self-assembly of TCOF1 requires homotypic interactions between its S/E-rich LCRs
- TCOF1 self-assembly is necessary for its ability to scaffold the FC
- Human TCOF1 is sufficient to form an FC-like condensate in fish, which lack an FC

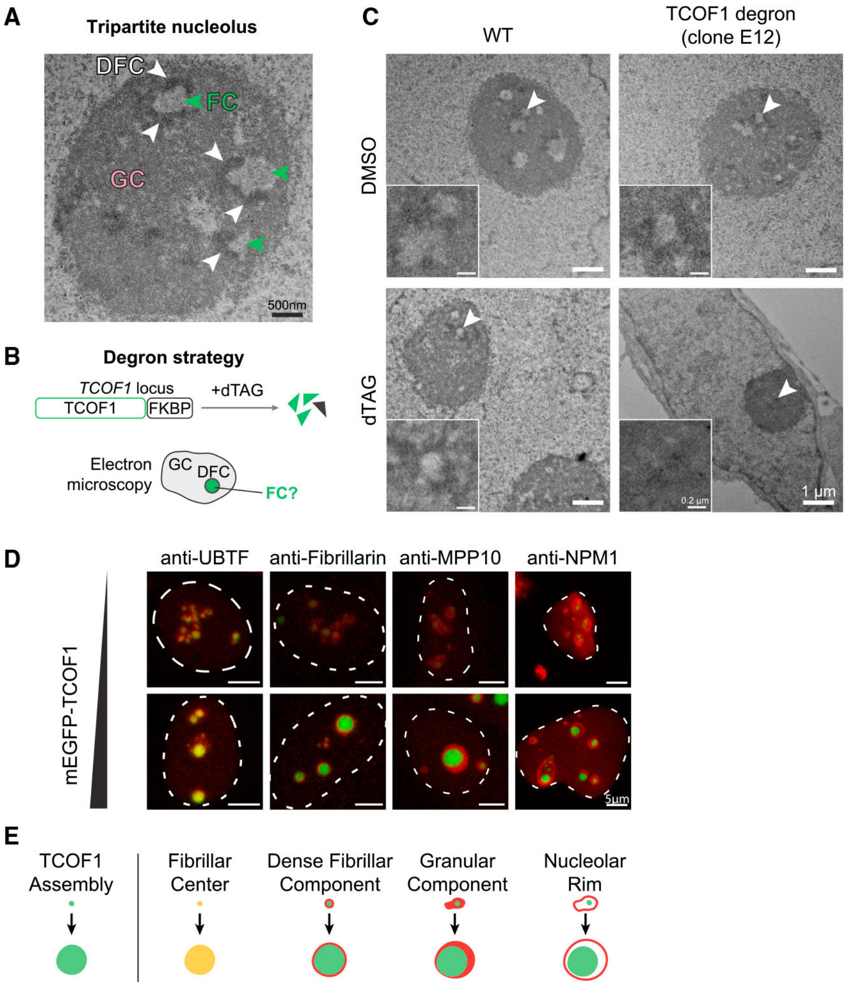


Figure 1. TCOF1 structurally defines the nucleolar fibrillar center

(A) Transmission electron microscopy (TEM) of a HeLa cell nucleolus with the GC, DFC, and FC labeled.

(B) Strategy for endogenous tagging of TCOF1 with the FKBP degron, degradation of TCOF1 with dTAG-13, and subsequent TEM.

(C) TEM of WT HeLa cells and TCOF1-degron HeLa cells in DMSO and dTAG-13 treatment. Scale bar: 1 μm. Inset shows a close up of the region indicated (arrowhead). Inset scale bar: 0.2 μm.

(D) Immunofluorescence of endogenous nucleolar markers across varying mEGFP-TCOF1 levels. Nuclei are outlined (dotted line). Scale bars: 5 μm.

(E) Model of TCOF1 structurally defining the FC and how TCOF1 assembly impacts nucleolar sub-compartments.

See also Figures S1–S3.

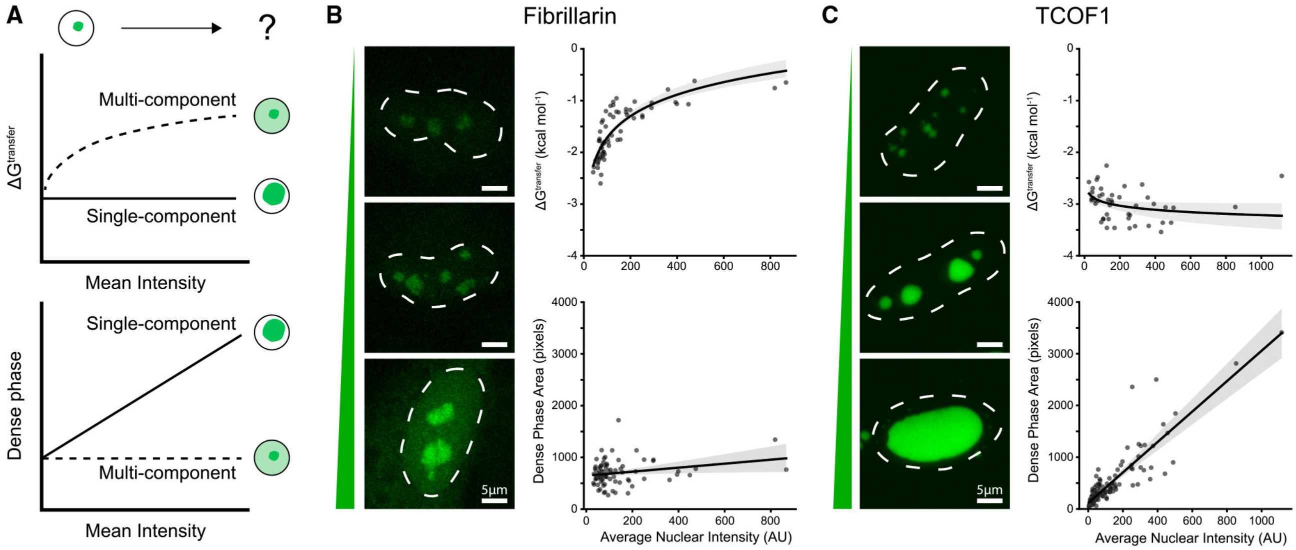


Figure 2. TCOF1 displays single-component assembly in cells

(A) Schematic illustrating expected ΔG^{transfer} and dense phase size trends for multi-component or single-component assembly behavior.

(B) Quantification of ΔG^{transfer} and dense phase size vs. average nuclear intensity for mEGFP-Fibrillarin. Cells with different expression levels are shown (left). Nuclei are outlined (dotted line). Scale bars: 5 μm. n = 66 nuclei for ΔG^{transfer}, n = 81 for dense phase area. Logarithmic and linear fits with 95% confidence intervals are shown.

(C) Same as (B) but for mEGFP-TCOF1. n = 46 nuclei for ΔG^{transfer}, n = 167 for dense phase area.

See also Figure S4.

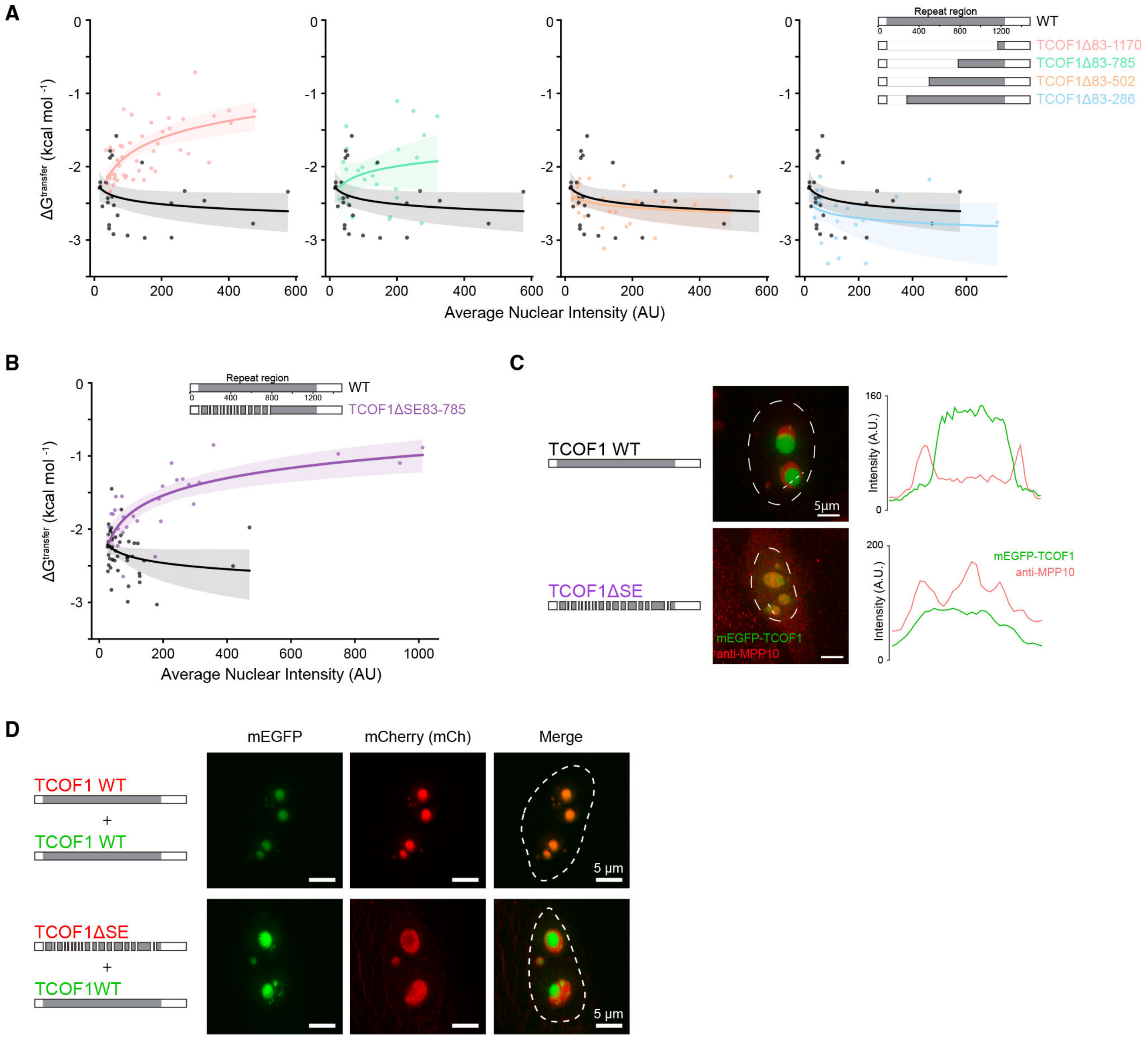


Figure 3. TCOF1 structurally defines the FC through self-interacting S/E-rich LCRs

(A) Quantification of $\Delta G^{\text{transfer}}$ vs. average nuclear intensity for WT TCOF1 and TCOF1 mutants with repeat region truncations. WT TCOF1 data are shown on each plot for clarity. n= 26 nuclei for WT TCOF1, n = 48 for TCOF1Δ83–1170, n = 25 for TCOF1Δ83–785, n = 27 for TCOF1Δ83–502, n = 24 for TCOF1Δ83–286. Logarithmic fits and 95% confidence intervals are shown.

(B) Quantification of $\Delta G^{\text{transfer}}$ vs. average nuclear intensity for WT TCOF1 and TCOF1ΔSE83–785. n = 46 nuclei for WT TCOF1, n = 29 for TCOF1ΔSE83–785.

(C) Localization of GC marker MPP10 with mEGFP-WT TCOF1 or mEGFP-TCOF1DSE. Nuclei are outlined (dotted line). Scale bars: 5 μm. Line profiles are of regions indicated by the finely dotted line for mEGFP (green) and MPP10 channels (red).

(D) Fluorescence microscopy of co-assembly experiments in cells. Nuclei are outlined (dotted line). Scale bars: 5 µm.

See also Figures S5 and S6.

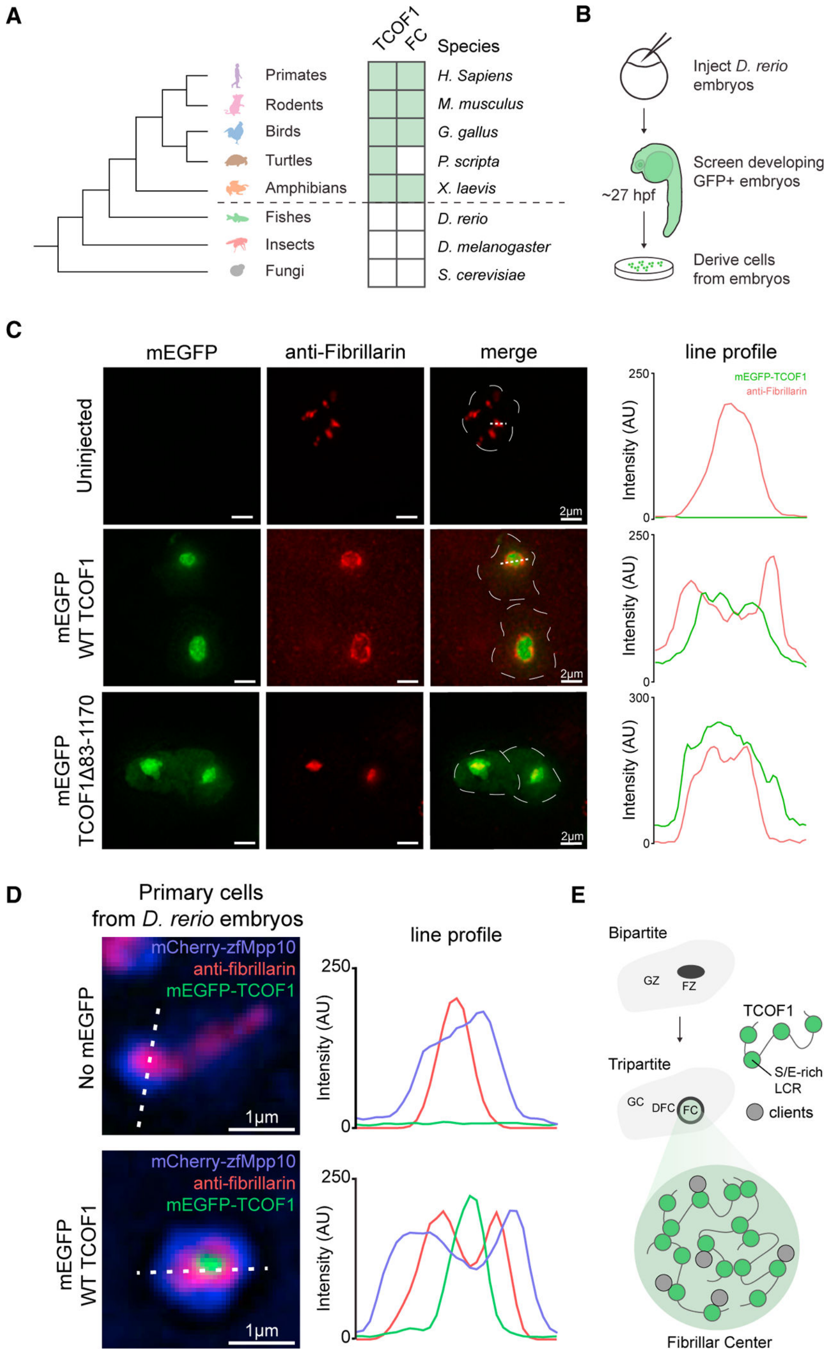


Figure 4. Emergence of the fibrillar center can be explained by emergence of TCOF1

(A) Analysis of presence of TCOF1 and FC-containing (tripartite) nucleoli in species (green). Tree (left) illustrates evolutionary relationships between species.

(B) Schematic of approach to express TCOF1 in cells derived from zebrafish embryos.

(C) Fluorescence microscopy of cells derived from uninjected zebrafish embryos or from zebrafish embryos injected with mEGFP-TCOF1 mRNA or mEGFP-TCOF1Δ83–1170 mRNA. Nuclei are outlined (dotted line). Scale bars: 2 μm. Line profiles are of regions indicated by the finely dotted line for mEGFP (green) and Fibrillarin channels (red).

(D) Fluorescence microscopy of nucleoli of cells derived from zebrafish embryos injected with mCherry-zfMpp10 alone (top) to mark the GZ or co-injected with mCherry-zfMpp10 and mEGFP-WT TCOF1 (bottom). Scale bars: 1 μm. Line profiles are of regions indicated by the finely dotted line for mEGFP (green), Fibrillarin (red), and mCherry (blue).

(E) Model for the emergence and underlying structure of the FC.

See also Figure S7 and Table S1.

KEY RESOURCES TABLE

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
Antibodies		
Rabbit polyclonal anti-TCOF1	Proteintech	Cat#11003-1-AP; RRID:AB_2202516
Rabbit monoclonal anti-HA (C29F4)	Cell Signaling	Cat#3724S; RRID:AB_10693385
Mouse monoclonal anti-Tubulin	Thermo Scientific	Cat#MA5-16308; RRID:AB_2537819
Goat polyclonal secondary anti-rabbit HRP	Invitrogen	Cat#32260; RRID:AB_1965959
Goat polyclonal secondary anti-mouse HRP	Invitrogen	Cat#32230; RRID:AB_1965958
Rabbit polyclonal anti-UBTF	Novus Biologicals	Cat#NBP1-82545; RRID:AB_11032609
Rabbit polyclonal anti-POLR1A	Novus Biologicals	Cat#NBP2-56122
Mouse monoclonal anti-Fibrillarin	EMD Millipore	Cat#MABE1154
Rabbit polyclonal anti-MPP10	Novus Biologicals	Cat#NBP1-84341; RRID:AB_11026825
Mouse monoclonal anti-Nucleophosmin	Abcam	Cat#ab86712; RRID:AB_10675692
Rabbit polyclonal anti-TCOF1	Novus Biologicals	Cat#NBP1-86909; RRID:AB_11007370
Mouse monoclonal anti-HA	Invitrogen	Cat#26183; RRID:AB_10978021
Goat polyclonal secondary anti-rabbit Alexa 647	Invitrogen	Cat#A-27040; RRID:AB_2536101
Goat polyclonal secondary anti-mouse Alexa 647	Invitrogen	Cat#A-32728; RRID:AB_2633277
Goat polyclonal secondary anti-mouse Alexa 488	Invitrogen	Cat#A-32723; RRID:AB_2633275
Critical commercial assays		
Click-iT™ RNA Alexa Fluor™ 594 Imaging Kit	Fisher Scientific	Cat#C10330
mMESSAGE mMACHINE T7 ULTRA Transcription Kit	Invitrogen	Cat#AM1345
Experimental models: Cell lines		
Human: HeLa S3	ATCC	ATCC CCL-2.2
Human: TCOF1-degron HeLa Cells - Clone E6	This paper	N/A
Human: TCOF1-degron HeLa Cells - Clone E12	This paper	N/A
Experimental models: Organisms/strains		
Zebrafish: *Danio rerio*, AB/Tübingen (TAB5/14) genetic background	Koch Institute Zebrafish Core Facility	ZFIN: ZDB-GENO-010924-10
Oligonucleotides		
sgRNA: DNA sequence of sgRNA for generation of degron cell line: GTATGACGAGCACCAGCACC	This paper	N/A
Primer: PCR Screening for degron cell line validation, Forward: GCTGGCCTCCAGGGGGCAGGTGAA	This paper	RPO335

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Primer: PCR Screening for degron cell line validation, Reverse: ACAGGGGACACCAGAGCTGT	This paper	RPO338
Primer: Generation of PCR DNA for IVT Template, Forward: GCGTAATACGACTCACTATAGGGAGACCCAA	This paper	RPO378
Primer: Generation of PCR DNA for IVT Template, Reverse: GTGGATCCGAGCTCGGTACCAA	This paper	RPO379
Recombinant DNA		
pcDNA3.1(+) EGFP-NPM	This paper	RP14 (NPM1 WT)
pcDNA3.1(+) mEGFP - FBRL	This paper	RP162 (FBRL WT)
pcDNA3.1(+) mEGFP - RPA43	Lee, Jaberi-Lashkari et al.[13]	RP104 (RPA43WT)
pcDNA3.1(+) mEGFP - TCOF1	Lee, Jaberi-Lashkari et al.[13]	RP133 (TCOF1 WT)
pcDNA3.1(+) mEGFP - TCOF1 ΔSE (ΔS83-E97, S171-S178, D197-D211, E267-E282, E336-E352, E407-E422, E472-E487, E542-E559, S608-E622, E691-E704, E760-E773, E865-E882, D930-E944, S990-D1005, E1141-E1161, E1185-D1195)	This paper	RP136 (TCOF1 ΔSE)
pcDNA3.1(+) mEGFP - TCOF1 (ΔS83-G286)	This paper	RP165 (TCOF1ΔS83-G286)
pcDNA3.1(+) mEGFP - TCOF1 (ΔS83-K502)	This paper	RP166 (TCOF1ΔS83-K502)
pcDNA3.1(+) mEGFP - TCOF1 (ΔS83-V785)	This paper	RP167 (TCOF1ΔS83-V785)
pcDNA3.1(+) mEGFP - TCOF1 (ΔS83-H1170)	This paper	RP168 (TCOF1ΔS83-H1170)
pcDNA3.1(+) mEGFP - TCOF1 (ΔR688-S1234)	This paper	RP170 (TCOF1AR688-S1234)
pcDNA3.1(+) mEGFP - TCOF1 (ΔR405-S1234)	This paper	RP171 (TCOF1AR405-S1234)
pcDNA3.1(+) mEGFP - TCOF1 (ΔT167-S1234)	This paper	RP172 (TCOF1AT167-S1234)
pcDNA3.1(+) mEGFP - TCOF1 (ΔS503-Q687)	This paper	RP177 (TCOF1ΔS503-Q687)
pcDNA3.1(+) mEGFP - TCOF1 ΔSE 83-785 (ΔS83-E97, S171-S178, D197-D211, E267-E282, E336-E352, E407-E422, E472-E487, E542-E559, S608-E622, E691-E704, E760-E773)	This paper	RP192 (TCOF1ΔSE83-785)
pcDNA3.1(+) mCherry - TCOF1	This paper	RP197 (TCOF WT)
pcDNA3.1(+) mCherry - TCOF1 ΔSE (ΔS83-E97, S171-S178, D197-D211, E267-E282, E336-E352, E407-E422, E472-E487, E542-E559, S608-E622, E691-E704, E760-E773, E865-E882, D930-E944, S990-D1005, E1141-E1161, E1185-D1195)	This paper	RP198 (TCOFASE)
pcDNA3.1(+) mEGFP - TCOF1 (ΔK1390-K1406, K1438-K1468, K1476-K1483)	Lee, Jaberi-Lashkari et al.[13]	RP157 (TCOF1AK)
pcDNA3.1(+) mCherry - zfMpp10	This paper	RP226 (zfMpp10 WT)
pSpCas9(BB)-2A-GFP (PX458)	Ran et al.[34]	Addgene: 48138
Software and algorithms		
SnapGene	N/A	https://www.snapgene.com/
Fiji	Schindelin et al.[35]	https://imagej.net/software/fiji; RRID:SCR_002285
Uniprot BLAST	Uniprot	https://www.uniprot.org/blast
NCBI BLAST	NCBI	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Custom Python code for image analysis	This paper	https://doi.org/10.5281/zenodo.7236056