A Systems-Level Study Reveals Regulators of Membrane-less Organelles in Human Cells

Berchtold, Doris ; Battich, Nico ; Pelkmans, Lucas

Abstract: Membrane-less organelles (MLOs) are liquid-like subcellular compartments that form through phase separation of proteins and RNA. While their biophysical properties are increasingly understood, their regulation and the consequences of perturbed MLO states for cell physiology are less clear. To study the regulatory networks, we targeted 1,354 human genes and screened for morphological changes of nucleoli, Cajal bodies, splicing speckles, PML nuclear bodies (PML-NBs), cytoplasmic processing bodies, and stress granules. By multivariate analysis of MLO features we identified hundreds of genes that control MLO homeostasis. We discovered regulatory crosstalk between MLOs, and mapped hierarchical interactions between aberrant MLO states and cellular properties. We provide evidence that perturbation of pre-mRNA splicing results in stress granule formation and reveal that PML-NB abundance influences DNA replication rates and that PML-NBs are in turn controlled by HIP kinases. Together, our comprehensive dataset is an unprecedented resource for deciphering the regulation and biological functions of MLOs.

DOI: https://doi.org/10.1016/j.molcel.2018.10.036

A systems-level study reveals regulators of membrane-less organelles in human cells

Doris Berchtold[1], Nico Battich[1], Lucas Pelkmans[1,*]

[1]Institute of Molecular Life Sciences, University of Zurich, 8057 Zurich, Switzerland
*Lead Contact: lucas.pelkmans@imls.uzh.ch

Summary

Membrane-less organelles (MLOs) are liquid-like subcellular compartments that form through phase separation of proteins and RNA. While their biophysical properties are increasingly understood their regulation and the consequences of perturbed MLO states for cell physiology are less clear. To study the regulatory networks, we targeted 1,354 human genes and screened for morphological changes of nucleoli, Cajal bodies, splicing speckles, PML nuclear bodies (PML-NBs), cytoplasmic processing bodies and stress granules. By multivariate analysis of MLO features we identified hundreds of genes that control MLO homeostasis. We discovered regulatory crosstalk between MLOs, and mapped hierarchical interactions between aberrant MLO states and cellular properties. We provide evidence that perturbation of pre-mRNA splicing results in stress granule formation, and reveal that PML-NB abundance influences DNA replication rates and that PML-NBs are in turn controlled by HIP kinases. Together, our comprehensive dataset is an unprecedented resource for deciphering the regulation and biological functions of MLOs.

Introduction

Membrane-less organelles (MLOs) are subcellular compartments in the cyto- and nucleoplasm of eukaryotic cells. MLOs contain proteins with intrinsically disordered regions that together with RNA phase separate from the surrounding milieu (Banani et al., 2017; Brangwynne et al., 2009). These local concentrations of proteins and RNAs have two major functional implications. First, MLOs may concentrate components to facilitate biochemical reactions. Nucleoli, for instance, are nucleated at sites of actively transcribed ribosomal RNAs and control most aspects of ribosome biogenesis (Sirri et al., 2007). Likewise, Cajal bodies (CBs) are small nuclear domains that form on active loci of small nuclear (sn)RNA transcription and are typically present in highly proliferative cells that have a high demand in splicing (Cioce and Lamond, 2005; Machyna et al., 2012). The second functional implication of MLOs is that they could act as dynamic buffers for both protein and RNAs (Saunders et al., 2012; Stoeger et al., 2016). Nuclear splicing speckles, for instance, are enriched in spliceosomal snRNPs and components of the pre-mRNA splicing machinery but are not considered as sites of pre-mRNA splicing. Instead, they might constitute storage sites of splicing factors to supply them to adjacent active transcription sites according to need (Spector and Lamond, 2011). Analogous to splicing speckles, nuclear bodies that form

around promyelocytic leukemia protein (PML-NBs) harbor proteins involved in transcriptional regulation, DNA damage response and apoptosis and are considered as storage sites and hubs for protein modification (Bernardi and Pandolfi, 2007; Hsu and Kao, 2018; Lallemand-Breitenbach and de The, 2010). Last, cytoplasmic processing bodies (PBs) form around translationally repressed mRNAs and were initially considered as sites of mRNA degradation as they contain numerous proteins associated with mRNA deadenylation, decapping, and the 5'-to-3' mRNA decay pathway (Luo et al., 2018; Sheth and Parker, 2003). However, there is increasing evidence that mRNA degradation might not occur in PBs and that PBs are rather storage sites for repressed mRNAs that can be released and translated according to the cell's needs (Horvathova et al., 2017; Hubstenberger et al., 2017).

Most constitutive MLOs are regulated according to the physiological state of the cell. Size and abundance of MLOs change along the cell cycle, and morphology and composition of MLOs are altered upon stress (Boulon et al., 2010; Courchaine et al., 2016; Dellaire and Bazett-Jones, 2004). One key factor that contributes to MLO formation is the concentration of both protein and RNA components. Active rDNA transcription, for instance, is essential for nucleolar assembly (McCann and Baserga, 2014), and depletion of proteins involved in ribosome biogenesis leads to diminished nucleolar number, as discovered by a recent genome-wide siRNA screen (Farley-Barnes et al., 2018). Another key factor in the assembly of MLOs are posttranslational modifications of proteins, such as phosphorylation, that can alter the number of transient interactions between MLO components and thus, influence the phase separation of MLO scaffolds or the recruitment of individual components
(Bah and Forman-Kay, 2016; Dundr and Misteli, 2010; Hebert, 2013; Hernandez-Verdun, 2011). However, only few kinases and phosphatases have been identified as MLO regulators so far. We previously discovered that DYRK3 kinase controls the dissolution of stress granules upon stress relief likely by phosphorylating multiple RNA-binding proteins, and is required for splicing speckle disassembly in pre-mitotic cells (Wippich et al., 2013; Rai et al., 2018). Similarly, MBK-2, the *C. elegans* homologue of DYRK3, controls the dissolution of P-granules (Wang et al., 2014) and over expression of the human dual-specificity kinases CLK and DYRK1A leads to a phosphorylation-dependent disassembly of splicing speckles (Alvarez and la Luna, 2003; Sacco-Bubulya, 2002).

Here, we present a parallel image-based RNAi screening approach with single cell resolution, in which we targeted 1,354 genes including the kinome and phosphatome, as well as multiple known MLO components to uncover regulators of six major MLOs in human cells, similar to an approach we previously applied to membrane-bound organelles (Liberali et al., 2014). We applied computer vision and machine learning methods to quantify the morphological changes of nucleoli, CBs, splicing speckles, PML-NBs, cytoplasmic PBs and SGs in thousands of single cells per perturbation. Single cell clustering allowed us to identify more than 500 genes involved in multiple biological processes that regulate one or more MLOs. We identified gene perturbations that provoke the assembly of SGs and discovered a link between perturbation of pre-mRNA splicing and SG formation. Moreover, our approach

allowed us to uncover co-occurrence of perturbed MLO states and their functional consequences on the cellular state. In particular, we found that the abundance of PML-NBs controls the rate of DNA replication, and that size and composition of PML-NBs are regulated by the dual-specificity kinases HIPK1 and HIPK2.

Results

Image-based RNAi screens on six membrane-less organelles with single-organelle resolution

To uncover the regulatory complexity underlying the homeostasis of MLOs, we targeted 1,354 human genes with pools of three siRNAs per gene and stained the cells against key marker proteins of nucleoli, CBs, PML-NBs, splicing speckles, PBs and SGs (Figures 1A and 1B). The gene library contained protein kinases and phosphatases, and known MLO components that are modified by phosphorylation (Figures 1C, S1A and Table S1). After imaging thousands of cells per condition, we applied computer vision approaches to segment nuclei and cells, and extracted multiple single cell features such as size and shape, or protein and DNA content. We used support vector machine classification (Ramo et al., 2009) for quality control and data cleanup on multiple levels (Figure S1B). We fully excluded 31 perturbations (2.3 percent of the gene library) that strongly impaired cell viability (Figure S1C). Cell numbers showed no plate position bias and both cell numbers and cell cycle classifications were highly reproducible between the screens (Figures S1D-S1F). 103 gene perturbations (7.6 percent of the gene library) displayed significantly altered fractions of cells in G1, S and G2 in all three screens along with reduced fitness (Figures S1G-S1I). In order to quantify MLO morphology in each of the single cells we used a custom pixel classification-based software to accurately segment individual MLOs across a wide range of intensities, sizes, shapes, and background signals (Figures 1D and 1E). We combined segmented MLOs per cell and extracted phenotypic features (see STAR Methods). All cell and MLO intensity measurements were corrected for plate positional staining biases on the single-cell level. Finally, we confirmed that the resulting numbers of MLOs per cell derived from more than 700,000 unperturbed cells were in agreement with previous reports (Figure S1J and (Mao et al., 2011)).

Identification of gene perturbations with aberrant MLO morphologies

The number and morphology of MLOs is to a large extent determined by the cell cycle and thus highly heterogeneous even in unperturbed cell populations. To account for this cell-to-cell variability, we analyzed the screens by means of unsupervised clustering of single cells based on their morphological MLO features (Figure 2A, STAR Methods, and (Van Gassen et al., 2015)). Each MLO was analyzed separately. For instance, about 3.7 million single cells from unperturbed and perturbed conditions of the nucleoli screen were analyzed together and clustered into 30 phenotypic nodes according to their nucleolar features (Figures 2B and S2A).

First, we explored the cell cycle-dependent MLO heterogeneity of unperturbed cells using cell cycle trajectories (CCTs) (Gut et al., 2015). CCTs are inferred from the cellular state information of fixed cell populations and reflect cell cycle progression. We mapped features describing the cellular state as well as features of MLOs on the CCTs (Figures 2C, and S2B-S2D) and observed their largest fluctuations during S phase. The intensity of the nucleolar marker NPM, for example, transiently drops in early S-phase cells but recovers towards late S phase (Figure 2C), a phenomenon that can not be revealed by simply binning cells into G1, S and G2 phases (Figure S2E). Consistently, unperturbed cells in early S phase were assigned to different phenotypic nodes than unperturbed cells in late S phase (Figure 2D). These data underscore the high quality of our screens and the sensitivity of our single cell clustering approach to distinguish even subtle phenotypic changes of MLOs.

Next, we calculated the median MLO feature values from all perturbed and unperturbed cells that were sorted into one phenotypic node and subsequently clustered the phenotypic nodes of each MLO (Figures 3A, and S3A-S3D). When we analyzed how cells from control populations distribute over the nodes we found that unperturbed cells primarily enrich in nodes with intermediate feature values and are absent in nodes with rather extreme MLO features (Figures 3B, S3A-S3D, and Table S2). In order to identify gene perturbations that drastically alter MLO morphology we then defined two globally perturbed states for each MLO, namely an increased and a decreased state, which each contain cells from one or more phenotypic nodes (see STAR Methods). Notably, exclusively for P-bodies and nucleoli we were able to discern a third perturbed MLO state, reflecting the absence of MLOs similar to the decreased state but with elevated levels of the marker proteins localizing diffuse in the cyto- or nucleoplasm, respectively. We then calculated for each gene perturbation the fraction of single cells that was clustered into the nodes belonging to each of the two or three perturbed MLO states, and used these values for further analysis (Figures 3B, S3A-S3D, and Table S1). For instance, 72 percent of *NPM1* knock down cells were found in nodes of the perturbed state ‘decreased NPM’ while the remaining cells were clustered into nodes with intermediate nucleolar features that do not reflect a perturbed state. In contrast, down regulation of *RPL11* results in ‘increased nucleoli’ for 97 percent of the cells. Gene perturbations resulting in perturbed MLO states for fractions higher than 2.5 standard deviations from the mean across the whole screen were then considered as hits (Figures 3C-3G). For all MLOs, the gene perturbation of the respective marker protein was identified among the highest scoring hits in the ‘decreased MLO marker’ state. Importantly, our single-cell clustering analysis has the potential to identify more than the perturbed MLO states portrayed here. For instance, node 5 of the Cajal body (CB) screen contains cells with fragmented CBs (high CB numbers without altered CB intensities), and node 11 of the splicing speckle screen contains cells with SRRM2 droplets in the cytoplasm (Figures S3E, S3F, and Table S2).

To assess the technical reproducibility of our analysis we compared how cells of three replicate conditions distributed over the nodes (Figures S3G-S3H). Reassuringly, the

fractions of single cells assigned to each phenotypic node were nearly identical between the biological replicates. Finally, to evaluate the biological consistency of our results with previous literature, we visualized 48 genes that are present in our library and are known components of P-bodies as a network in which the color and shape of nodes indicate PB morphology (Figure 3H). We identified *LSM14A*, *PATL1* (Pat1b) and *EDC4* (Ge-1) among the genes required for PB assembly as previously reported (Ozgur et al., 2010; Yang et al., 2006; Yu et al., 2005). We further identified four genes coding for proteasomal subunits (*PSMA2*, *PSMA4*, *PSMA6* and *PSMC2*) as positive regulators of PBs, supporting previous findings that exposure of cells to the proteasome inhibitor MG132 resulted in PB dissolution (Mazroui, 2007). Moreover, we found that depletion of the mRNA decapping factors *DCP1A* and *DCP2* leads to increased PB formation, as it was previously shown in yeast (Sheth and Parker, 2003). Together, these results provide confidence that our single-cell clustering approach is of high quality both in terms of technical reproducibility and biological outcome.

Perturbation of pre-mRNA splicing leads to stress granule formation

Stress granules (SGs) are typically absent in unperturbed cells and form either when cells experience stress or upon over expression of certain SG components, such as G3BP-1 or TIA-1 (Anderson and Kedersha, 2002; Anderson and Kedersha, 2009; Dewey et al., 2011). We included the SG marker G3BP-1 to screen for the formation of SGs upon gene perturbation. We did detect some SG-containing cells among unperturbed cell populations, which we used to train a pixel classification-based model for SG segmentation. Although SGs were accurately segmented (Figure 1D), we could not entirely prevent the classifier to also segment granule-like G3BP-1 staining resembling cytoplasmic blebs. However, cluster analysis of the morphological features of the segmented objects allowed the separation of gene perturbations resulting in blebs from perturbations resulting in actual SGs (Figures 4A, S4A-S4D, and STAR Methods). We identified 53 genes, whose perturbation caused a marked increase in the fraction of cells with SGs as compared to unperturbed cells (Figures 4B and 4C). Remarkably, when we explored the biological function of these genes, we found that eight out of 53 genes encode for either spliceosome components or contribute to its assembly, seven genes function as splicing regulators and six genes play roles in RNA editing, processing and export, respectively (Figure 4D). To corroborate this finding, we treated cells with either mRNA transcription or mRNA splicing inhibitors. Strikingly, in all three cell lines tested we observed formation of SGs in about 5-15 percent of cells when pre-mRNA splicing is inhibited, but not in cells where mRNA transcription is blocked (Figures 4E, 4F, and S4E). In addition, we performed fluorescence *in situ* hybridization in these cell lines to detect polyA-mRNAs. While polyA-mRNAs are almost absent in transcriptionally inhibited cells, they accumulate in the nucleus of cells treated with splicing inhibitors and can be detected in cytoplasmic SGs in HeLa and A-431 cells (Figure S4F). Thus, our screen revealed a previously unrecognized functional link between pre-mRNA splicing and SG formation, which can be observed in various cell lines originating from different tissues.

Shared and distinct regulatory pathways control MLO formation

One main goal of generating this dataset was to reveal systems-level properties that go beyond the identification of particular genes involved in the regulation of one MLO but emerge from the cross-comparable scoring of genes. First, we calculated functional annotation enrichments for all identified hits per MLO separately and visualized them as networks of gene ontology (GO)-terms (Figures 5A-5F, S5A and S5B). Comparing the resulting patterns revealed that the six different MLOs have distinct as well as shared control pathways. To relate the functional enrichments to individual genes we created interaction networks of the hit genes in which the color and shape of nodes indicates the respective MLO morphology upon perturbation (Figures 5G-5K). The genes whose perturbation induces SG formation are particularly enriched in splicing-related functions, but also in transcription, mRNA binding, processing, and transport (Figure 5A), and display eleven-fold enrichment in the KEGG pathway term 'spliceosome' (Figure S5C). Genes that regulate nucleolar morphology are also enriched in splicing-related functions, but in addition cover a wide range of other functions, such as transcriptional regulation, proteasome and APC-dependent protein degradation, and Wnt and other signaling pathways (Figures 5B and 5G). For both PBs and Cajal bodies (CBs), we observed a particular functional enrichment in second messenger signaling (Figures 5C and 5E). We found that down-regulation of the cAMP phosphodiesterases PDE4C and PDE3B leads to increased PBs, while several other phosphodiesterase-encoding genes (*PDE2A*, *PDE3A*, *PDE6A*, *PDE6D*) as well as *PRKAA1* (AMPK) are required for PB assembly (Figure 5H). A study in yeast previously linked cAMP signaling to PB regulation by demonstrating that cAMP-dependent protein kinase (PKA) could dissociate PBs via phosphorylation of Pat1 (Ramachandran et al., 2011) indicating that this regulation is evolutionary conserved. For CBs, we identified *PRKACB*, the catalytic subunit of PKA, and *PRKX*/*PKX1*, a protein kinase with similarity to PKA, to be required for CB formation (Figure 5J). Unexpectedly, we found that genes whose silencing resulted in altered splicing speckle morphology show less functional enrichment in splicing as compared to genes whose silencing affected nucleolar morphology or resulted in SG formation (Figure 5D). Instead, regulators of splicing speckles show a strong enrichment in cell cycle-related functions. We identified Aurora kinase A (*AURKA*) and *CDC25B* (a phosphatase and substrate of Aurora-A) among other centrosome-related genes, as well as genes coding for regulatory subunits of the phosphatase PP2A (*PPP2R2B*, *PPP2R2C* and *PPP2R2D*) that was shown to regulate Aurora-A stability in mitosis (Figure 5I, and Horn et al., 2007). Last, for PML nuclear bodies (PML-NBs) we observed a strong enrichment for plasma membrane/endosome-related terms (Figures 5F and S5C). Among the strongest negative regulators are *PIP5K1C*, a phosphatidylinositol-4-phosphate 5-kinase, and *PIK3CG*, the catalytic subunit of phosphatidylinositol-4,5-bisphosphate 3-kinase, as well as *INPP5F*, an inositol polyphosphate-4-phosphatase (Figure 5K). Consistently, we also identified the growth factor receptor *TGFBR1*, which acts upstream of these signaling kinases, as a negative regulator of PML-NB formation. Previous work has shown that TGF-β induces the expression of a

cytoplasmic isoform of PML, which is required for the accumulation of TGFBR1 in early endosomes (Lin et al., 2004), but how that connects to the control of PML-NB abundance remains to be investigated. Thus, our screen provides a resource for exploring numerous biological processes that were previously not linked to the regulation of MLOs. In addition, our data demonstrates that some genes act specifically on one MLO, whereas other regulatory pathways control more than one MLO, indicating regulatory crosstalk.

Co-occurrence of deregulated MLO states

To investigate this coordination in more detail, we asked whether two or more perturbed states of MLOs co-occur and which gene perturbations are causing them. We pooled all genes that were scored as a hit in at least one of the six MLO screens. For these 453 genes we used the phenotypic scores of the twelve different perturbed MLO states (Table S1) to compile a dataset that can be visualized as a two-dimensional t-SNE map (Figures 6A, S6A, S6B, and STAR Methods). In such a map, genes are positioned relative to all other genes based on their similarity in effects across the various perturbed MLO states. By coloring these genes according to their phenotypic scores for each of the 12 perturbed MLO states and comparing the resulting patterns, we discovered partial co-occurrences. For instance, a subset of gene perturbations that lead to increased PML-NBs do not form PBs (Figure 6A, encircled in left panels). Another example is a set of gene perturbations that lead to both disturbed nucleoli (NPM diffuse in the nucleoplasm) and increased formation of CBs (Figure 6A, encircled in right panels). We next calculated how often a gene was scored as a hit for any of the perturbed MLO states (including SG formation) and highlighted the result on the gene t-SNE map (Figure 6B). Of the 453 gene perturbations more than 70 percent (325 genes) affected only one MLO, 20 percent (92 genes) affected two MLOs and less than 10 percent (36 genes) affected the morphology of three or more MLOs (Figure S6C). The most abundantly perturbed MLO states observed for the 128 gene perturbations with pleiotropic effects were increased nucleoli, followed by SG formation, increased CBs, and increased splicing speckles (Figure S6D).

We next calculated which of the co-occurrences of perturbed MLO states are statistically significant and visualized the results as a network in which the edge width indicates the number of genes that are hits for both perturbed MLO states (Figure 6C). This shows that both the co-occurrence between increased PML-NBs and absence of PBs, and between disturbed nucleoli and increased CBs (as highlighted in Figure 6A) were significant. In addition, the network reveals an interesting anti-correlation between nucleolar (light blue nodes) and PB (dark blue nodes) morphology. Some gene perturbations that lead to increased nucleoli also display dissolved PBs where DDX6 is either cytoplasmic diffuse or present at very low levels (Figure 6D, groups 1 and 2). Conversely, gene perturbations leading to increased PBs negatively affect nucleoli (decreased NPM; Figure 6D, group 3). Since the major function of nucleoli is to regulate ribosome biogenesis, this finding could

suggest the existence of a system that adjusts ribosome availability to the needs of mRNA translation.

Hierarchical functional interactions between perturbed MLOs and cellular states

An important and often still poorly understood question is how MLOs are integrated into cellular physiology. While the uncovered genetic networks controlling the formation of MLOs may suggest certain links, they do not directly infer this from the data. Our multivariate image-based approach extracts besides multiple properties of MLOs also multiple features describing the phenotypic state and physiology of cells, allowing us to directly map such interactions. To achieve this, we applied a statistical method termed the hierarchical interaction score (HIS) that we previously developed to infer hierarchical interactions from multivariate datasets (Snijder et al., 2013). We calculated the HIS between 13 perturbed MLO states and 17 features describing properties of the cellular state and physiology derived from 1,326 gene perturbations and 218 control populations (see STAR Methods). We obtained 50 functional interactions with HIS scores greater than zero and visualized them as a hierarchical interaction network (Figure 6E). It is important to note that the inferred directionality of the interaction between two properties can be interpreted as the statistical likelihood of a cell having the downstream property if a gene perturbation caused the upstream property and does not by itself reveal a direct causality.

To substantiate the inferred functional interactions we explored three examples from the HIS network in more detail. We plotted the values of cellular state properties on the t-SNE map that was generated from data on perturbed MLO states and does not contain any information about the cellular state (see STAR Methods). By doing so, we observed patterns of cellular states that explain the inferred hierarchy of interactions. For example, cell area is increased for the majority of gene perturbations but only a subset also has increased CBs (Figures 6F and 6G). In contrast, almost all gene perturbations that lead to increased CBs also display increased cell area (Figure 6H). Another example is the hierarchical interaction between increased PML-NBs and the fraction of cells in S phase. While the majority of gene perturbations in the t-SNE map displays reduced fractions of S phase cells only a subset have increased PML-NBs (Figures 6I and 6J). Conversely, all gene perturbations that lead to increased PML-NBs have lower fractions of S phase cells (Figure 6K). As a third example, we highlight the directionality between the absence of PBs and increased cytoplasmic protein concentration (Figures 6L-6N). Almost all gene perturbations that result in the dissolution of PBs with DDX6 being diffuse in the cytoplasm have a higher cytoplasmic protein concentration, while not all gene perturbations that lead to high cytoplasmic protein concentration have this perturbed PB state.

Increased PML-NBs cause a delay in DNA replication

To characterize one of these interactions in more detail, we decided to explore the hierarchical interaction between increased PML-NBs and the fraction of cells in S phase.

First, we analyzed whether this correlation holds true for single cells. Indeed, when we compared cell populations of hit genes for increased PML-NBs to cells of gene perturbations that lead to G1 arrest we observed the same hierarchical interaction as shown before. Single cells with increased PML-NBs have a low level of 5-ethynyl-2'-deoxyuridine (EdU) incorporation (Figure 7A). G1-arrested cells, in contrast, though displaying a similar reduction in EdU intensity, do not have increased PML-NBs.

We next included the DNA content to explore whether cells with increased PML-NBs are arrested in G1 (Figures 7B and S7A). Surprisingly, we found that many cells with increased PML-NBs have the DNA content of S phase cells but very low levels of EdU incorporation (dashed black box in Figure 7B) suggesting that these cells have a slower DNA replication rate. To confirm these findings in independent experiments, we targeted five genes that are both components of PML-NBs and the strongest negative regulators of PML-NBs (Figure S7B) with siRNAs in two cell lines and stained the cells with antibodies against two PML-NB markers, PML and Sp100. We reproduced the findings of the screen in HeLa cells and confirmed that cells with increased PML-NBs also contained elevated levels of Sp100, as well as reduced EdU incorporation (Figures 7C-7E). In A-431 cells three out of five gene perturbations resulted in elevated fractions of cells with increased PML-NBs and the reduced EdU incorporation was observable albeit less pronounced due to the smaller proportion of cells with increased PML-NBs (Figure S7C and S7D). To verify that EdU incorporation rates are slower in cells with increased PML-NBs we allowed cells to incorporate EdU for 60 minutes instead of 15. Indeed, cells with increased PML-NBs displayed higher EdU intensity as compared to the 15 minute pulse but the levels remained lower as compared to control cells (Figure 7F), underscoring that aberrantly increased PML-NBs correlate with decreased DNA replication rates.

Finally, we addressed the directionality inferred by the HIS network. If increased PML-NBs were the cause and not the consequence of a slower DNA replication rate, other perturbations of DNA replication would not lead to increased PML-NBs. To test this, we treated cells with either hydroxyurea (HU) or deoxythymidine (dT), which both stall replication forks. EdU incorporation was abolished in cells treated with either drug and the replication factor RPA1/p70 changed from a diffuse to a speckled nuclear pattern localizing with PCNA at stalled replication forks (Figure 7G) as previously described (Urban et al., 2017). However, we did not observe increased PML-NBs in cells treated with HU or dT (Figures 7H and 7I). HU- and dT-treated cell populations displayed the same fraction of cells with increased PML-NBs as control populations (below five percent) and did not reach the values observed for cell populations depleted of *HIPK1* or *HIPK2* (grey dotted boxes in Figure 7I). Consistently, genetically perturbing the DNA helicases BLM (Bloom syndrome protein) or WRN (Werner syndrome ATP-dependent helicase), which promote replication fork movement (Urban et al., 2017), did not lead to increased PML-NBs (Table S1). This indicates that aberrantly large PML-NBs are likely to cause slower DNA replication rates and that homeodomain-interacting protein kinases (HIPKs) are required to prevent this effect.

HIPKs are known to interact with and phosphorylate PML-NB components (Rinaldo et al., 2008) and Figure S7B), and are also members of the dual-specificity protein kinase family, like DYRK3 and MBK-2, which promote the dissolution of other MLOs in a kinase-dependent manner (Wippich et al., 2013; Rai et al., 2018; Wang et al., 2014). To test whether HIPKs perform their action in a similar manner as DYRK3 by promoting the dissolution of PML-NB constituents through phosphorylation we overexpressed HIPK1 and HIPK2 and quantified their effect on PML-NBs. Intriguingly, and consistent with previous work (Ecsedy et al., 2003; Engelhardt et al., 2003), HIPK1 and HIPK2 both partition into PML-NBs (Figures S7E and S7F). While their overexpression does not cause a dissolution of the scaffold protein PML (Figures S7G-S7I) it does dissolve other components of PML-NBs, such as Sp100, in a concentration-dependent manner (Figures S7J-S7L). Thus, HIPKs can be added to the growing list of dual-specificity kinases that control the condensation of proteins into MLOs, which in their case are PML-NBs.

Discussion

Biological processes that regulate MLO formation

Among the many biological processes we found to control MLO formation, one unexpected finding was the discovery that several genes involved in growth factor signaling at the plasma membrane impact on the abundance of PML nuclear bodies (PML-NBs). Although the molecular aspects of this control remain to be elucidated it suggests a mechanism by which growth factor signaling induces cell proliferation through modulating PML-NB size, which may act upstream of DNA replication rates, as indicated by our findings. As another example, we identified multiple genes involved in pre-mRNA splicing whose genetic perturbation induced SG formation. We currently do not know whether SG formation is induced by a leakage of unspliced pre-mRNA into the cytoplasm, or by the protein products of these unspliced mRNAs. One mechanism by which cells could link spliceosome integrity to SG formation would involve a protein sensor that shuttles between the nucleus and the cytoplasm. Indeed, some SG components, such as TIA-1, regulate splicing in the nucleus under normal conditions and might translocate to the cytoplasm upon stress where they contribute to SG formation. Thus, the nuclear-cytoplasmic translocation of certain RNA-binding proteins could be a feedback mechanism by which cells sense splicing stress and arrest mRNA translation.

Regulatory crosstalk between MLOs

The translocation of MLO components could be a general principle by which cells sense stress in one MLO and signal it to other, functionally related MLOs. For example, previous work has shown that inhibition of transcription induces the reorganization of nucleoli and the formation of nucleolar caps, which recruit numerous non-resident proteins including coilin and PML, the scaffold proteins for CBs and PML-NBs (Boulon et al., 2010; Shav-Tal et al., 2005). The recruitment of coilin and PML to nucleolar caps might in turn alter the integrity and

function of CBs and PML-NBs, respectively. Here, we identified an anti-correlation between the morphology of nucleoli and cytoplasmic P-bodies (PBs), which likewise could be achieved by relocating components between the two MLOs. PATL1, for instance, is primarily a PB component and required for PB formation (Figure 3H). However, PATL1 shuttles between the nucleus and the cytoplasm and, upon inhibition of transcription, was found in nucleolar caps when its nuclear export was prevented (Marnef et al., 2011). Thus, aberrantly perturbed nucleoli might segregate PATL1, and its nucleolar detention then disrupts PB assembly in the cytoplasm. Regulatory crosstalk might also be mediated by RNAs. In the case of dysfunctional ribosome biogenesis, higher levels of ribosome-free mRNAs in the cytoplasm might induce an increased formation of PBs. Indeed, treatment of cells with puromycin, which inhibits translation by triggering premature release of mRNAs from ribosomes, enhances P-body assembly (Eulalio et al., 2007). Such crosstalk between nucleoli and PBs might be an elegant mechanism to adapt the production of ribosomes to the levels of mRNAs. Importantly, our identification of regulatory crosstalk between different MLOs complements recent advances in deciphering the molecular composition of MLOs (Fong et al., 2013; Hubstenberger et al., 2017; Youn et al., 2018). It will be interesting to monitor such compositional changes upon stress or gene perturbations that together with our findings might illuminate how MLOs sense and react to stressful conditions and how this impacts physiological processes in the cell.

Functional consequences of deregulated MLO morphology

A particularly unique nature of our systems-level study is the ability to infer interactions between MLO states and physiological properties of cells. In general, if MLOs serve as ‘bioreactors’ that concentrate components to facilitate biochemical reactions, their absence would decrease the efficiency of reactions and reduce proliferation rates and/or the cell’s ability to respond to stress, while an aberrant increase could imply increased production rates of, for instance, RNPs. If MLOs serve as storage compartments, however, the up regulation of the scaffold proteins or core components might enhance the recruitment of other MLO residents through multivalent interactions, and thus deplete the cell from these molecules. If in this case the proteins or RNAs function elsewhere in the cell, abnormally increased MLOs have negative implications for cellular processes. Our data revealed that cells with increased PML-NBs have decreased DNA replication rates. Multiple scenarios might explain our observation. First, PML-NBs contain many proteins required for DNA replication and repair (Van Damme et al., 2010) hence increased PML-NBs could lead to increased segregation or residence time of replication or DNA-repair factors. Second, PML-NBs are considered as platforms for protein modification, in particular sumoylation (Bernardi and Pandolfi, 2007; Van Damme et al., 2010). Sumoylation is a key modification for factors involved in DNA replication and repair (Lecona and Fernandez-Capetillo, 2016). Increased PML-NBs might interfere with the sumoylation and de-sumoylation cycles of replication factors and thus slow down DNA replication progression. Sentrin-specific proteases (SENPs) catalyze the maturation of SUMO

proteins and de-conjugation of SUMO-linked proteins, and knock down of *SENP1, SENP2* or *SENP6* leads to increased PML-NBs and accumulation of SUMO1 and SUMO2 in PML-NBs (Figure S7B, and Hattersley et al., 2011; Yates et al., 2008). This suggests that PML-NB homeostasis requires turnover of SUMO. Last, recent work demonstrated a key role for the ubiquitin hydrolase USP7/HAUSP in regulating replication fork progression through de-ubiquitinating SUMO-2 (Lecona et al., 2016). We identified USP7 as a negative regulator of PML-NBs (Figure 5K). Since USP7 partially colocalizes with PML-NBs and is repressed by the PML-NB component DAXX, increased PML-NBs might segregate and inhibit USP7, which consequently decreases DNA replication rates. Although untangling the exact molecular mechanisms of how increased PML-NBs impact on replication rates requires further studies, this example illustrates the importance of controlling the properties of MLOs in this process.

Dual-specificity kinases as key regulators of MLO formation

Our finding that depletion of HIPK1 or HIPK2 leads to increased PML-NBs strengthened a key role for HIPKs in the control of PML-NB homeostasis and/or composition. HIPKs belong to the DYRK kinase family and share some interesting parallels with DYRK kinases (van der Laden et al., 2015). First, HIPKs and DYRKs seem to phosphorylate several MLO components which may contribute to global changes in MLO size and composition. Second, HIPKs and DYRKs have intrinsically disordered domains and low complexity regions, by which they might associate with specific MLOs. This could be one mechanism for homeostatic size control of MLOs in that larger MLOs would recruit more kinases, which then counteract a further size increase. Third, DYRKs and HIPKs are constitutively active but appear to be tightly regulated by proteasomal degradation. We recently demonstrated that DYRK3 is present at low levels in G1 and S phase but its presence increases in G2 and during mitosis where it contributes to the dissolution of splicing speckles (Rai et al., 2018). HIPK2 is largely unstable in unperturbed cells and becomes stabilized upon DNA damage, which might contribute to the compositional and spatial rearrangements observed for PML-NBs (Dellaire and Bazett-Jones, 2004; Eskiw and Bazett-Jones, 2003; Winter et al., 2008). Having such broad impact on diverse MLOs, low levels of DYRKs and HIPKs might be sufficient under normal conditions to achieve homeostasis of splicing speckles and PML-NBs, respectively. Thus, our finding that HIPKs act as regulators of PML-NB homeostasis adds to the growing importance of DYRK family kinases as central regulators of MLOs.

Acknowledgments

We thank all members of the Pelkmans and Klemm labs for critical comments on the manuscript. We thank R. Holtackers, Y. Yakimovich and R. Murri for technical support. We thank K. Shin-ya (AIST, Japan) for providing FR901464, and K. Koide (University of Pittsburgh, USA) for providing Meayamycin. This work was supported by an ERC-replacement Consolidator Grant from the Swiss National Science Foundation to L.P., and a Human Frontier Science Program Long-Term Fellowship to D.B..

Author Contributions

Conceptualization, D.B. and L.P.; Software, D.B. and N.B.; Validation, D.B.; Formal Analysis, D.B.; Investigation, D.B.; Writing – Original Draft, D.B.; Writing – Review & Editing, D.B., N.B. and L.P.; Visualization, D.B.; Funding Acquisition, D.B. and L.P.

Declaration of Interests

The authors declare no competing interests.

References

Alvarez, M., and la Luna, de, S. (2003). DYRK1A accumulates in splicing speckles through a novel targeting signal and induces speckle disassembly. Journal of Cell Science *116*, 3099–3107.

Anderson, P. and Kedersha, N. (2002). Stressful initiations. Journal of Cell Science *115*, 3227–3234.

Anderson, P., and Kedersha, N. (2009). RNA granules: post-transcriptional and epigenetic modulators of gene expression. Nature Publishing Group *10*, 430–436.

Bah, A., and Forman-Kay, J.D. (2016). Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. Journal of Biological Chemistry *291*, 6696–6705.

Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. Nature Reviews Molecular Cell Biology *18*, 285–298.

Bernardi, R., and Pandolfi, P.P. (2007). Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. Nat Rev Mol Cell Biol *8*, 1006–1016.

Boulon, S., Westman, B.J., Hutten, S., Boisvert, F.-M., and Lamond, A.I. (2010). The Nucleolus under Stress. Molecular Cell *40*, 216–227.

Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. Science *324*, 1729–1732.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biology 7:R100.

Cioce, M., and Lamond, A.I. (2005). CAJAL BODIES: A Long History of Discovery. Annu. Rev. Cell Dev. Biol. *21*, 105–31.

Courchaine, E.M., Lu, A., and Neugebauer, K.M. (2016). Droplet organelles? Embo J. *35*, 1603–1612.

Dellaire, G., and Bazett-Jones, D.P. (2004). PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. Bioessays *26*, 963–977.

Dewey, C.M., Cenik, B., Sephton, C.F., Dries, D.R., Mayer, P., Good, S.K., Johnson, B.A., Herz, J., and Yu, G. (2011). TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. Molecular and Cellular Biology *31*, 1098–1108.

Dundr, M., and Misteli, T. (2010). Biogenesis of Nuclear Bodies. Cold Spring Harbor Perspectives in Biology *2*, a000711.

Ecsedy, J.A., Michaelson, J.S., and Leder, P. (2003). Homeodomain-interacting protein kinase 1 modulates Daxx localization, phosphorylation, and transcriptional activity. Molecular and Cellular Biology *23*, 950–960.

Engelhardt, O.G., Boutell, C., Orr, A., Ullrich, E., Haller, O., and Everett, R.D. (2003). The homeodomain-interacting kinase PKM (HIPK-2) modifies ND10 through both its kinase domain and a SUMO-1 interaction motif and alters the posttranslational modification of PML. Experimental Cell Research *283*, 36–50.

Eskiw, C.H., and Bazett-Jones, D.P. (2003). Size, position and dynamic behavior of PML nuclear bodies following cell stress as a paradigm for supramolecular trafficking and assembly. Journal of Cell Science *116*, 4455–4466.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Molecular and Cellular Biology *27*, 3970–3981.

Farley-Barnes, K.I., McCann, K.L., Ogawa, L.M., Merkel, J., Surovtseva, Y.V., and Baserga, S.J. (2018). Diverse Regulators of Human Ribosome Biogenesis Discovered by Changes in Nucleolar Number. CellReports *22*, 1923–1934.

Fong, K.W., Li, Y., Wang, W., Ma, W., Li, K., Qi, R.Z., Liu, D., Songyang, Z., and Chen, J. (2013). Whole-genome screening identifies proteins localized to distinct nuclear bodies. The Journal of Cell Biology *203*, 149–164.

Gut, G., Tadmor, M.D., Pe'er, D., Pelkmans, L., and Liberali, P. (2015). Trajectories of cell-cycle progression from fixed cell populations. Nat Meth *12*, 951–954.

Hattersley, N., Shen, L., Jaffray, E., and Hay, R.T. (2011). The SUMO protease SENP6 is a direct regulator of PML nuclear bodies. Molecular Biology of the Cell *1*, 78–90.

Hebert, M.D. (2013). Signals controlling Cajal body assembly and function. International Journal of Biochemistry and Cell Biology *45*, 1314–1317.

Hernandez-Verdun, D. (2011). Assembly and disassembly of the nucleolus during the cell cycle. Nucleus *2*, 189–194.

Horn, V., Thelu, J., Garcia, A., Albiges-Rizo, C., Block, M.R., Viallet, J. (2007). Functional Interaction of Aurora-A and PP2A during Mitosis. Molecular Biology of the Cell *18*, 1233–1241.

Horvathova, I., Voigt, F., Kotrys, A.V., Zhan, Y., Artus-Revel, C.G., Eglinger, J., Stadler, M.B., Giorgetti, L., and Chao, J.A. (2017). The Dynamics of mRNA Turnover Revealed by Single-Molecule Imaging in Single Cells. Molecular Cell 68, 615–625.

Hsu, K.-S., and Kao, H.-Y. (2018). PML: Regulation and multifaceted function beyond tumor suppression. Cell & Bioscience *8*:5.

Hubstenberger, A., Courel, M., Bénard, M., Souquere, S., Ernoult-Lange, M., Chouaib, R., Yi, Z., Morlot, J.-B., Munier, A., Fradet, M., et al. (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. Molecular Cell *68*, 144–157.e145.

Lallemand-Breitenbach, V., and de The, H. (2010). PML Nuclear Bodies. Cold Spring Harbor Perspectives in Biology *2*, a000661.

Lecona, E., and Fernandez-Capetillo, O. (2016). A SUMO and ubiquitin code coordinates protein traffic at replication factories. Bioessays *38*, 1209–1217.

Lecona, E., Rodriguez-Acebes, S., Specks, J., Lopez-Contreras, A.J., Ruppen, I., Murga, M., Mu oz, J., Mendez, J., and Fernandez-Capetillo, O. (2016). USP7 is a SUMO deubiquitinase essential for DNA replication. Nat Struct Mol Biol *23*, 270–277.

Liberali, P., Snijder, B., and Pelkmans, L. (2014). A Hierarchical Map of Regulatory Genetic Interactions in Membrane Trafficking. Cell *157*, 1473–1487.

Lin, H.-K., Bergmann, S., and Pandolfi, P.P. (2004). Cytoplasmic PML function in TGF-b signalling . Nature *431*, 200–205.

Luo, Y., Na, Z., and Slavoff, S.A. (2018). P-Bodies: Composition, Properties, and Functions. Biochemistry *57*, 2424–2431.

Machyna, M., Heyn, P., and Neugebauer, K.M. (2012). Cajal bodies: where form meets function. WIREs RNA *4*, 17–34.

Mao, Y.S., Bin Zhang, and Spector, D.L. (2011). Biogenesis and function of nuclear bodies. Trends in Genetics *27*, 295–306.

Marnef, A., Weil, D., and Standart, N. (2011). RNA-related nuclear functions of human Pat1b, the P-body mRNA decay factor. Molecular Biology of the Cell *23*, 213–224.

Mazroui, R., Di Marco, S., Kaufmann, R.J., and Gallouzi, I.-E. (2007). Inhibition of the Ubiquitin-Proteasome System Induces Stress Granule Formation. Molecular Biology of the Cell *18*, 2603–2618.

McCann, K.L., and Baserga, S.J. (2014). Driving nucleolar assembly. Genes & Development

28, 211–213.

Ozgur, S., Chekulaeva, M., and Stoecklin, G. (2010). Human Pat1b Connects Deadenylation with mRNA Decapping and Controls the Assembly of Processing Bodies. Molecular and Cellular Biology *30*, 4308–4323.

Rai, A.K., Chen, J.-X., Selbach, M., and Pelkmans, L. (2018). Kinase-controlled phase transition of membraneless organelles in mitosis. Nature *559*, 211-216.

Ramachandran, V., Shah, K.H., and Herman, P.K. (2011). The cAMP-Dependent Protein Kinase Signaling Pathway Is a Key Regulator of P Body Foci Formation. Molecular Cell *43*, 973–981.

Ramo, P., Sacher, R., Snijder, B., Begemann, B., and Pelkmans, L. (2009). CellClassifier: supervised learning of cellular phenotypes. Bioinformatics *25*, 3028–3030.

Rinaldo, C., Siepi, F., Prodosmo, A., and Soddu, S. (2008). HIPKs: Jack of all trades in basic nuclear activities. Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research *1783*, 2124–2129.

Sacco-Bubulya, P. and Spector, D.L. (2002). Disassembly of interchromatin granule clusters alters the coordination of transcription and pre-mRNA splicing. The Journal of Cell Biology *156*, 425–436.

Saunders, T.E., Pan, K.Z., Angel, A., Guan, Y., Shah, J.V., Howard, M., and Chang, F. (2012). Noise Reduction in the Intracellular Pom1p Gradient by a Dynamic Clustering Mechanism. Developmental Cell *22*, 558–572.

Shav-Tal, Y., PattonJ, and Singer, R.H. (2005). Dynamic Sorting of Nuclear Components into Distinct Nucleolar Caps during Transcriptional Inhibition. Molecular Biology of the Cell *16*, 2395–2413.

Sheth, U., and Parker, R. (2003). Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies. Science *300*, 805–808.

Sirri, V., Urcuqui-Inchima, S., Roussel, P., and Hernandez-Verdun, D. (2007). Nucleolus: the fascinating nuclear body. Histochem Cell Biol *129*, 13–31.

Snijder, B., Liberali, P., Frechin, M., Stoeger, T., and Pelkmans, L. (2013). Predicting functional gene interactions with the hierarchical interaction score. Nat Meth *10*, 1089–1092.

Spector, D.L., and Lamond, A.I. (2011). Nuclear Speckles. Cold Spring Harbor Perspectives in Biology *3*, a000646.

Stoeger, T., Battich, N., and Pelkmans, L. (2016). Passive Noise Filtering by Cellular Compartmentalization. Cell *164*, 1151–1161.

Stoeger, T., Battich, N., Herrmann, M.D., Yakimovich, Y., and Pelkmans, L. (2015). Computer vision for image-based transcriptomics. Methods *85*, 44–53.

Urban, V., Dobrovolna, J., and Janscak, P. (2017). Distinct functions of human RecQ helicases during DNA replication. Biophysical Chemistry *225*, 20–26.

Van Damme, E., Laukens, K., Dang, T.H., and Van Ostade, X. (2010). A manually curated network of the PML nuclear body interactome reveals an important role for PML-NBs in SUMOylation dynamics. International Journal of Biological Sciences *6*, 51–67.

van der Laden, J., Soppa, U., and Becker, W. (2015). Effect of tyrosine autophosphorylation on catalytic activity and subcellular localisation of homeodomain-interacting protein kinases (HIPK). Cell Communication and Signaling *13:3*, 1–11.

van der Maaten, L., and Hinton, G. (2008). Visualizing Data using t-SNE. Journal of Machine Learning Research *9*, 2579–2605.

Van Gassen, S., Callebaut, B., Van Helden, M.J., Lambrecht, B.N., Demeester, P., Dhaene, T., and Saeys, Y. (2015). FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. Cytometry *87*, 636–645.

Wang, J.T., Smith, J., Chen, B.-C., Schmidt, H., Rasoloson, D., Paix, A., Lambrus, B.G., Calidas, D., Betzig, E., and Seydoux, G. (2014). Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in C. elegans. eLife *3*, e04591.

Winter, M., Sombroek, D., Dauth, I., Moehlenbrink, J., Scheuermann, K., Crone, J., and Hofmann, T.G. (2008). Control of HIPK2 stability by ubiquitin ligase Siah-1 and checkpoint kinases ATM and ATR. Nature Cell Biology *10*, 812–824.

Wippich, F., Bodenmiller, B., Trajkovska, M.G., Wanka, S., Aebersold, R., and Pelkmans, L. (2013). Dual Specificity Kinase DYRK3 Couples Stress Granule Condensation/Dissolution to mTORC1 Signaling. Cell *152*, 791–805.

Yang, W.-H., Yu, J.H., Gulick, T., Bloch, K.D., and Bloch, D.B. (2006). RNA-associated protein 55 (RAP55) localizes to mRNA processing bodies and stress granules. RNA *12*, 547–554.

Yates, K.E., Korbel, G.A., Shtutman, M., Roninson, I.B., and DiMaio, D. (2008). Repression of the SUMO-specific protease Senp1 induces p53-dependent premature senescence in normal human fibroblasts. Aging Cell *7*, 609–621.

Youn, J.-Y., Dunham, W.H., Hong, S.J., Knight, J.D.R., Bashkurov, M., Chen, G.I., Bagci, H., Rathod, B., MacLeod, G., Eng, S.W.M., et al. (2018). High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. Molecular Cell *69*, 517–532.

Yu, J.H., Yang, W.-H., Gulick, T., Bloch, K.D., and Bloch, D.B. (2005). Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body. RNA *11*, 1795–1802.

Figure 1

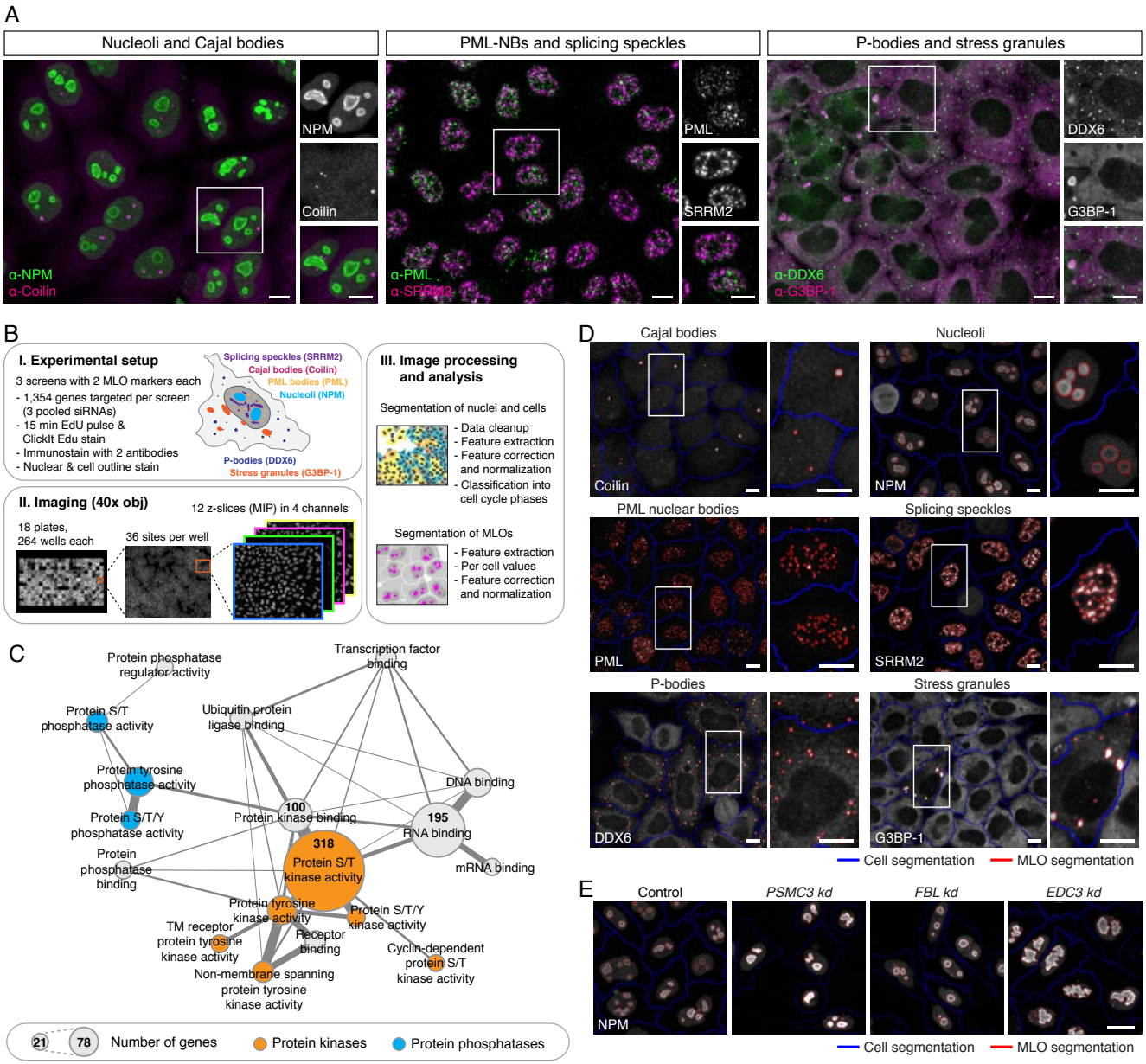
A
Nucleoli and Cajal bodies
PML-NBs and splicing speckles
P-bodies and stress granules
NPM
Coilin
α-NPM
α-Coilin
PML
SRRM2
α-PML
α-SRRM2
DDX6
G3BP-1
α-DDX6
α-G3BP-1
B
I. Experimental setup
3 screens with 2 MLO markers each
- 1,354 genes targeted per screen (3 pooled siRNAs)
- 15 min EdU pulse & ClickIt Edu stain
- Immunostain with 2 antibodies
- Nuclear & cell outline stain
Splicing speckles (SRRM2)
Cajal bodies (Coilin)
PML bodies (PML)
Nucleoli (NPM)
P-bodies (DDX6)
Stress granules (G3BP-1)
II. Imaging (40x obj)
18 plates, 264 wells each
36 sites per well
12 z-slices (MIP) in 4 channels
III. Image processing and analysis
Segmentation of nuclei and cells
- Data cleanup
- Feature extraction
- Feature correction and normalization
- Classification into cell cycle phases
Segmentation of MLOs
- Feature extraction
- Per cell values
- Feature correction and normalization
C
Protein phosphatase regulator activity
Transcription factor binding
Protein S/T phosphatase activity
Ubiquitin protein ligase binding
Protein tyrosine phosphatase activity
DNA binding
Protein S/T/Y phosphatase activity
100
Protein kinase binding
195
RNA binding
Protein phosphatase binding
318
Protein S/T kinase activity
mRNA binding
Protein tyrosine kinase activity
Protein S/T/Y kinase activity
TM receptor protein tyrosine kinase activity
Receptor binding
Cyclin-dependent protein S/T kinase activity
Non-membrane spanning protein tyrosine kinase activity
21
78
Number of genes
Protein kinases
Protein phosphatases
D
Cajal bodies
Nucleoli
Coilin
NPM
PML nuclear bodies
Splicing speckles
PML
SRRM2
P-bodies
Stress granules
DDX6
G3BP-1
Cell segmentation
MLO segmentation
E
Control
PSMC3 kd
FBL kd
EDC3 kd
NPM
Cell segmentation
MLO segmentation

Figure 2

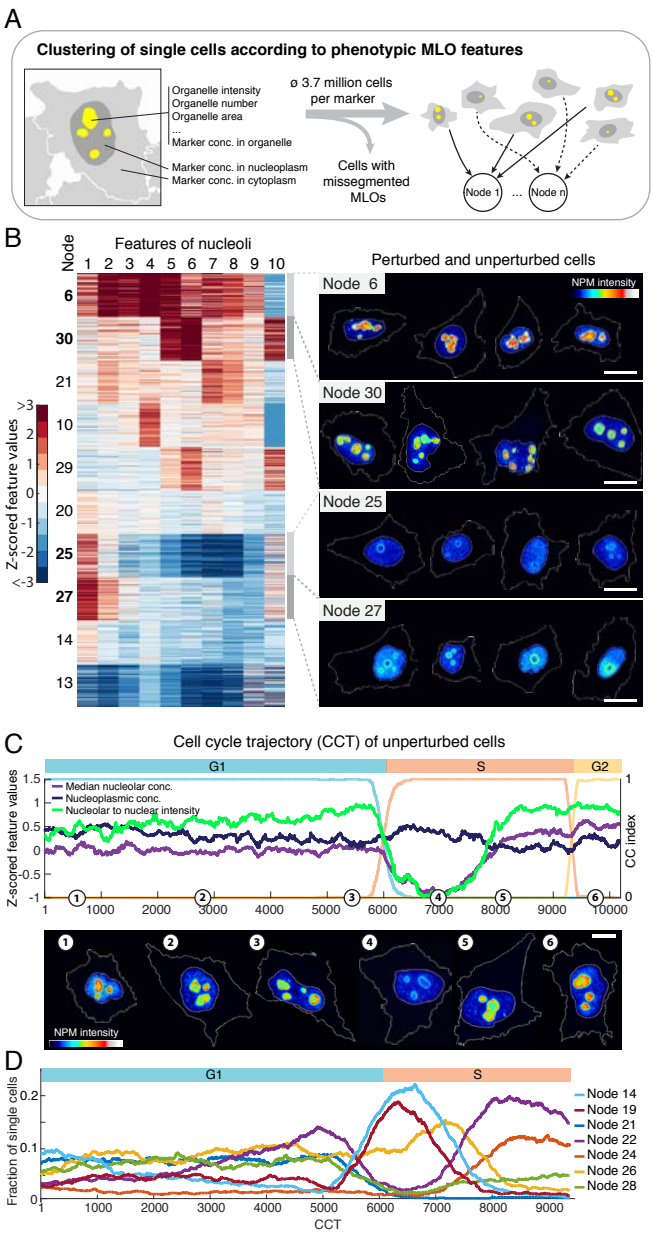
A
Clustering of single cells according to phenotypic MLO features
Organelle intensity
Organelle number
Organelle area
...
Marker conc. in organelle
Marker conc. in nucleoplasm
Marker conc. in cytoplasm
ø 3.7 million cells per marker
Cells with missegmented MLOs
Node 1
...
Node n
B
Features of nucleoli
Node
1 2 3 4 5 6 7 8 9 10
6
30
21
10
29
20
25
27
14
13
Z-scored feature values
>3
2
1
0
-1
-2
<-3
Perturbed and unperturbed cells
Node 6
NPM intensity
Node 30
Node 25
Node 27
C
Cell cycle trajectory (CCT) of unperturbed cells
G1
S
G2
Median nucleolar conc.
Nucleoplasmic conc.
Nucleolar to nuclear intensity
Z-scored feature values
CC index
1.5
1
0.5
0
-0.5
-1
1
0
1 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
1 2 3 4 5 6
NPM intensity
D
G1
S
Fraction of single cells
0.2
0.1
0
1 1000 2000 3000 4000 5000 6000 7000 8000 9000
CCT
Node 14
Node 19
Node 21
Node 22
Node 24
Node 26
Node 28

Figure 3

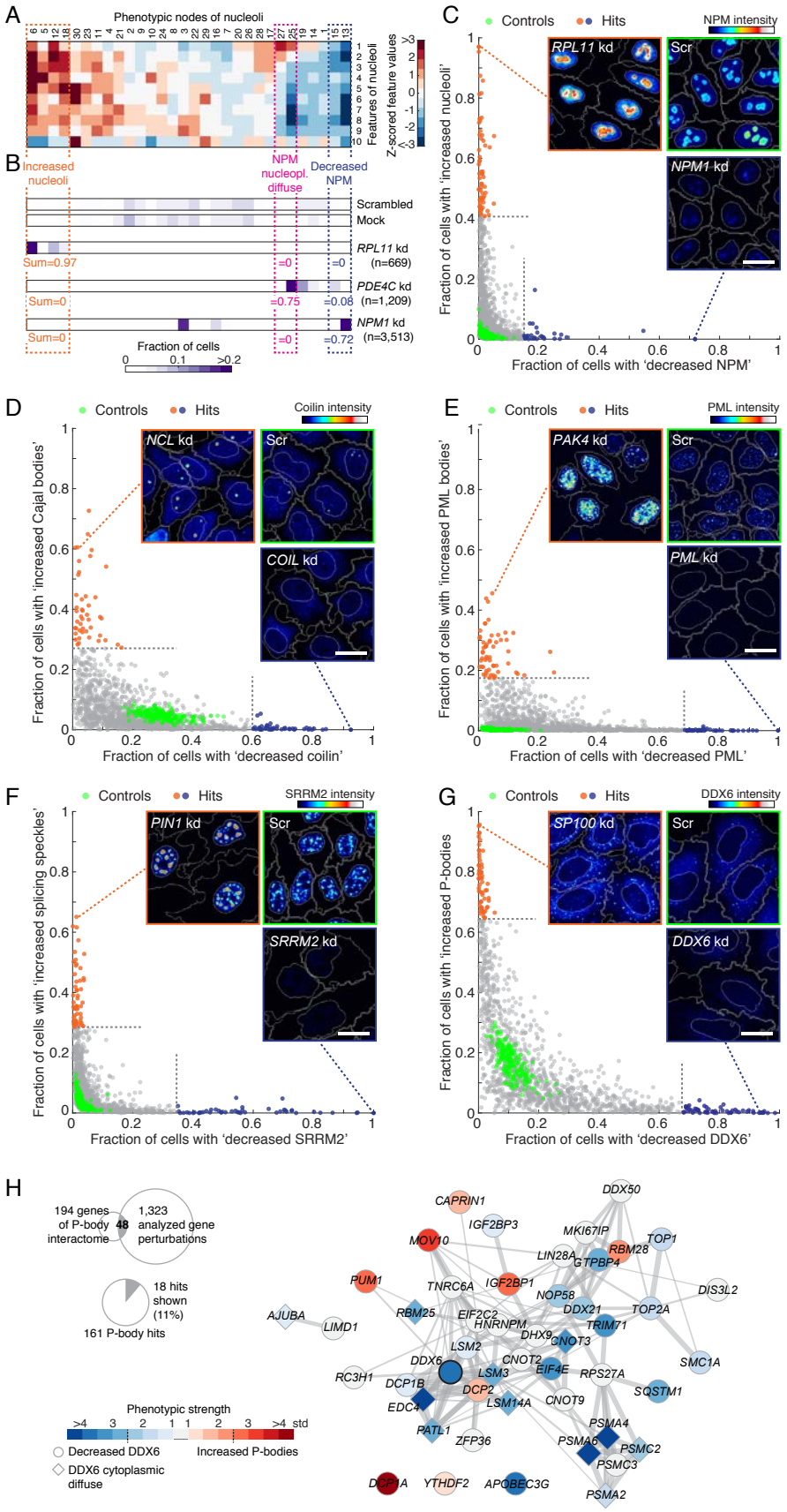
A
Phenotypic nodes of nucleoli
Features of nucleoli
Z-scored feature values
B
Increased nucleoli
NPM nucleopl. diffuse
Decreased NPM
Scrambled
Mock
RPL11 kd (n=669)
PDE4C kd (n=1,209)
NPM1 kd (n=3,513)
Sum=0.97
=0
=0
Sum=0
=0.75
=0.08
Sum=0
=0
=0.72
Fraction of cells
C
Controls
Hits
NPM intensity
RPL11 kd
Scr
NPM1 kd
Fraction of cells with 'increased nucleoli'
Fraction of cells with 'decreased NPM'
D
Controls
Hits
Coilin intensity
NCL kd
Scr
COIL kd
Fraction of cells with 'increased Cajal bodies'
Fraction of cells with 'decreased coilin'
E
Controls
Hits
PML intensity
PAK4 kd
Scr
PML kd
Fraction of cells with 'increased PML bodies'
Fraction of cells with 'decreased PML'
F
Controls
Hits
SRRM2 intensity
PIN1 kd
Scr
SRRM2 kd
Fraction of cells with 'increased splicing speckles'
Fraction of cells with 'decreased SRRM2'
G
Controls
Hits
DDX6 intensity
SP100 kd
Scr
DDX6 kd
Fraction of cells with 'increased P-bodies'
Fraction of cells with 'decreased DDX6'
H
194 genes of P-body interactome
48
1,323 analyzed gene perturbations
18 hits shown (11%)
161 P-body hits
Phenotypic strength
Decreased DDX6
Increased P-bodies
DDX6 cytoplasmic diffuse
CAPRIN1
DDX50
IGF2BP3
MKI67IP
MOV10
TOP1
LIN28A
RBM28
GTPBP4
PUM1
TNRC6A
IGF2BP1
NOP58
DIS3L2
AJUBA
LIMD1
RBM25
EIF2C2
HNRNPM
DDX21
TOP2A
TRIM71
DHX9
CNOT3
LSM2
SMC1A
DDX6
CNOT2
EIF4E
RPS27A
RC3H1
LSM3
DCP1B
DCP2
SQSTM1
EDC4
LSM14A
CNOT9
PSMA4
PATL1
ZFP36
PSMA6
PSMC2
PSMC3
DCP1A
YTHDF2
APOBEC3G
PSMA2

Figure 4

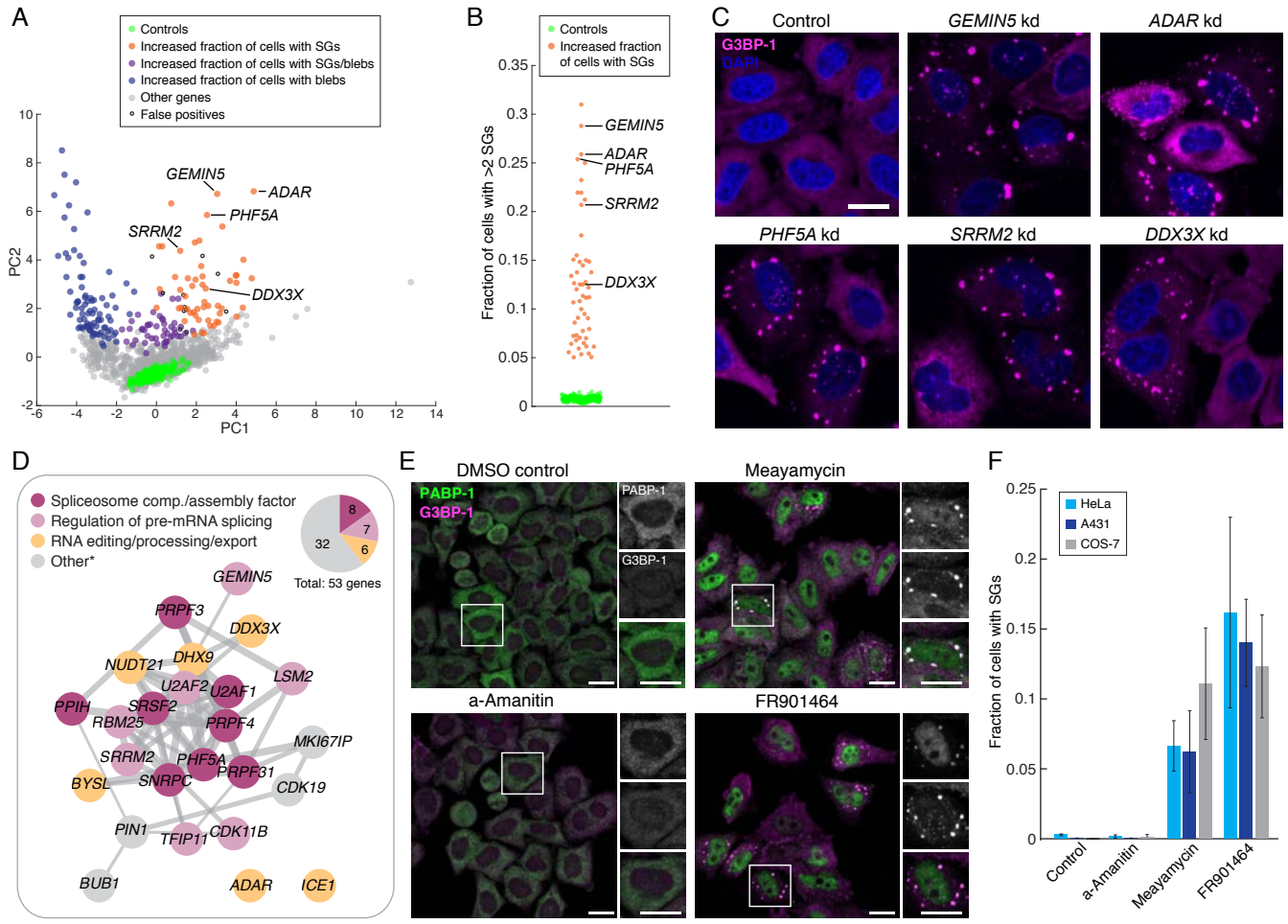
A
Controls
Increased fraction of cells with SGs
Increased fraction of cells with SGs/blebs
Increased fraction of cells with blebs
Other genes
False positives
GEMIN5
ADAR
PHF5A
SRRM2
DDX3X
PC1
PC2
B
Controls
Increased fraction of cells with SGs
Fraction of cells with >2 SGs
GEMIN5
ADAR
PHF5A
SRRM2
DDX3X
C
Control
GEMIN5 kd
ADAR kd
G3BP-1
DAPI
PHF5A kd
SRRM2 kd
DDX3X kd
D
Spliceosome comp./assembly factor
Regulation of pre-mRNA splicing
RNA editing/processing/export
Other*
8
7
6
32
Total: 53 genes
GEMIN5
PRPF3
DDX3X
NUDT21
DHX9
LSM2
U2AF2
U2AF1
PPIH
SRSF2
RBM25
PRPF4
MKI67IP
SRRM2
PHF5A
PRPF31
BYSL
SNRPC
CDK19
PIN1
TFIP11
CDK11B
BUB1
ADAR
ICE1
E
DMSO control
Meayamycin
PABP-1
G3BP-1
a-Amanitin
FR901464
F
HeLa
A431
COS-7
Fraction of cells with SGs
Control
a-Amanitin
Meayamycin
FR901464

Figure 5

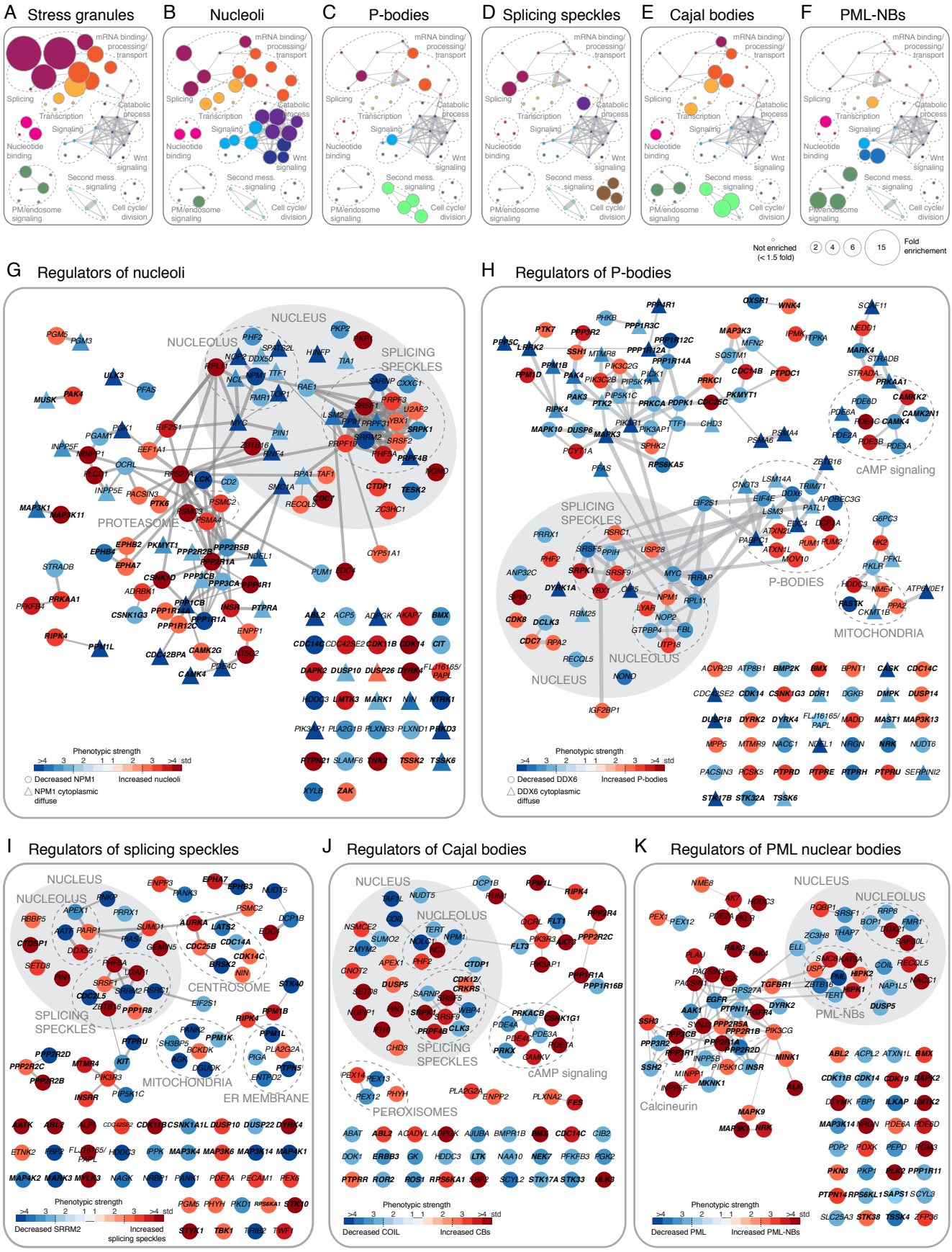
A Stress granules
B Nucleoli
C P-bodies
D Splicing speckles
E Cajal bodies
F PML-NBs
mRNA binding/ processing/ transport
Splicing
Transcription
Signaling
Catabolic process
Nucleotide binding
Wnt signaling
Second mess. signaling
PM/endosome signaling
Cell cycle/ division
Not enriched (< 1.5 fold)
2
4
6
15
Fold enrichement
G Regulators of nucleoli
NUCLEUS
NUCLEOLUS
SPLICING SPECKLES
PROTEASOME
Phenotypic strength
Decreased NPM1
Increased nucleoli
NPM1 cytoplasmic diffuse
H Regulators of P-bodies
cAMP signaling
P-BODIES
MITOCHONDRIA
Decreased DDX6
Increased P-bodies
DDX6 cytoplasmic diffuse
I Regulators of splicing speckles
CENTROSOME
MITOCHONDRIA
ER MEMBRANE
Decreased SRRM2
Increased splicing speckles
J Regulators of Cajal bodies
PEROXISOMES
Decreased COIL
Increased CBs
K Regulators of PML nuclear bodies
PML-NBs
Calcineurin
Decreased PML
Increased PML-NBs

Figure 6

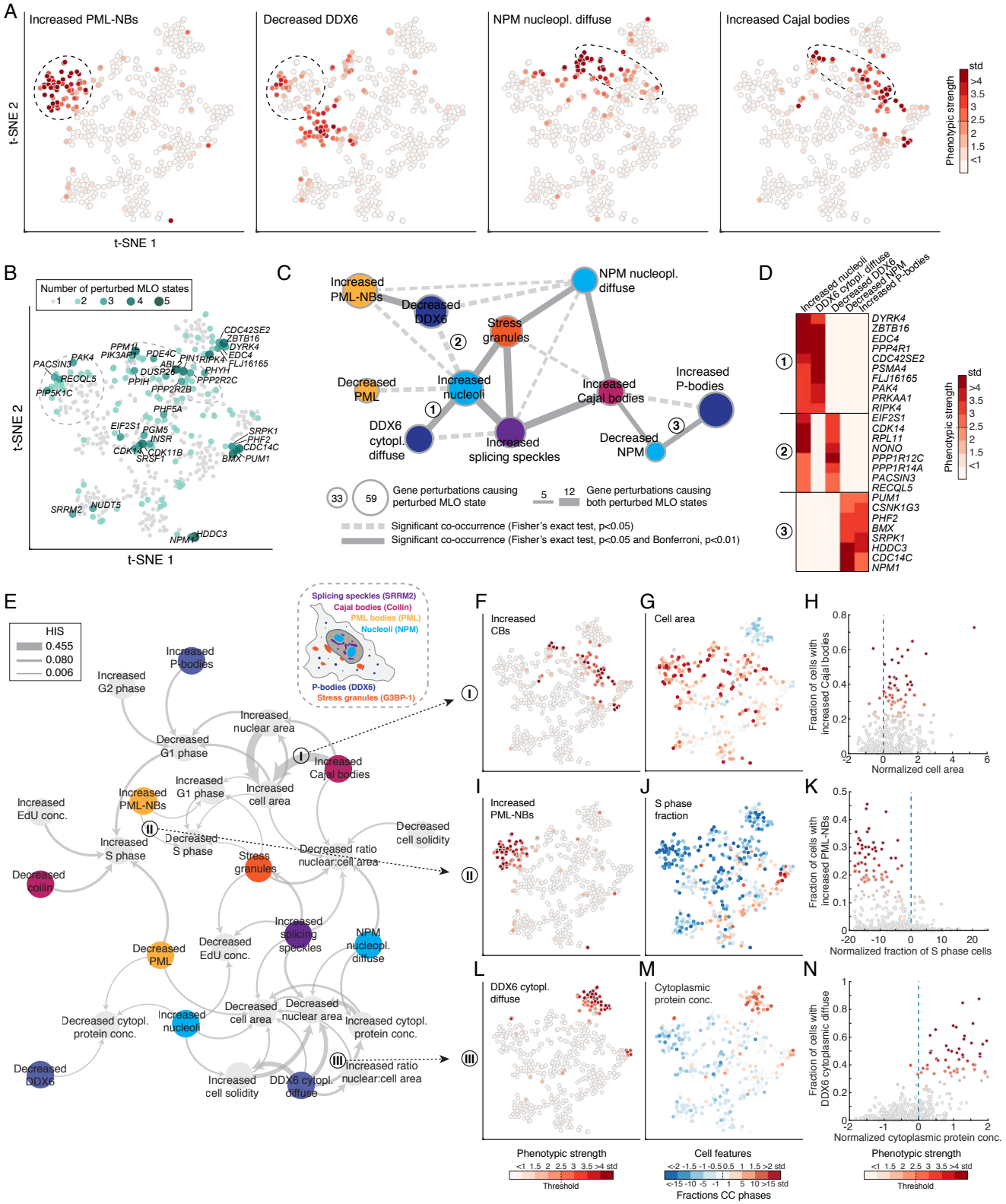
A
Increased PML-NBs
Decreased DDX6
NPM nucleopl. diffuse
Increased Cajal bodies
t-SNE 2
t-SNE 1
Phenotypic strength
std
B
Number of perturbed MLO states
C
Increased PML-NBs
NPM nucleopl. diffuse
Decreased DDX6
Stress granules
Decreased PML
Increased nucleoli
Increased Cajal bodies
Increased P-bodies
DDX6 cytopl. diffuse
Increased splicing speckles
Decreased NPM
Gene perturbations causing perturbed MLO state
Gene perturbations causing both perturbed MLO states
Significant co-occurrence (Fisher's exact test, p<0.05)
Significant co-occurrence (Fisher's exact test, p<0.05 and Bonferroni, p<0.01)
D
E
HIS
Splicing speckles (SRRM2)
Cajal bodies (Coilin)
PML bodies (PML)
Nucleoli (NPM)
P-bodies (DDX6)
Stress granules (G3BP-1)
F
Increased CBs
G
Cell area
H
Fraction of cells with increased Cajal bodies
Normalized cell area
I
Increased PML-NBs
J
S phase fraction
K
Fraction of cells with increased PML-NBs
Normalized fraction of S phase cells
L
DDX6 cytopl. diffuse
M
Cytoplasmic protein conc.
N
Fraction of cells with DDX6 cytoplasmic diffuse
Normalized cytoplasmic protein conc.
Phenotypic strength
Threshold
Cell features
Fractions CC phases

Figure 7

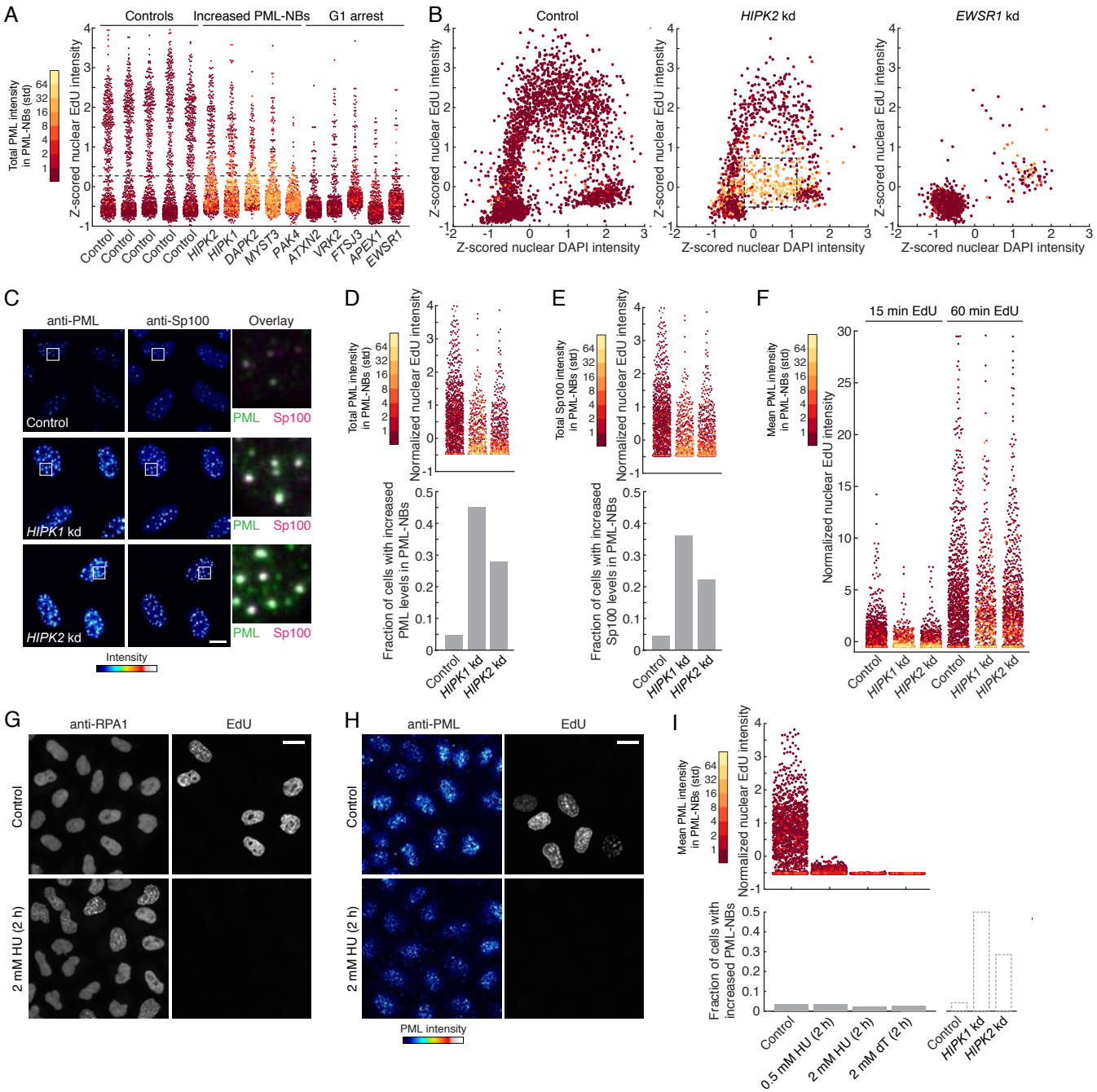
A
Controls
Increased PML-NBs
G1 arrest
Total PML intensity in PML-NBs (std)
Z-scored nuclear EdU intensity
Control
HIPK2
HIPK1
DAPK2
MYST3
PAK4
ATXN2
VRK2
FTSJ3
APEX1
EWSR1
B
Control
HIPK2 kd
EWSR1 kd
Z-scored nuclear EdU intensity
Z-scored nuclear DAPI intensity
C
anti-PML
anti-Sp100
Overlay
Control
HIPK1 kd
HIPK2 kd
PML
Sp100
Intensity
D
Total PML intensity in PML-NBs (std)
Normalized nuclear EdU intensity
Fraction of cells with increased PML levels in PML-NBs
E
Total Sp100 intensity in PML-NBs (std)
Fraction of cells with increased Sp100 levels in PML-NBs
F
15 min EdU
60 min EdU
Mean PML intensity in PML-NBs (std)
Normalized nuclear EdU intensity
G
anti-RPA1
EdU
2 mM HU (2 h)
H
anti-PML
EdU
PML intensity
I
Mean PML intensity in PML-NBs (std)
Normalized nuclear EdU intensity
Fraction of cells with increased PML-NBs
0.5 mM HU (2 h)
2 mM HU (2 h)
2 mM dT (2 h)

Figure Legends

Figure 1: Image-based RNAi screens with single organelle resolution

A) Images of unperturbed HeLa cells stained with antibodies against the indicated marker proteins of the six screened MLOs. Scale bars 20 µm.

B) Overview of the experimental set up and the computational workflow.

C) Network view of the molecular function of the 1,354 screened genes. Node sizes represent the number of genes in the library with the indicated GO-term, and nodes are connected (gray lines) when at least five genes overlap.

D) Gray scale images of unperturbed cells stained with antibodies against the indicated MLO markers. Cell segmentation is shown in blue and MLO segmentation in red. Scale bars 20 µm.

E) Gray scale images of control and perturbed cells stained with antibodies against NPM. Scale bar 20 µm.

See also Figure S1.

Figure 2: Single cell clustering accounts for cell cycle-dependent heterogeneity in MLO morphology

A) Schematic of the computational workflow per MLO marker.

B) Phenotypic nodes contain perturbed and unperturbed single cells with similar MLO features. The clustergram depicts 1,000 subsampled single cell feature profiles of ten example nodes of the nucleoli screen. Features of nucleoli are 1) NPM concentration (conc.) around nucleoli, 2) nuclear NPM conc., 3) median NPM conc. in nucleoli, 4) median NPM intensity in nucleoli, 5) total NPM intensity in nucleoli, 6) total area of nucleoli, 7) ratio area of nucleoli to nucleus, 8) ratio nucleolar to nuclear NPM intensity, 9) ratio NPM conc. in to conc. around nucleoli, and 10) number of nucleoli. False color images of four example cells that were clustered into the indicated nodes are shown on the right. Scale bars 20 µm.

C) Morphological changes of nucleoli over the cell cycle. Nucleolar features are plotted along a cell cycle trajectory (CCT) of more than 10,000 unperturbed cells (upper panel). False colored images of representative cells derived from the indicated CCT positions are shown (lower panel). Scale bar 10 µm.

D) Unperturbed cells (same as in **C**) are clustered into different nodes that reflect the morphological changes of nucleoli over the cell cycle. Fractions were calculated as moving average of 1,000 cells. Note that not all node fractions are shown.

See also Figure S2.

Figure 3: Identification of gene perturbations that alter MLO morphology

A) Clustered median feature values of the 30 phenotypic nodes of the nucleoli screen. Features 1-10 as indicated in Figure 2B.

B) Fractions of cells clustered into the 30 phenotypic nodes of the nucleoli screen (Scrambled: n=726,715, mock: n=77,824). Nodes that were combined to perturbed states are indicated, as well as the sum of the fractions of the perturbed cell populations that clustered into these nodes.

C) Gene perturbations above the threshold (2.5 standard deviations (stds) of the mean of all fractions, gray dotted lines) are considered as hits. False colored images of cells from control (green) and hit genes (orange or blue) are shown. Scale bar 20 µm.

D-G) Fractions of cells plotted as in **C** for perturbed states of **D**) Cajal bodies, **E**) PML nuclear bodies, **F**) splicing speckles, and **G**) P-bodies.

H) Network of 48 genes representing the overlap of the P-body (PB) interactome (194 genes) with the 1,323 analyzed gene perturbations. Node colors and shape indicate the effect of the respective gene perturbation on PBs. Phenotypic strength indicates the total fraction of cells with the respective perturbed MLO state in stds from the mean of all perturbations and controls.

See also Figure S3, and Tables S1 and S2.

Figure 4: Genetic and chemical perturbation of pre-mRNA splicing induces stress granule formation

A) Identification of gene perturbations with an increased fraction of cells with segmented stress granules (SGs).

B) Fraction of cells with SGs of control populations (green, n=199 wells, mean fraction=0.0077, std=0.0024) and of gene perturbations (orange, n=53).

C) Representative images of control cells and cells with the indicated gene perturbations stained for SGs (G3BP-1, magenta) and nuclei (blue). Scale bar 20 µm.

D) Protein interaction network for 25 out of 53 genes whose perturbation leads to the formation of SGs. Node colors indicate the functional annotation of the genes.

E) Images of cells treated with either DMSO or the indicated chemical compounds and stained for SGs. Scale bars 20 µm.

F) HeLa, A-431 and COS-7 cells were treated with the indicated chemical compounds and the fractions of cells that form SGs upon treatment were quantified. Bars represent the mean fraction of cells derived from three independent experiments, error bars represent the std.

See also Figure S4 and Table S1.

Figure 5: Distinct regulatory pathways control MLOs

A-F) Network visualization of functional annotation enrichments calculated for gene perturbations resulting in **A)** stress granules, or increased or decreased states of **B)** nucleoli, **C)** P-bodies, **D)** splicing speckles, **E)** Cajal bodies, and **F)** PML nuclear bodies. Nodes represent GO-terms and are colored and grouped according to functional similarity. Node

edges (gray lines) are shown if two annotations share more than 20% of genes. Node sizes represent fold enrichment.

G-K) Gene networks of the regulators of **G)** nucleoli, **H)** P-bodies, **I)** splicing speckles, **J)** Cajal bodies, and **K)** PML nuclear bodies. Node colors and shape indicate the effect of the respective gene perturbation on MLO morphology.

See also Figure S5.

Figure 6: Co-occurrence of perturbed MLO states and their functional consequences on the cellular state

A) Gene t-SNE maps of 453 gene perturbations (nodes) that are scored as hit for one or more perturbed MLO states. Color indicates the respective phenotypic strength per gene.

B) Gene t-SNE map highlighting gene perturbations with pleiotropic effects on the six screened MLOs.

C) Network visualization of significant co-occurrences between two perturbed MLO states (nodes). Node colors represent individual MLOs (see cartoon in **E**). Edge width (gray lines) indicates the number of gene perturbations with both perturbed MLO states.

D) Heat map of the phenotypic strength of gene perturbations that affect both nucleoli and P-bodies (PBs). Gene groups 1-3 refer to the indicated edges of the network in **C**.

E) Network visualization of the hierarchical interaction score (HIS) between perturbed MLO states (colored nodes, see cartoon) and features describing the cellular state (gray nodes). Edges (gray arrows) indicate the directionality and strength of the inferred interaction.

F-H) Hierarchical interaction between increased Cajal bodies (CBs) and cell area. **F)** Gene t-SNE map as in **A** showing the distribution of gene perturbations with increased CBs. **G)** Gene t-SNE map as in **A**. Color indicates the median cell area of each gene perturbation normalized to controls. **H)** Correlation between increased CBs and cell area. Color indicates the phenotypic strength of increased CBs as in **F**. Light blue dashed line indicates the median cell area of controls.

I-K) Hierarchical interaction between increased PML nuclear bodies (PML-NBs) and fraction of cells in S phase. **I)** Gene t-SNE map as in **A** showing the distribution of gene perturbations with increased PML-NBs. **J)** Gene t-SNE map as in **A**. Color indicates the fraction of cells in S phase per gene perturbation normalized to controls. **K)** Correlation between increased PML-NBs and fraction of cells in S phase. Color indicates the phenotypic strength of increased PML-NBs as in **I**. Light blue dashed line indicates the fraction of control cells in S phase.

L-N) Hierarchical interaction between absent P-bodies (DDX6 cytoplasmic diffuse) and increased cytoplasmic protein concentration. **L)** Gene t-SNE map as in **A** showing the distribution of gene perturbations with DDX6 cytoplasmic diffuse. **M)** Gene t-SNE map as in **A**. Color indicates the median cytoplasmic protein concentration of each gene perturbation normalized to controls. **N)** Correlation between 'DDX6 cytoplasmic diffuse' and cytoplasmic protein concentration. Color indicates the phenotypic strength of 'DDX6 cytoplasmic diffuse'

as in **L**. Light blue dashed line indicates the median cytoplasmic protein concentration of controls.
See also Figure S6.

Figure 7: Increased PML nuclear bodies cause a decrease in DNA replication rates
A) EdU intensity of 800 subsampled cells of control wells (derived from five different plates), five hits for increased PML-NBs, and five perturbations that lead to G1 arrest. Color indicates total PML intensity in segmented PML-NBs as std from the mean of all control wells. Dashed line indicates approximate threshold for SVM-based classification of S phase cells. Note that the total nuclear EdU intensity also includes weak signal of the protein stain succinimidylester.
B) DAPI and EdU intensity of single cells. Color indicates total PML intensity in segmented PML-NBs as in **A**. Box (dashed black outline) highlights cells with increased PML-NBs and reduced EdU incorporation. Note that the total nuclear EdU intensity also includes weak signal of the protein stain succinimidylester.
C) Sp100 (magenta) is enriched with PML (green) in increased PML-NBs.
D) EdU intensity and total PML intensity in segmented PML-NBs of control cells and *HIPK1* or *HIPK2* depleted cells (upper panel). Lower panel shows the fraction of cells with a total PML intensity in segmented PML-NBs of more than two stds from the mean of control cells (n=1,336 to 3,086 per condition).
E) EdU intensity of the same cells as in **D** but color indicates the total Sp100 intensity in segmented PML-NBs (upper panel). Lower panel shows the fraction of cells with a total Sp100 intensity in segmented PML-NBs of more than two stds from the mean of control cells.
F) EdU intensity and mean PML intensity in segmented PML-NBs of control cells and *HIPK1* or *HIPK2* depleted cells. Cells were incubated with EdU for 15 or 60 minutes before fixation. (n=1,336 to 3,086 per condition).
G-I) Stalled replication forks do not induce an increase in PML-NBs. Representative images of unperturbed cells or cells treated with 2 mM hydroxyurea (HU) were labeled with EdU and stained with antibodies against RPA1 (**G**) or PML (**H**). Scale bars 20 µm. **I)** EdU intensity and mean PML intensity in segmented PML-NBs of control cells or cells treated with the indicated concentration of HU or deoxythymidine (dT) (upper panel). Lower panel shows the fraction of cells with a mean PML intensity in segmented PML-NBs of more than two stds from the mean of control cells (n=2,214 to 3,070 per condition). For comparison, dotted grey boxes show the fractions as calculated in **D** but for mean PML intensity in segmented PML-NBs.
See also Figure S7.

STAR Methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Lucas Pelkmans (lucas.pelkmans@imls.uzh.ch).

MODEL AND SUBJECT DETAILS

HeLa MZ cells were cultivated at 37ºC and 5% CO_2 in DMEM (ThermoFisher) supplemented with 1% GlutaMAX (ThermoFisher) and 10% FBS (Merck). A-431 and COS-7 cells were cultivated at 37ºC and 5% CO_2 in DMEM containing sodium pyruvate (PAN-Biotech) supplemented with 1% GlutaMAX (ThermoFisher) and 10% FBS (Merck).

METHOD DETAILS

Transfections

For the screens, about 900 HeLa cells were plated per well in 384-well plates (Greiner) for reverse transfection on top of a mixture of pooled siRNAs (5 nM final concentration) and RNAiMAX (0.08 µl per well in OptiMEM; ThermoFisher) according to manufacturer's specifications. Cells were subsequently grown for 72 hours at 37ºC in complete DMEM to establish efficient knock down of the targeted genes. For plasmid transfections cells were seeded in 96-well plates (Greiner) and incubated at 37ºC till cells reached ~70% confluency. Cells were transfected with 100 ng of plasmid per well using Lipofectamine 2000 (ThermoFisher) according to manufacturer's specifications and incubated for 24 hours at 37ºC.

Assays and drug treatments

For measuring DNA synthesis rates cells were incubated for 15 or 60 minutes at 37ºC in DMEM containing 200 µm 5-ethynyl-2'-deoxyuridine (see Click-iT EdU Alexa Fluor 647 Imaging Kit, ThermoFisher). For stalling DNA replication forks cells were incubated for 2 hours at 37ºC with the indicated concentrations of hydroxyurea or deoxythymidine. For chemical inhibition of pre-mRNA splicing cells were incubated with 100 ng/ml FR901464 or 10 nM Meayamycin for 20 hours at 37ºC. For inhibition of mRNA transcription cells were incubated with 10 ug/ml alpha-Amanitin for 20 hours at 37ºC.

siRNA library

The siRNA library consists of three pooled siRNAs against each of 1,354 human genes (listed in Table S1). Among the genes targeted are the human kinome and phosphatome (Plates 1-3 of the Ambion Silencer Select Drugable Genome Library, ThermoFisher) and a custom set of 300 genes (Ambion Silencer Select, ThermoFisher) encoding for phosphoproteins with GO-term annotations for any of the six screened MLOs. GO-terms and gene candidates were retrieved from QuickGO (https://www.ebi.ac.uk/QuickGO/). 199 wells with negative controls

(Silencer Select Negative Control No. 1, ThermoFisher), 19 wells with mock controls (no siRNAs) and 10 wells with positive controls (siRNA against *KIF11*) were positioned throughout the screen plates.

Cell stains

All staining and washing steps were performed on a semi-automated liquid handling platform (BioTek). Between each staining step cells were washed four times with PBS. Cell were fixed with 4% PFA in PBS for 30 min and permeabilized with 0.25% Triton X-100 in PBS for 30 min. EdU Click-iT reactions were performed before blocking according to manufacturer's specifications (Click-iT EdU Alexa Fluor 647 Imaging Kit, ThermoFisher) and as previously described (Gut et al., 2015). Cells were blocked in 5% donkey serum (Abcam) for 1 hour at room temperature (RT) and incubated over night at 4°C in primary antibodies diluted in 5% donkey serum. Cells were incubated for 90 minutes at RT in secondary antibodies diluted 1:500 in 5% donkey serum. Cells were subsequently stained for 10 minutes with DAPI (1:1,000) and for 5 minutes with succinimidyl ester (Alexa Fluor 647 NHS Ester; ThermoFisher; 1:80,000). For detection of polyA-mRNA fluorescence in situ hybridization (FISH) was performed according to manufacturer's specifications using Stellaris RNA FISH buffers (LGC Biosearch Technologies) and a custom ATTO 488-labelled 18-nucleotide long oligo-dT probe (Microsynth). Cells were fixed with 4% PFA in PBS for 10 min and permeabilized with 70% ethanol for 6 h at 4°C. Cells were washed and incubated over night at 37°C in hybridization buffer containing the probe before they were washed and processed for immunofluorescence.

Imaging

Images were acquired with an automated spinning disc microscope (CellVoyager 7000, Yokogawa) using a 40x air objective (0.95 NA, Olympus) and two Neo sCMOS cameras (Andor).

Images from two wavelengths were acquired simultaneously (405 and 568 nm, 488 and 647 nm). Twelve z-planes spaced by 1 µm were acquired per site and channel, and maximum intensity projections were saved. All images presented in the same panel of a figure are rescaled to the same brightness/contrast values.

Image processing

Two image processing pipelines were set up with CellProfiler (Carpenter et al., 2006). In a first pipeline, images were illumination corrected and camera-dependent invariant background signal was subtracted (Stoeger et al., 2015). A five-pixel y-shift between two band pass filters was corrected for 488 and 647 nm images. Nuclei and cells were segmented based on DAPI and succinimidyl ester signal intensity, respectively (Stoeger et al., 2015). Intensity, texture, area and shape features were extracted from segmented nuclei, cells and cytoplasms. In a second pipeline, segmentations of nuclei, cells, and MLOs were loaded and segmented

MLOs were related to either parent nuclei or parent cytoplasms. Intensity, texture, area and shape features were extracted from each single segmented MLO. MLOs with an area smaller than three pixels were discarded. For all segmented MLOs related to one parent nucleus or parent cytoplasm, mean, median and total values of intensity and area features were calculated. Segmented MLOs were subtracted from their parent nuclei or cytoplasms to measure the MLO marker intensity around MLOs.

Segmentation of MLOs using pixel classification

For pixel classification-based MLO segmentation a customized GUI was written in MATLAB. The code is available as Supplemental Software and can be downloaded from GitHub (https://github.com/pelkmanslab/PixelClassification). Analogous to the pixel classifier *ilastik* (http://ilastik.org/) the pixel features used for classification are based on a variety of image filters outlined in detail in the *GeneratePixelFeature.m* function of the Supplemental Software. Briefly, image filters include Gaussian filters, difference-of-Gaussian and Laplacian-of-Gaussian filters, mean and median filters, entropy filters, morphological opening and closure, as well as measurements of the local background, blind deconvolution of images, watershed lines (Stoeger et al., 2015), and top hat filters. Image filters and scales were selected prior to training according to the morphology of the six different MLOs (see Supplemental Software). For the segmentation of nucleoli, features of nuclei and cell segmentations were also included. For each MLO the training set was composed of more than 1,000 manually selected pixels from 150 to 200 different wells associated with a binary label (signal or background) and the list of feature values for each pixel. The classification model was trained using the inbuilt support vector machine functionality of MATLAB (*fitsvm.m*). In the case of ring- or donut-shaped MLOs, such as Cajal bodies or nucleoli, the segmentation algorithm was trained in such a way that resulting holes in segmented structures were filled.

Data clean up and classification of cell cycle phases

Cells that extended beyond the field of view (border cells) were discarded from all images. For the remaining cells supervised machine learning was applied to discard apoptotic, wrongly segmented and multinucleated cells (Ramo et al., 2009; Stoeger et al., 2015). To remove gene perturbations that strongly decreased cell number a mixture-model of three Gaussians was fit to the cell number distribution of all wells. The minimal cell number per well was set to the 95th percentile of the first Gaussian. We excluded 31 perturbations, among them *PLK1*, *WEE1*, *AURKB* and the positive control *KIF11* (in total 2.3 percent of the gene library) that seemed to be essential for cell viability as their knock down resulted in low numbers of vital cells (<246). Supervised machine learning was applied to classify cells in S-phase based on intensity and texture of the 5-ethynyl-2'-deoxyuridine (EdU) Click-iT stain (Gut et al., 2015). Support vector machines were trained to classify mitotic (pro, meta, ana, and telo phase) and early G1 cells to exclude them from this analysis. To assign the remaining cells to either G1 or G2 phase, histograms of the DNA content (total nuclear DAPI

intensity) were plotted for all interphase cells of one 384-well plate. Cells classified as S-phase were subtracted and two Gaussians were fitted using the *gmdistribution.fit* function. The minimal point between the two means was used as split point and populations on either side of the split point were classified as G1 and G2, respectively. Subsequently, for each well a Gaussian was fit on either histogram of G1 and G2 cell populations using the *sgm.fit_distribution* function and outliers were discarded.

Correction of single cell intensity-based features

To correct single cell intensity and texture measurements for positional staining biases within each plate a well correction factor was derived as follows: Correction factor $f_{i,j} = median(row_i) + median(column_j) - 2 \times median(plate)$, where $median(row_i)$ is the median value for the given feature of all single cells in row_i, likewise $median(column_j)$ is the median value for the given feature of all single cells in $column_j$ and $median(plate)$ is the median of all single cell feature values of the plate. The corrected single cell value $C_{i,j,k}$ was given by $C_{i,j,k} = f_{i,j} + O_{i,j,k}$, where $O_{i,j,k}$ is the original value of the given feature for cell k of well(row_i, $column_j$). To normalize feature values between different plates we computed $N_{i,j,k} = (C_{i,j,k} - median(C_{plate})) / mad(C_{plate})$, where $median(C_{plate})$ is the median of all corrected cell values of a given plate and $mad(C_{plate})$ is the median average deviation of all corrected cell values of the plate.

Single cell clustering in multivariate MLO feature space

In addition to the derived single cell MLO features from CellProfiler (see Image Processing), such as 'Total marker intensity in MLOs', or 'Median marker concentration in MLO' (concentration = total intensity divided by area), the following ratios were calculated: MLO area to total area of parent nucleus or cytoplasm, MLO marker intensity in segmented MLOs to marker intensity in parent nucleus or cytoplasm, and marker intensity in MLOs to marker intensity around MLOs within the parent nucleus or cytoplasm. These ratios are particularly important to avoid bias in hit detection due to size scaling effects. For each MLO marker only features that displayed sufficient variability over all single cells were included in the analysis. The features are indicated in Figures 3A, S3A-S3D, S4A, and Table S2. For instance, PML intensity around segmented PML nuclear bodies was close to background in all single cells, thus this feature was excluded. This selection resulted in 8-12 features per MLO. Unperturbed and perturbed conditions were analyzed together per MLO and about ~3.7 million single cells per screen were clustered in a multivariate feature space using the FlowSOM algorithm in R (Van Gassen et al., 2015). FlowSOM was used with the Euclidean distance function. First, self-organizing maps (SOMs) were built with a high node number (e.g. 100 nodes) and the result was visualized as minimal spanning tree to identify single nodes with large distances to the majority of nodes. These nodes contained cells with wrongly segmented MLOs, as confirmed by visual image inspection, and all cells assigned to these nodes were removed from the dataset. The procedure was repeated until all cells with wrongly segmented MLOs were excluded. For the final clustering the node number was reduced in such a way that cells

within one node had highly similar MLO morphological features while MLO features of cells between two nodes were noticeably different. For splicing speckles, PML nuclear bodies and P-bodies the node number was set to 15 as this node number was sufficient to represent all possible phenotypes of these MLOs. For Cajal bodies and nucleoli 30 nodes were used since these MLOs have a higher phenotypic variability.

Definition of perturbed MLO states

To identify gene perturbations that affected MLO formation we aimed to analyze two states of each MLO: an increased state which includes a high intensity of the MLO marker within the segmented MLO and/or increased MLO area, and a decreased state where the MLO is either smaller or absent, and/or the marker intensity is markedly decreased within detectable MLOs. In many cases, we combined multiple phenotypic nodes with the strongest or weakest MLO feature values, respectively, into one perturbed MLO state, e.g. 'Increased nucleoli'. The decision which nodes to combine was made on the basis of the median feature strength of the nodes, the absence of control cells in these nodes, and a visual inspection of single cells that clustered into these nodes. The decreased MLO state, meaning the absence or size reduction of a MLO can occur in two scenarios: 1) when the marker protein is down regulated to a level where no MLOs are detectable anymore, and 2) when the marker protein is present yet no MLOs form (diffuse marker localization). We observed down regulation and concomitant absence of MLOs for all five MLOs. However, only for P-bodies and nucleoli we observed that the marker proteins were diffuse in the cyto- or nucleoplasm. We never observed cells with nucleoplasmic diffuse SRRM2, PML or coilin. Thus, only P-bodies and nucleoli have a third MLO state termed 'DDX6 cytoplasmic diffuse' and 'NPM nucleoplasmic diffuse', respectively.

Identification of gene perturbations that induce stress granule formation

Single stress granules were segmented by pixel classification and SG features were extracted as previously described for the other MLOs. To separate cells with segmented SGs from cells with falsely segmented cytoplasmic blebs, we first calculated the fraction of cells per well that had at least two segmented objects (SGs or blebs). From these cells, we calculated the following five features per well: The mean number of segmented objects, the median of the total area of segmented objects per cell, the median of the total G3BP-1 intensity in segmented objects per cell, and the median of the G3BP-1 concentration around segmented objects per cell. The latter is an important feature to distinguish segmented SGs from segmented blebs since the G3BP-1 signal around blebs is high while it is very low around SGs. Perturbations with at least 5 percent of cells with at least two segmented objects were hierarchically clustered (Euclidean distance) and split into three groups based on the five features. Cells with segmented SGs show high feature values for number, area and intensity but low G3BP-1 concentration around segmented objects. Blebby cells show low feature values for number, area, and intensity but have high G3BP-1 concentration around

segmented objects. The third group consists of cell populations that show both SGs and blebs. To plot all wells we performed a principal component (PC) analysis on the same five features but included all controls and perturbations. The first two PCs are plotted and colors indicate the three classes as defined from the clustergram. The group of gene perturbations with high fractions of SGs was subsequently confirmed by visual inspection of the images and false positives were excluded.

Cell cycle trajectories

Cell cycle trajectories (CCTs) were constructed with Cycler (Gut et al., 2015). In brief, for each CCT around 10,000 unperturbed cells derived from three adjacent wells were pooled and sorted by Cycler in a multivariate feature space consisting of single cell measurements of DNA (DAPI) content, EdU content, DNA replication pattern (texture), and nuclear area corrected for local cell crowding. Classification of cells into discrete cell cycle phases (G1, S, G2) was achieved independently (see above) and was used to identify the cell-cycle phase transition points along the CCT. Phenotypic features of cells or MLOs were plotted along the constructed CCT as moving average of 200 cells unless otherwise stated.

Functional Enrichment Analysis

Functional enrichment analysis was performed with DAVID version 6.8. For each MLO, gene perturbations that were classified as hit for any of the perturbed MLO states (increased MLO, decreased marker, or diffuse staining) were pooled and analyzed for GO-term enrichments against the background of 1,323 genes in the siRNA library. From the functional annotation clustering only GO-terms or KEGG pathways with a fold enrichment higher than 1.5 for at least 4 genes were considered. The network of GO-terms was created with Cytoscape 3.5.1.

Two-dimensional maps of perturbed MLO states and cellular features

For the two-dimensional visualization of perturbed MLO states 453 gene perturbations that were classified as hit for at least one of the following 13 perturbed MLO states were considered: 1) Increased nucleoli, 2) NPM nucleoplasmic diffuse, 3) decreased NPM, 4) increased Cajal bodies, 5) decreased coilin, 6) increased PML nuclear bodies, 7) decreased PML, 8) increased splicing speckles, 9) decreased SRRM2, 10) increased P-bodies, 11) decreased DDX6, 12) DDX6 cytoplasmic diffuse, and 13) formation of stress granules. For these 453 genes we combined the phenotypic scores (sum of the fraction of single cells per well that were assigned to the nodes that reflect the perturbed MLO state) of the first twelve perturbed MLO states, excluding the Boolean classification for stress granule formation. After z-score normalization of the values we applied t-Distributed Stochastic Neighbor Embedding (t-SNE) dimensionality reduction (van der Maaten and Hinton, 2008) using the MATLAB *drtoolbox* and visualized the compiled dataset as a two-dimensional map. Features describing the cellular state were not used to create the two-dimensional gene map but were plotted on the generated map afterwards. Cell features were calculated as median per well but only for

the subpopulation of single cells that were assigned to nodes, which reflect one of the aforementioned perturbed MLO states. Likewise, fractions of cells in G1, S, or G2 phase were calculated for these subpopulations. In the case of genes with pleiotropic effects, cell features were calculated for each subpopulation (single cells with either perturbed MLO state) separately and values were then averaged. The resulting cell features were z-score normalized to the mean values of unperturbed cell populations, which were calculated from 199 controls wells (scrambled) in each of the three parallel screens.

Calculation of hierarchical interaction scores (HIS)

Hierarchical interaction scores (HIS) (Snijder et al., 2013) were calculated between 13 perturbed MLO states and 17 features describing the cellular state derived from 1,326 gene perturbations and 218 controls. Feature values describing cell morphology, such as cell area, shape, or protein content, were calculated as mean per well. From the cell cycle classifications of single cells fractions of cells per well in G1, S, or G2 phase were calculated. Each of the 17 features describing the cellular state was calculated separately for each of the three screens and the resulting values were then averaged. After z-score normalization values were split at zero and positive values were used as score for 'increased' states (e.g. increased cell area), while negative values were used as score for 'decreased' states. Perturbed MLO states were scored using the z-scored sum of fraction of single cells per well that was assigned to the respective phenotypic nodes. Instead of the Boolean classification for stress granule formation, the fraction of stress granule-positive cells per well of positive scored genes was used and normalized to the minimal and maximal threshold of the HIS. HIS scores were calculated with the following settings: The minimal threshold (intMin) was set to 1.5, the maximal threshold (intMax) was set to 5, and the step number was set to 1,000 for single tail (one directional) distributions. 50 out of 435 possible pairwise combinations (functional interactions) between the 30 MLO or cell features had a HIS greater than zero and were visualized as network using Cytoscape 3.5.1.

QUANTIFICATION AND STATISTICAL ANALYSIS

All data analysis was done using MATLAB unless otherwise stated. If statistical values were calculated they are indicated in the Figure Legends. N represents the number of single cells unless otherwise stated and all exact values of n are indicated in the figures or figure legends. Single cells were excluded as outlined in detail in STAR Methods.

DATA AND SOFTWARE AVAILABILITY

The raw images of the figures have been deposited in Mendeley Data (http://dx.doi.org/10.17632/h8byr7w3sx.1). The full image data set (>680,000 images, approx. 3.5 TB) is available upon reasonable request. The MATLAB code required for the pixel classification-based MLO segmentation can be downloaded from GitHub

(https://github.com/pelkmanslab/PixelClassification). Additional information can be found in the ‘README’ file.

Supplemental Information

Table S1: Fractions of cells with perturbed MLO states. Related to Figure 3 and 4

This table contains the fractions of single cells per perturbation that were clustered into the respective phenotypic nodes, which reflect one out of twelve perturbed MLO states. Fractions of gene perturbations that were identified as hits for the respective perturbed MLO state are highlighted (bold). For gene perturbations that were identified as hits for stress granule (SGs) formation the fraction of cells with SGs is indicated.

Table S2: Fractions of cells clustered into phenotypic nodes. Related to Figure 3

This table contains the fractions of single cells per perturbation that were clustered into any phenotypic node of the nucleoli, Cajal body, P-body, splicing speckle, and PML nuclear body screen, respectively. MLO features and their median values per phenotypic node are indicated.

Figure S1

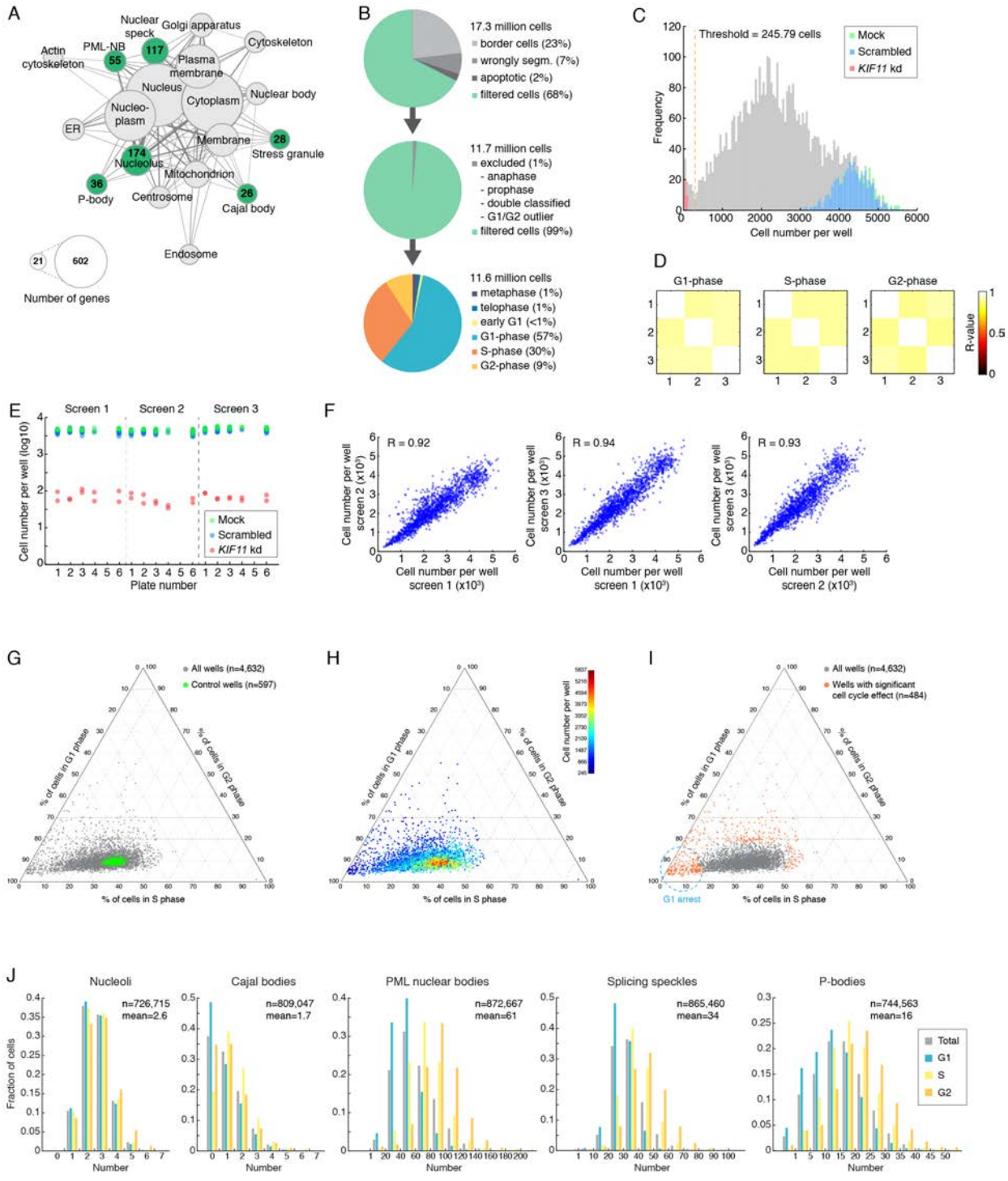

Figure S1: Quality controls of the siRNA screens. Related to Figure 1.

A) Network view of the subcellular localization of the 1,354 screened genes. Node sizes represent the number of genes in the library with the indicated GO-term for cellular component, and nodes are connected (gray lines) when at least five genes overlap. Nodes that represent one of the six screened MLOs are colored green. **B)** Diagram indicating the total numbers of segmented cells in all screens and the percentages of excluded cells due to various quality controls (for details see STAR Methods). **C)** Distribution of cell number per well. Only cells that passed the quality controls are used for further analysis. Control wells with unperturbed cells are highlighted in green (mock, n=57 wells) and blue

(Scrambled siRNA, n=597 wells). Positive controls are highlighted in red (*KIF11* kd, n=30 wells). 31 perturbations (including *KIF11*) were excluded from further analysis since knock down of these genes resulted in low numbers of vital cells (n<246, orange dashed line). **D)** Heat map showing the correlations of cell cycle classifications between the three screens. The R-value was calculated on the fraction of cells per well classified as G1, S, and G2, respectively. **E)** Cell numbers of mock (green), negative (Scrambled, blue) and positive controls (*KIF11*, red) are plotted for each plate of the three siRNA screens. Note that plate 5, 11 and 17 do not contain control wells. **F)** Correlations of cell numbers per well between siRNA screens each. **G)** Fractions of cells in G1, S or G2 phase were calculated per well (n=4,632), and normalized so that their sum equals one. Control wells (scrambled, n=597) are highlighted in green. **H)** Normalized fractions of cells in G1, S or G2 phase per well are plotted as in **S1G**. The total cell number per well is indicated in color. **I)** Normalized fractions of cells in G1, S or G2 phase per well are plotted as in **S1G**. Gene perturbations with significant alterations in cell cycle fractions are shown in orange and were calculated as follows: The mean and standard deviation (std) of G1, S and G2 fractions were calculated from all scrambled control wells (n=597; meanG1=0.59, meanS=0.28, meanG2=0.1; stdG1=0.098, stdS=0.086, stdG2=0.04). Perturbations with any fraction more than 2 std away from the mean of controls are considered significantly (alpha=0.05) different. Gene perturbations that lead to G1 arrest are encircled (light blue dashed line). **J)** Frequency plots of MLO numbers derived from more than 700,000 unperturbed cells. Segmented nucleoli, Cajal bodies, PML nuclear bodies, splicing speckles and P-bodies were counted per nucleus and cytoplasm, respectively, and numbers are plotted as fractions. For the latter three MLOs, numbers were binned and bin edges are indicated. The frequency of MLO numbers is shown for all interphase cells (gray bars) and for cells classified as G1, S or G2, respectively. Total number of cells and average MLO number are indicated.

Figure S2

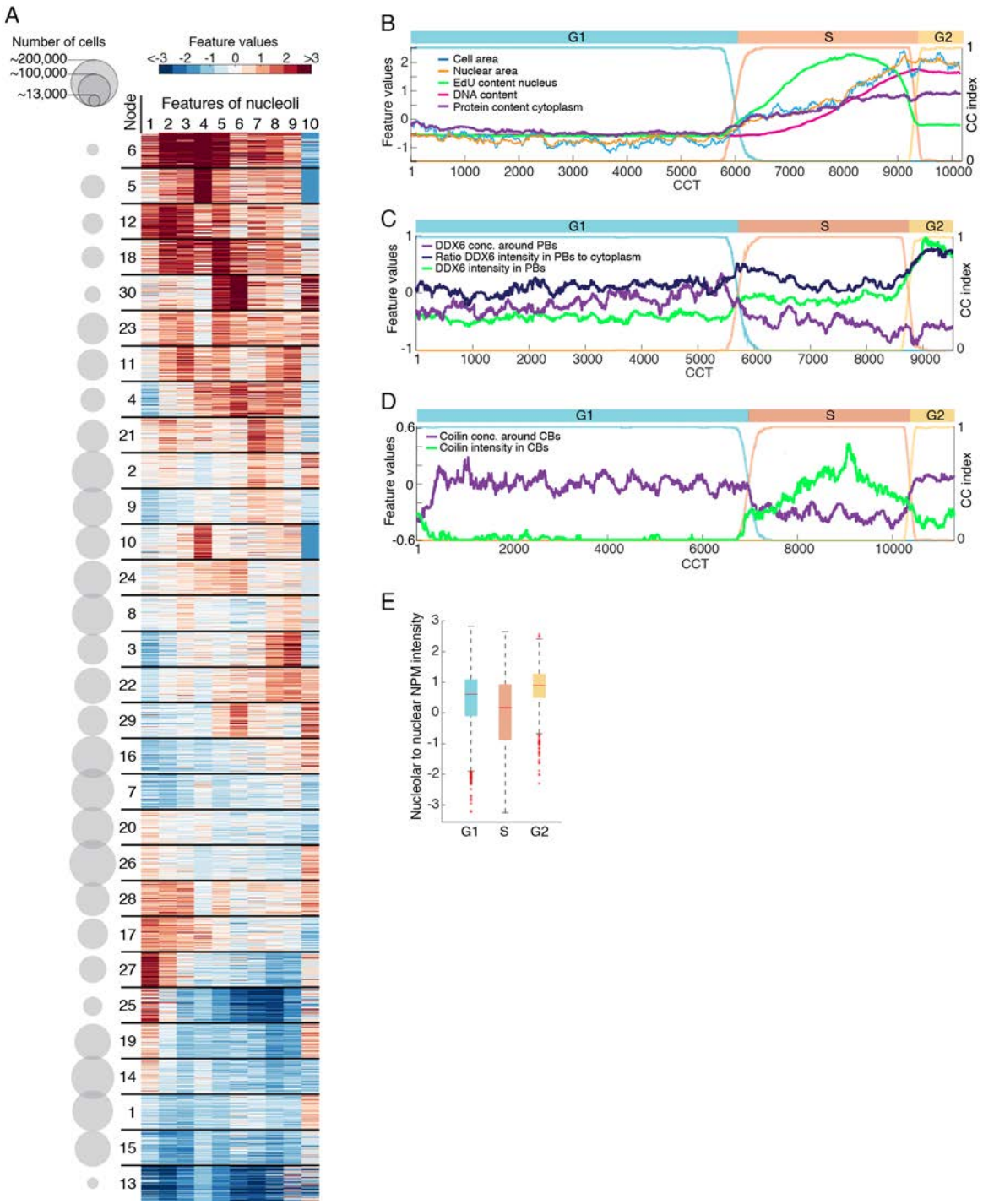

Figure S2: Single cell clustering accounts for cell cycle-dependent heterogeneity in MLO morphology. Related to Figure 2.

A) Clustergram of 1,000 subsampled single cell profiles of all 30 phenotypic nodes of the nucleoli screen. Unperturbed and perturbed cells were sorted together into the 30 nodes based on 10 z-scored feature values: 1) NPM concentration (conc.) around nucleoli, 2) Nuclear NPM conc., 3) Median NPM conc. in nucleoli, 4) Median NPM intensity in nucleoli, 5) Total NPM intensity in nucleoli, 6) Total area of nucleoli, 7) Ratio area of nucleoli to nucleus, 8) Ratio nucleolar to nuclear NPM intensity, 9) Ratio NPM conc. in nucleoli to conc. around, and 10) Number of nucleoli. Gray circles indicate the total number of single cells

clustered into each node. **B)** Morphological changes of cells over the cell cycle. Cellular features are plotted along a cell cycle trajectory (CCT) of more than 10,000 unperturbed cells. **C)** Morphological changes of P-bodies (PBs) over the cell cycle. PB features are plotted along a CCT of more than 9,000 unperturbed cells. **D)** Cajal body features are plotted along a CCT of more than 10,000 unperturbed cells. **E)** Binning of unperturbed cells into three discrete cell cycle phases does not reveal the morphological changes of nucleoli during early S-phase. Boxplots show the same data as plotted in Figure **2C** (green line, 'Nucleolar to nuclear NPM intensity') but binned into G1, S, and G2 cells, respectively.

Figure S3

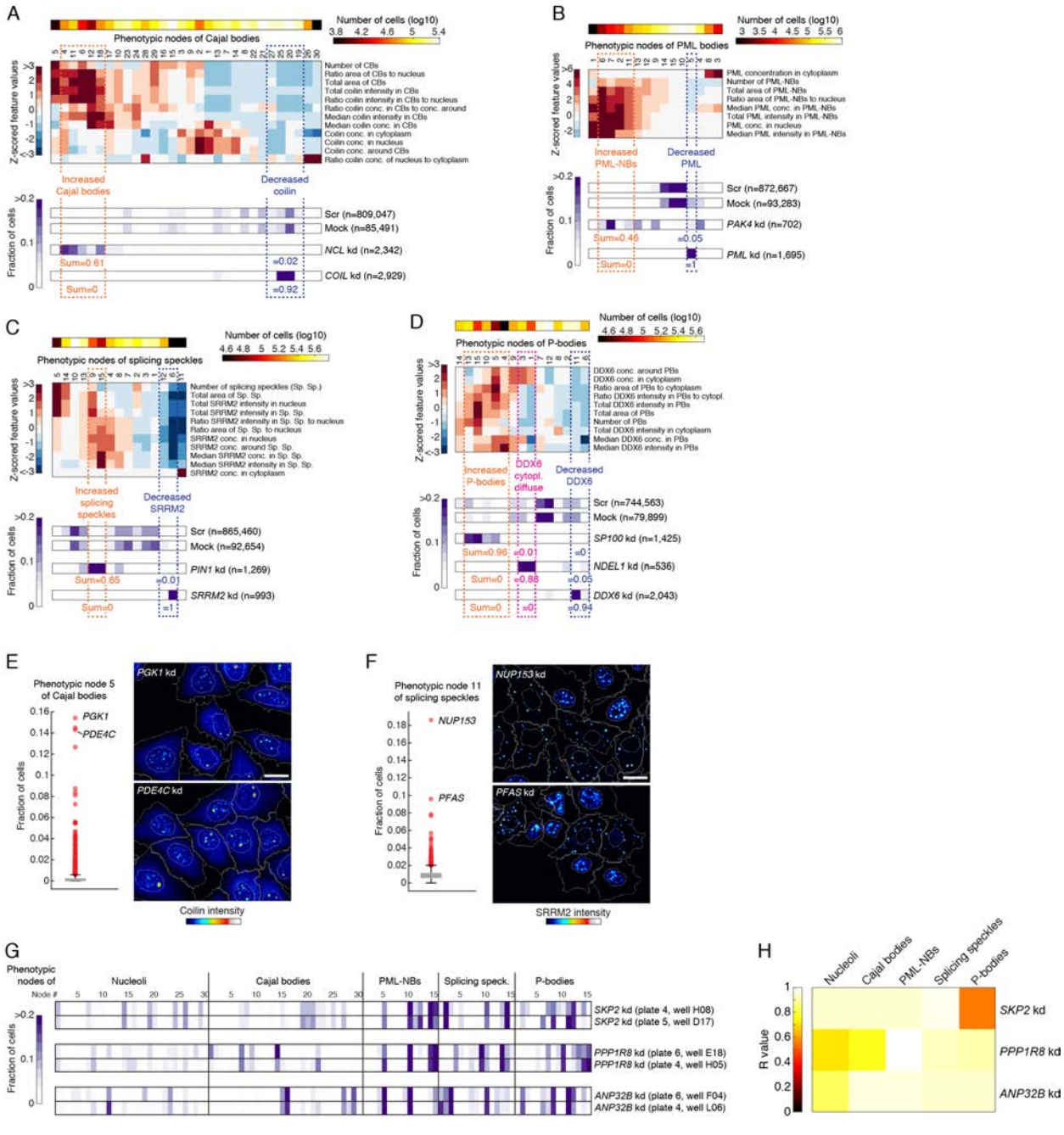

Figure S3: Identification of gene perturbations with perturbed MLO states. Related to Figure 3.

A) Clustered median feature values of the 30 phenotypic nodes of the Cajal body (CB) screen (middle panel). The top panel indicates the total number of unperturbed and perturbed cells clustered into each phenotypic node. The lower panels (purple heat maps) show the node occupancy profiles for controls (Scrambled and mock) and two gene perturbations (n indicates the number of cells). Nodes that were combined to perturbed CB states are indicated, as well as the sum of the fractions of the perturbed cell populations that clustered into these nodes. **B)** Clustered median feature values of the 15 phenotypic nodes of the PML nuclear body (PML-NBs) screen (middle panel). The top panel indicates the total number of unperturbed and perturbed cells clustered into each phenotypic node. The lower panels (purple heat maps) show the node occupancy profiles for controls (Scrambled and mock) and two gene perturbations (n indicates the number of cells). Nodes that were combined to perturbed PML-NB states

are indicated, as well as the sum of the fractions of the perturbed cell populations that clustered into these nodes. **C)** Clustered median feature values of the 15 phenotypic nodes of the splicing speckle screen (middle panel). The top panel indicates the total number of unperturbed and perturbed cells clustered into each phenotypic node. The lower panels (purple heat maps) show the node occupancy profiles for controls (Scrambled and mock) and two gene perturbations (n indicates the number of cells). Nodes that were combined to perturbed splicing speckle states are indicated, as well as the sum of the fractions of the perturbed cell populations that clustered into these nodes. **D)** Clustered median feature values of the 15 phenotypic nodes of the P-body (PB) screen (middle panel). The top panel indicates the total number of unperturbed and perturbed cells clustered into each phenotypic node. The lower panels (purple heat maps) show the node occupancy profiles for controls (Scrambled and mock) and three gene perturbations (n indicates the number of cells). Nodes that were combined to perturbed PB states are indicated, as well as the sum of the fractions of the perturbed cell populations that clustered into these nodes. **E)** Gene perturbations ranked by the highest fraction of cells clustered into phenotypic node 5 of the CB screen. Two gene perturbations are highlighted and representative false colored images are shown on the right. Cells were stained with antibodies against coilin and cell and nuclear segmentation is shown in white. Scale bar 20 μm. **F)** Gene perturbations ranked by the highest fraction of cells clustered into phenotypic node 11 of the splicing speckle screen. Two gene perturbations are highlighted and representative false colored images are shown on the right. Cells were stained with antibodies against SRRM2 and cell and nuclear segmentation is shown in white. Scale bar 20 μm. **G-H)** Node occupancy profiles of three gene duplicates in the screens underscore the high technical reproducibility of the analysis method. **G**) Heat maps of the fraction of cells of three duplicate gene perturbations (*SKP2*, *PPP1R8* and *ANP32B*) clustered into the phenotypic nodes of the five MLO screens. **H**) Heat map of the correlation coefficient (R-values) between the occupancy profiles of the three duplicate gene perturbations.

Figure S4

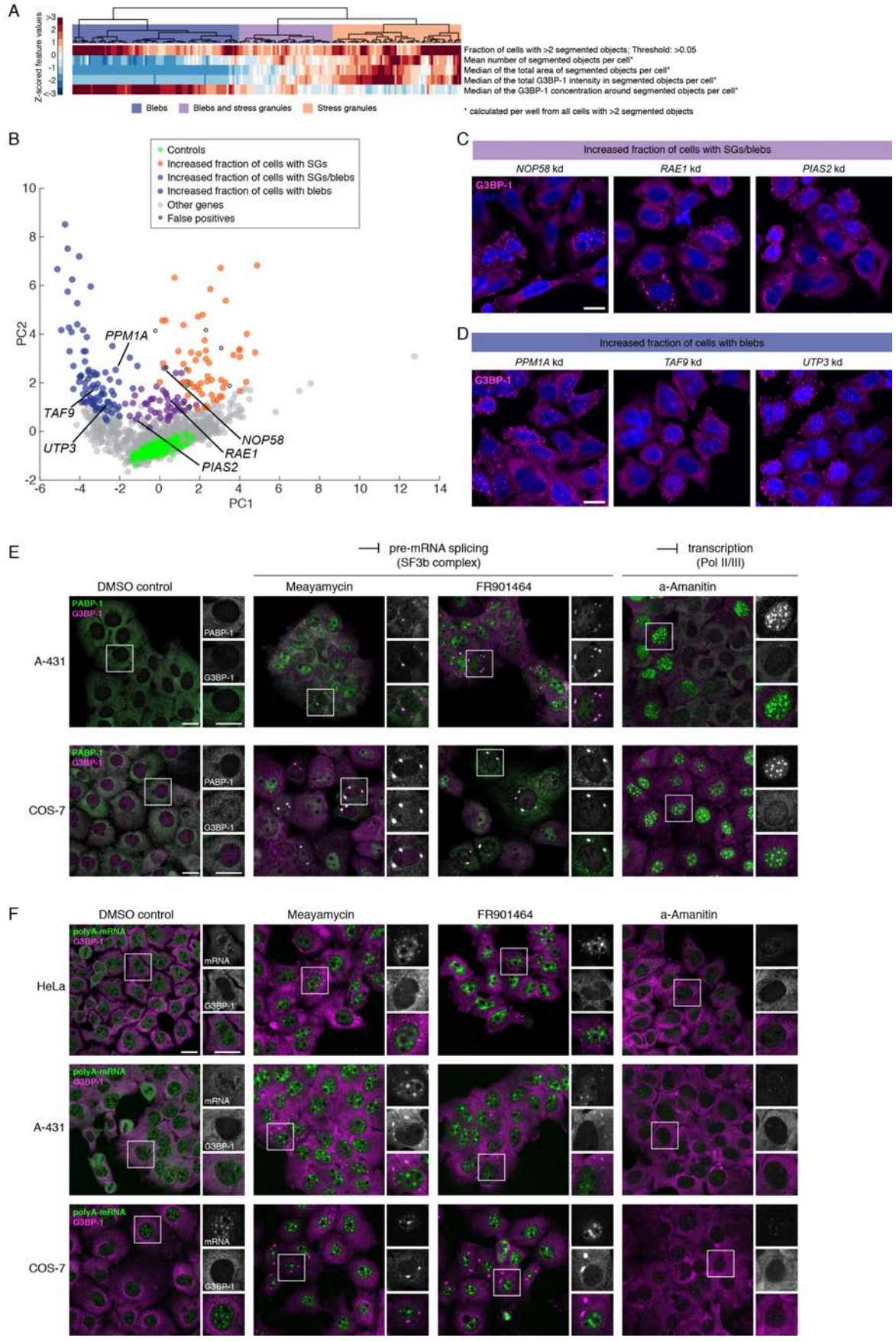

Figure S4: Genetic and chemical perturbation of pre-mRNA splicing induces stress granule formation. Related to Figure 4.

A) Clustergram of gene perturbations that have at least five percent of cells with at least two segmented objects (Stress granules (SGs) or cytoplasmic blebs). Gene perturbations were clustered based on the five indicated features per well, and grouped into three phenotypic classes: i) Perturbations with an increased fraction of cells with segmented SGs (orange), ii) perturbations with an increased fraction of cells with segmented blebs (blue), and iii) perturbations with an increased fraction of cells with both segmented SGs and blebs (purple). **B)** Principal component (PC) analysis was performed on the same features as used in **A** but for controls and perturbations. The first two PCs are plotted. Colors indicate the three classes as defined in **A**. Controls are shown in green. **C)** Representative images of cells with the indicated gene perturbations that lead to an increased fraction of cells with SGs and cytoplasmic blebs. The SG marker G3BP-1 is shown in magenta and nuclei in blue. Scale bar 20 μm. **D)** Representative images of cells with the indicated gene perturbations that lead to an increased fraction of cells with cytoplasmic blebs. The SG marker G3BP-1 is shown in magenta and nuclei in blue. Scale bar 20 μm. **E)** Images of A-431 (upper panel) and COS-7 (lower panel) cells treated with either DMSO or the indicated chemical compounds. Cells were stained with antibodies against the SG markers G3BP-1 (magenta) and PABP-1 (green). Scale bars 20 μm. **F)** Images of HeLa, A-431, and COS-7 cells treated with either DMSO or the indicated chemical compounds. RNA FISH was performed with probes against polyA-mRNAs (green) and cells were stained with antibodies against the SG marker G3BP-1 (magenta). Scale bars 20 μm.

Figure S5

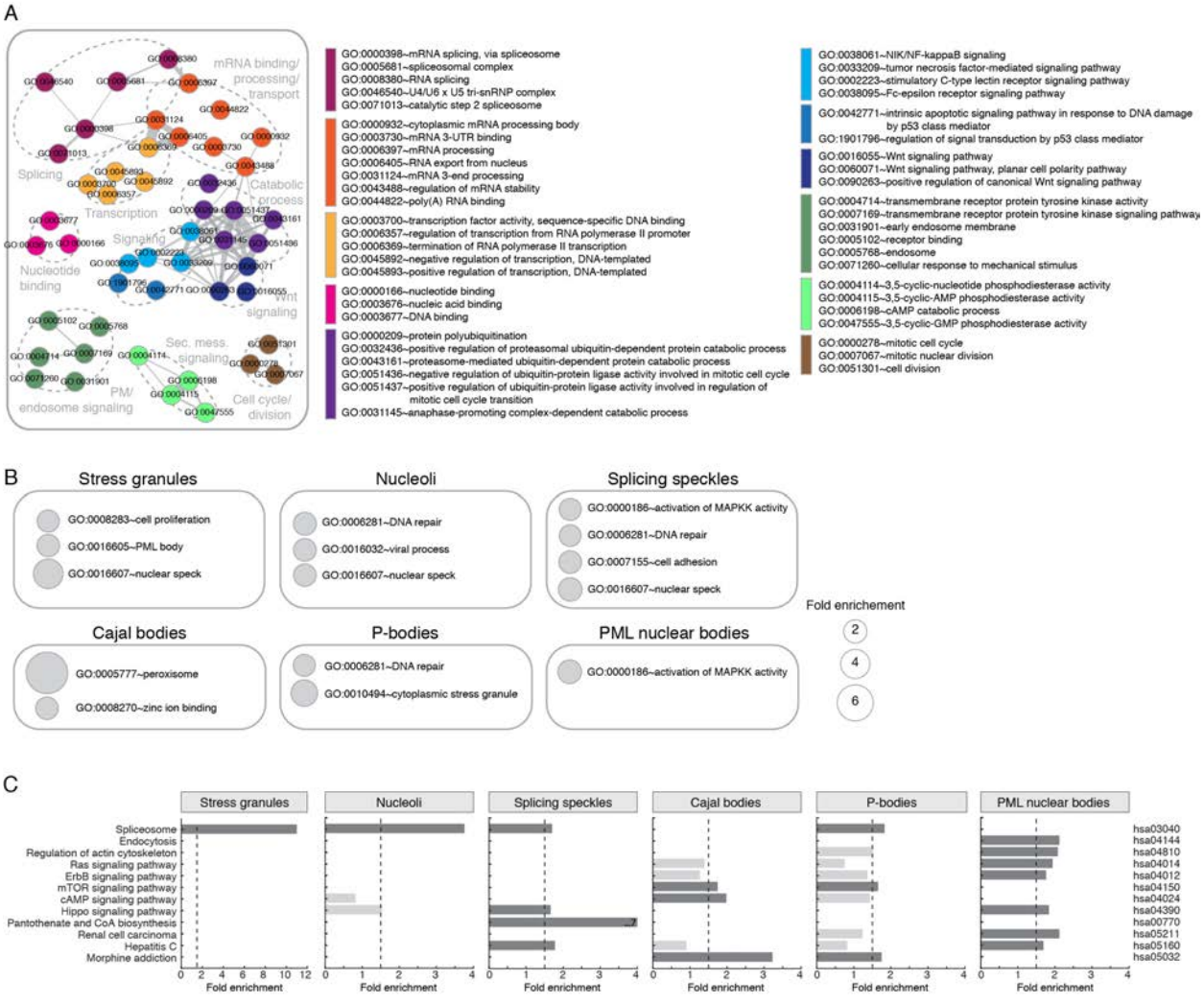

Figure S5: Functional annotations of MLO regulators. Related to Figure 5.

A) Network visualization of the functional annotation enrichments as shown in Figures **5A-5F**. Nodes represent the indicated Gene Ontology (GO)-terms and are grouped and colored according to functional similarity. Node edges (gray lines) are shown if two annotations share more than 20 percent of genes. **B)** Functional annotation enrichments calculated for gene perturbations resulting in either stress granules (SGs), or up-/down-regulation of nucleoli, P-bodies (PBs), splicing speckles, Cajal bodies (CBs), and PML nuclear bodies (PML-NBs), respectively, that were omitted in **A**. Circle size represents fold enrichment of the indicated GO-term. **C)** KEGG pathway enrichments calculated for gene perturbations resulting in either SGs, or up-/down-regulation of nucleoli, PBs, splicing speckles, CBs, and PML-NBs, respectively. Bars are colored in dark gray if fold enrichment was higher than 1.5 (black dotted line) for at least four genes.

Figure S6

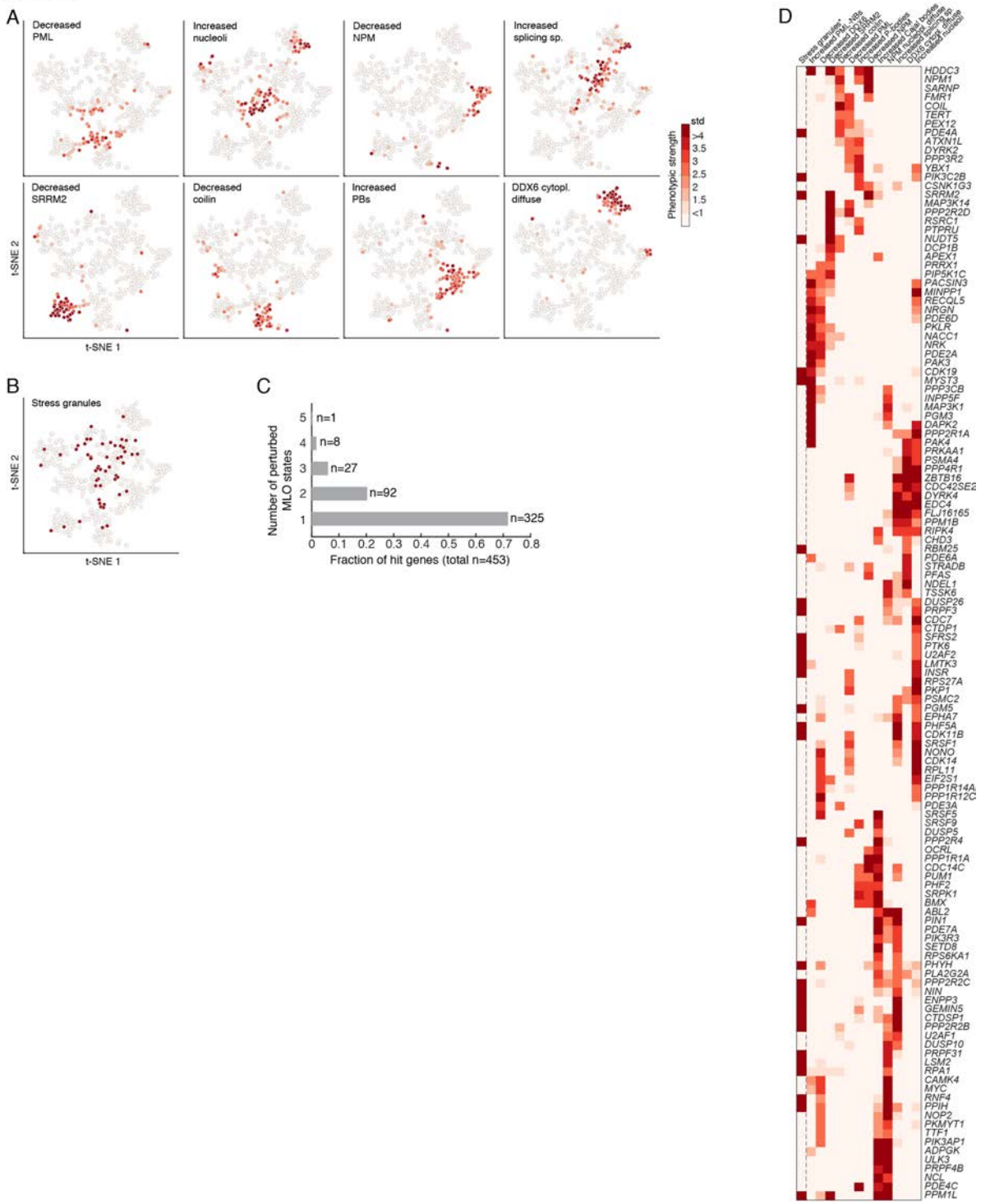

Figure S6: Gene perturbations with pleiotropic effects on MLOs. Related to Figure 6.

A) Gene t-SNE maps of 453 gene perturbations that are scored as hit for one or more perturbed MLO states. The distribution of eight perturbed MLO states is shown. Color indicates the respective phenotypic strength per gene. **B)** Gene t-SNE map of 453 gene perturbations that are scored as hit for one or more perturbed MLO states. Gene perturbations that resulted in increased fraction of cells with SGs are colored in red. **C)** Fractions of the 453 gene perturbations that display one or more perturbed MLO states. **D)** Clustered heat map of the phenotypic strength of the 128 gene perturbations that display more than two perturbed MLO states. The asterisk indicates that the phenotype of SG formation is Boolean, thus was set to the highest phenotypic strength and was not used for clustering.

Figure S7

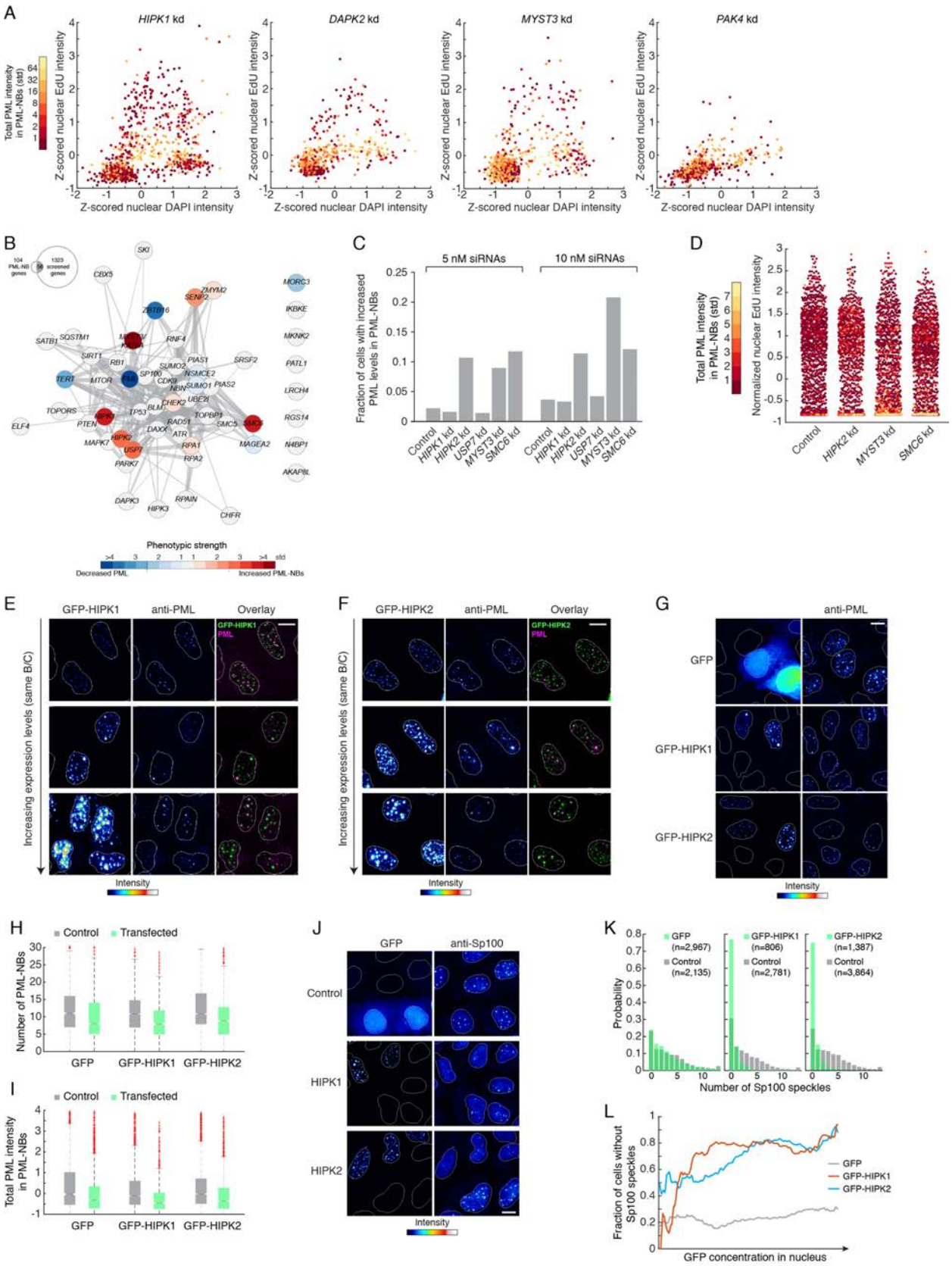

Figure S7: HIP kinases regulate PML-NB size and integrity. Related to Figure 7.

A) DAPI and EdU intensity of single cells treated for 72 hours with siRNAs against *HIPK1*, *DAPK2*, *MYST3* and *PAK4*, respectively. Color indicates total PML intensity in segmented PML-NBs as stds from the mean of all controls. Note that the total nuclear EdU intensity also includes weak signal of the protein

stain succinimidylester. **B)** Cell cycle classifications for control cells and cells treated for 72 hours with siRNAs against *HIPK2* and *EWSR1*, respectively. DAPI and EdU intensity of single cells are plotted as in Figure 7B. Colors indicate classification as G1, S or G2 cell. Box (dashed black outline) highlights cells with increased PML-NBs and reduced EdU incorporation as in Figure 7B. Note that the total nuclear EdU intensity also includes weak signal of the protein stain succinimidylester. **C)** STRING network visualization of 56 genes representing the overlap of 104 genes with a functional annotation for PML-NBs (GO:0016605~PML body, derived from QuickGO) with the set of 1323 screened genes. Edges (gray lines) indicate the relative physical and functional interaction strength derived from STRING database. Node colors indicate the effect of the respective gene perturbation on PML-NB morphology in stds from the mean of all perturbations and controls. Genes with a phenotypic strength higher than 2.5 stds from the mean are considered hits. **D)** Overexpressed GFP-tagged HIPK1 and HIPK2 partially colocalize with PML in PML-NBs. Cells were fixed after 24 hours of transfection and stained with antibodies against PML. Images showing the GFP signal (left column in both panels) are rescaled the same. Images showing the overlay of the GFP signal with PML (right column in both panels) are rescaled individually. Nuclear segmentation is shown in white. Scale bars 10 µm. **E-G)** Overexpressed GFP-tagged HIPK1 and HIPK2 do not alter PML speckle abundance or intensity. **E**) Representative images of cells transfected for 24 hours with GFP, GFP-HIPK1, or GFP-HIPK2, and stained with antibodies against PML. Nuclear segmentation is shown in white. Scale bar 10 µm. **F)** Number of segmented PML-positive speckles in control cells (gray boxes, n=1,445-4,029) and cells transfected (green boxes) for 24 hours with GFP (n=2,621), GFP-HIPK1 (n=975), and GFP-HIPK2 (n=1,143), respectively. **G)** Total PML intensity in segmented PML-positive speckles of the same cells as quantified in **F**. **H-J)** Overexpressed GFP-tagged HIPK1 and HIPK2 dissolve Sp100-positive speckles in a concentration-dependent manner. **H)** Representative images of cells transfected for 24 hours with GFP, GFP-HIPK1, or GFP-HIPK2, and stained with antibodies against Sp100. Nuclear segmentation is shown in white. Scale bar 10 µm. **I)** Probability distributions of the number of nuclear Sp100 speckles in control cells (gray) and cells transfected for 24 hours with GFP, GFP-HIPK1, and GFP-HIPK2, respectively. Numbers of quantified cells are indicated. **J)** Fractions of cells with no nuclear Sp100 speckles. Cells were treated as above and fractions were calculated for GFP-positive cells (GFP control: n=2,967, GFP-HIPK1: n=806, GFP-HIPK2: n=1,387). The fractions were calculated as a function of nuclear GFP concentration (sliding window method on log10-transformed intensity values) and smoothed afterwards.