Integrative Chemical Genetics Platform Identifies Condensate Modulators Linked to Neurological Disorders

Dylan Poch,[1] Chandrayee Mukherjee,[1] Sunanda Mallik,[1] Vanessa Todorow,[1] E.F. El-siena Kuiper,[1,2] Nalini Dhingra,[3] Yulia V Surovtseva,[3] Christian Schlieker.[1,4,†]

[1]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511, USA.

[2]Max Planck Institute of Biophysics, Mechanisms of Cellular Quality Control, Frankfurt a. M., Germany

[3]Yale Center for Molecular Discovery, Yale University, West Haven, CT 06516, USA.

[4]Department of Cell Biology, Yale School of Medicine, New Haven, CT 06520, USA.

[†]Correspondence: christian.schlieker@yale.edu

ABSTRACT

Aberrant biomolecular condensates are implicated in multiple incurable neurological disorders, including Amyotrophic Lateral Sclerosis, Frontotemporal Dementia, and DYT1 dystonia. However, the role of condensates in driving disease etiology remains incompletely understood. Here, we establish myeloid leukemia factor 2 (MLF2) as a disease-agnostic biomarker of phase transitions. Using MLF2 as a high-content screening readout enables the identification of condensate modulators across both chemical and genetic perturbations. We uncover FDA-approved drugs that potently modulate condensates, validating our pipeline for facile drug discovery. Leveraging our methodology for a genome-wide CRISPR/Cas9 screen reveals genes linked to primary microcephaly and related neurodevelopmental disorders that result in proteotoxic condensates. Machine learning establishes two phenotypic clusters, with RNF26 deletion provoking nuclear envelope condensates reminiscent of nuclear pore defects, while loss of ZNF335, a known microcephaly-causative gene, drives accumulation of distinct nucleoplasmic condensates. Our study establishes a link between aberrant phase transitions and microcephaly, highlighting corrective chemical modulators.

KEYWORDS

Biomolecular condensates, high-content screening, microcephaly, dystonia, ubiquitin, MLF2, nuclear pore, pyrithione zinc, machine learning, neurodegeneration, stress granules, ubiquitin ligase.

TEASER

Chemical genetic methods uncover link between aberrant biomolecular condensates and neurodevelopmental disorders.

INTRODUCTION

Dysregulation of physiological phase transitions is a driver of many neurological disorders, including Amyotrophic Lateral Sclerosis (ALS), Frontotemporal Dementia (FTD), and DYT1 dystonia.(*1–5*) Biomolecular condensates rely on a thermodynamically favorable separation between a dense and dilute phase.(*6*, *7*) When properly regulated, biomolecular condensates allow efficient, fine-tuned microenvironments often containing diverse components, ensuring maintenance of cellular homeostasis while buffering against external perturbations.(*8–10*) However, in pathological states referred to as condensatopathies, these condensates exhibit reduced dynamicity and undergo aberrant phase transitions, leading to protein aggregation or amyloid formation that is not easily reversible.(*11–13*) According to some estimates, up to one-third of common diseases can be linked to condensate-forming proteins.(*14*) Determining the mechanisms and cellular activities of aberrant phase transitions is essential for devising therapeutic strategies that tune condensate properties to counteract pathological transitions.

One favorable phase-state is observed in the central channel of the nuclear pore complex (NPC). The NPC serves as the intermediary gate at which nuclear transcription is separated from cytoplasmic translation in eukaryotic organisms. Within the central channel of the NPC, phenylalanine-glycine nucleoporins (FG-nups) have a high proportion of intrinsically disordered regions that set up a favorable phase, contributing to the selective barrier of the NPC.(*15–18*) Recently, our lab proposed that this phase may additionally represent an optimal environment that counteracts protein aggregation and promotes folding.(*19*) However, disruptions to NPC assembly or nucleocytoplasmic transport have been strongly implicated in neurological diseases, including ALS and DYT1 dystonia,(*1*, *20–22*) a debilitating movement disorder characterized by involuntary muscle contractions and painful sustained postures.(*23*, *24*)

DYT1 dystonia is caused by a mutation within the TOR1A gene that disrupts activation of the encoded AAA+ ATPase,(*25–29*) TorsinA, resulting in the formation of pathological nuclear envelope condensates.(*1*, *20*, *21*, *30*, *31*) As with many condensatopathies, there are only palliative interventions, with no curative or preventative treatments for dystonia.(*23*, *24*) These nuclear envelope condensates accumulate K48-linked polyubiquitinylated proteins that evade proteasomal degradation and sequester

chaperone proteins, causing cellular proteotoxicity.(*1*, *21*, *32*) Thus, modulation of aberrant nuclear condensates by either direct dissolution or modulation of their proteotoxic properties represents potential therapeutic approaches. At present, limitations to high-throughput screening campaigns targeting condensatopathies include a lack of well-defined condensate targets and broadly applicable condensate biomarkers.

In this study, we identify Myeloid leukemia factor 2 (MLF2) as a disease-agnostic marker of aberrant condensates. Previously found to be enriched in DYT1 dystonia-related nuclear envelope condensates linked to NPC assembly,(*21*) we discover MLF2 also accumulates in proteotoxic nuclear condensates and co-localizes with TDP-43 and G3BP1 in stress granules. Using high-content screening with MLF2-GFP as our readout, we uncovered both pharmacological and previously uncharacterized genetic modulators of aberrant phase transitions. Several FDA-approved small molecules modulate nuclear condensates, providing a tractable methodological framework for discovering therapeutics for condensatopathies. Combining CRISPR knockout (KO) genetic screening with deep learning approaches allowed us to identify RNF26 and ZNF335 as protective factors against phenotypically distinct aberrant condensates. Notably, 15 out of the top 50 genes in our screen are strongly associated with neurodevelopmental disorders, including seven directly linked to primary microcephaly or related congenital brain malformations. This work establishes a connection between nuclear condensates, perturbed nuclear proteostasis, and neurodevelopmental disorders and establishes a robust drug discovery platform to facilitate the development of innovative pharmacological strategies.

RESULTS

MLF2 is a Biomarker for Aberrant Condensates Associated with ALS/FTD and DYT1 Dystonia.

Having previously established a connection between impaired protein quality control and condensates in a torsin-deficient DYT1 dystonia HeLa cell model (4TorKO cells),(*1*, *21*) we aimed to develop biomarkers to facilitate the identification of condensate modulators across other condensatopathies. MLF2 represents a major protein constituent of nuclear envelope condensates within 4TorKO cells(*21*) and readily immerses into FG-

nup condensates *in vitro*, likely due to its highly enriched Methionine and Arginine residues.(*1*) Since many condensates share aromatic side chains as part of their molecular grammar,(*33*) we reasoned that MLF2 might localize to additional aberrant condensates, likely through π-cation interactions.

Dysregulation of stress granules (SGs) has been pathologically associated with several neurological disorders, including ALS and FTD.(*11*) To asses if MLF2 localizes with SGs, we co-expressed G3BP1 and TDP-43 in wild-type (WT) HeLa cells treated with 0.5 mM sodium arsenite ($NaAsO_2$) for 30 minutes, a well-established condition for SG induction.(*34–37*) MLF2 exhibited strong co-localization with both G3BP1-GFP and TDP-43-tdTomato, regardless of $NaAsO_2$ exposure (Figure 1A). These findings suggest that MLF2 serves as a robust biomarker for aberrant condensates associated with distinct neurological disorders. To exploit the propensity of MLF2 to immerse into both stable (nuclear envelope condensates) and transient (SGs) aberrant phases, we designed doxycycline (Dox) inducible fluorophore-derivatized variants of MLF2 as condensate biomarkers. This enabled us to develop high-throughput chemical and genetic assays to screen for condensate modulators (Figure 1B).

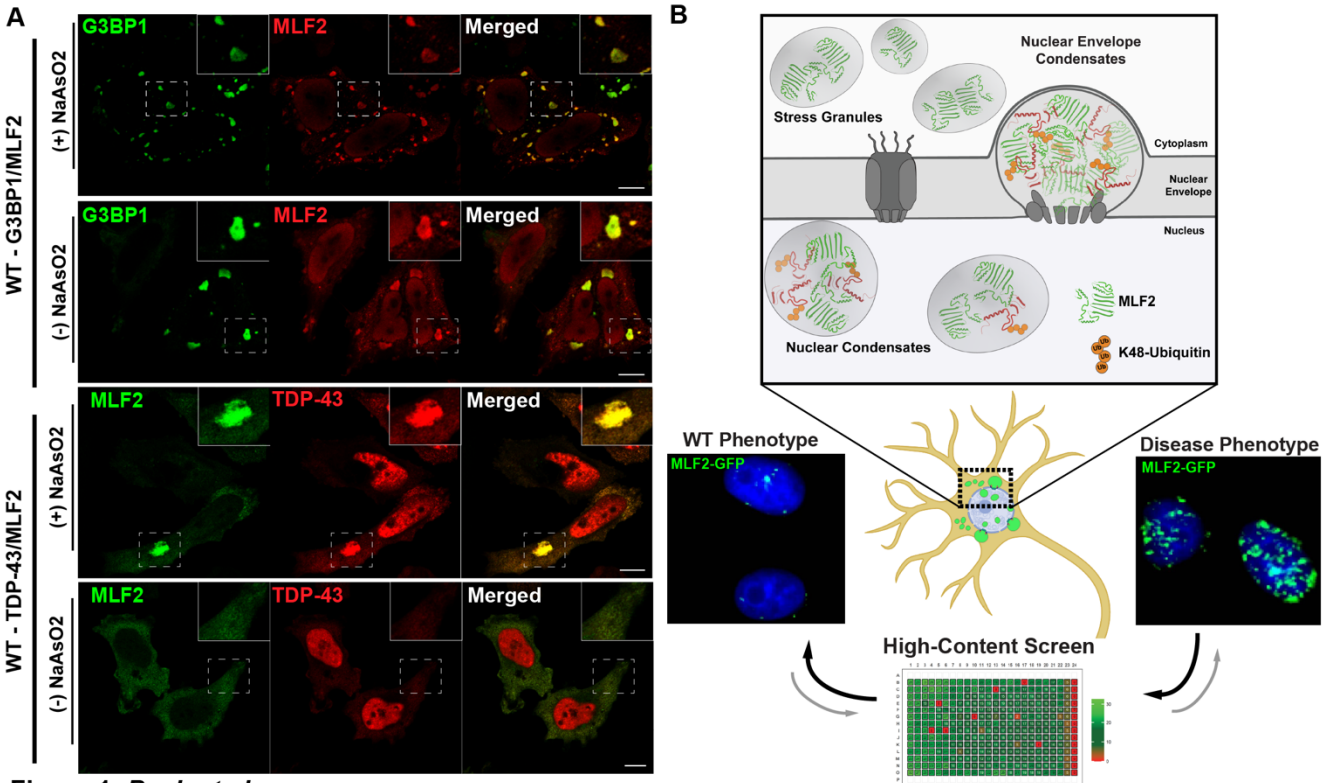

Figure 1, *Poch et al.*

Fig. 1. MLF2 is a biomarker for aberrant phase transitions across neurological disorders. (**A**) Utilizing MLF2 as a SG reporter. MLF2 co-localizes with G3BP1 (top panels) and TDP-43 (bottom panels) before and after SG formation in response to oxidative stress (0.5 mM $NaAsO_2$, 30 min). (**B**) MLF2 detects aberrant condensates associated with multiple disease states. The ability of MLF2 to immerse into condensates can be leveraged for high-content imaging platforms to systematically profile additional disease-relevant phase transitions. Arrows represent potential opportunities to manipulate disease phenotypes using chemical genetic methods. Representative images captured during high-content screening of WT HeLa (left) and 4TorKO HeLa (right) cells expressing MLF2-GFP.

Establishing MLF2-GFP as an Aberrant Condensates Reporter for High-Content Screening

Leveraging the condensate biomarker properties of MLF2, we developed a miniaturized, high-content assay using Dox-inducible MLF2-GFP cells in a 384-well format. This facilitated the design of phenotypic, high-content arrayed screens to identify chemical and genetic modulators of nuclear condensates. Specifically, each well contained either a small molecule applied to 4TorKO cells or synthetic single guide RNAs (sgRNAs) delivered to Cas9-integrated WT cells, designed to suppress or promote nuclear condensates, respectively. To analyze the high-resolution images (over 10k images/plate), we developed a computational condensate screening suite, *CondenScreen*. This open-source CellProfiler and R script was designed for the rapid quantification and statistical analysis of aberrant condensates (Figure S1A).

While many robust computational resources are available for monitoring condensates (see *PhaseMetrics* and *Punctatools*)(*38*, *39*), *CondenScreen* is specifically designed for large-scale, high-content screening analysis using a customizable, open-source *CellProfiler 4* (*40*, *41*) pipeline. Beyond quantifying the number of condensates at both single-cell and aggregate levels, *CondenScreen* automatically computes the robustness of the screen using statistical indicators such as the Z'-Score (and BZ'-Score), signal-to-background (SB), normalized effect, and the coefficient of variation (CV). *CondenScreen* provides a simple graphical user interface (GUI), allowing the user to select plate size, layout, and choose whether the user's readout is signal "ON" vs signal "OFF", altering downstream statistics (Figure S1A). Using our corresponding Bash script, multiple headless *CellProfiler* instances can be executed simultaneously on a Simple Linux Utility for Resource Management (SLURM) cluster. Integrating the MLF2-GFP inducible system with our *CondenScreen* analysis pipeline ensures rapid identification of phase-transition modulators.

108 **High-Content Screen Identifies FDA-approved Drugs that Modulate Condensates**
109 To identify chemical matter that can modulate aberrant nuclear envelope condensates,
110 we employed our MLF2-GFP inducible 4TorKO DYT1 cell model. We screened 1760
111 drugs from the Pharmakon library, comprising 1360 US and 400 international drugs, at
112 a 10 μM working concentration. To minimize false positives due to differential reporter
113 expression, Dox (0.5 μg/ml) was added for four hours, followed by an additional 2 hours
114 with compound incubation.
115 *CondenScreen* reported an average Z'-score of 0.73, indicating pronounced sep-
116 aration between the positive (0.1% DMSO-treated WT HeLa) and negative (0.1%
117 DMSO-treated 4TorKO HeLa) controls.(*42*) Additionally, the ratio between the negative
118 and positive controls (SB) was 930, while the SD and CV of DMSO-treated cells were
119 2.1% and 9.2%, respectively, further confirming assay robustness.
120 To control for plate-specific column, row, and edge effects, we assigned each
121 compound a BZ-score based on its ability to decrease the number of nuclear MLF2-
122 GFP foci (Figure 2A, see methods). The distribution of the average number of MLF2-
123 GFP foci per condition was near-normally distributed (Figures 2B, 2C). Using a standard
124 BZ-score cutoff of three, we identified 20 putative hit compounds that significantly de-
125 pleted MLF2-GFP condensate foci. Top compounds included plumbagin, sanguinarium
126 chloride, thiram, piroctone olamine, penfluridol, bronopol, thimerosal, triclosan, disulfi-
127 ram, celastrol, dihydrocelastryl diacetate, and pyrithione zinc (PZ). A full list of the
128 screened drugs can be found in the supplementary materials (Table S1).
129 Top-scoring compounds were resolubilized to construct a 12-point dose-re-
130 sponse (n = 4 replicates), covering over 5 orders of magnitude to determine the calcu-
131 lated IC_{50}. (Figure 2D, Figure S1B). The normalized percent effect and percent toxicity
132 (calculated as a function of viability, nuclear area, and nuclear morphology) were quanti-
133 fied for each concentration to determine potential therapeutic windows. Four structurally
134 diverse drugs demonstrated a potent dose-dependent relationship, which included
135 celastrol (IC_{50} = 124.8 nM), PZ (IC_{50} = 601.6 nM), plumbagin (IC_{50} = 7.4 μM), and disulfi-
136 ram (IC_{50} = 10.8 μM) (Figure 2D).
137
138

139

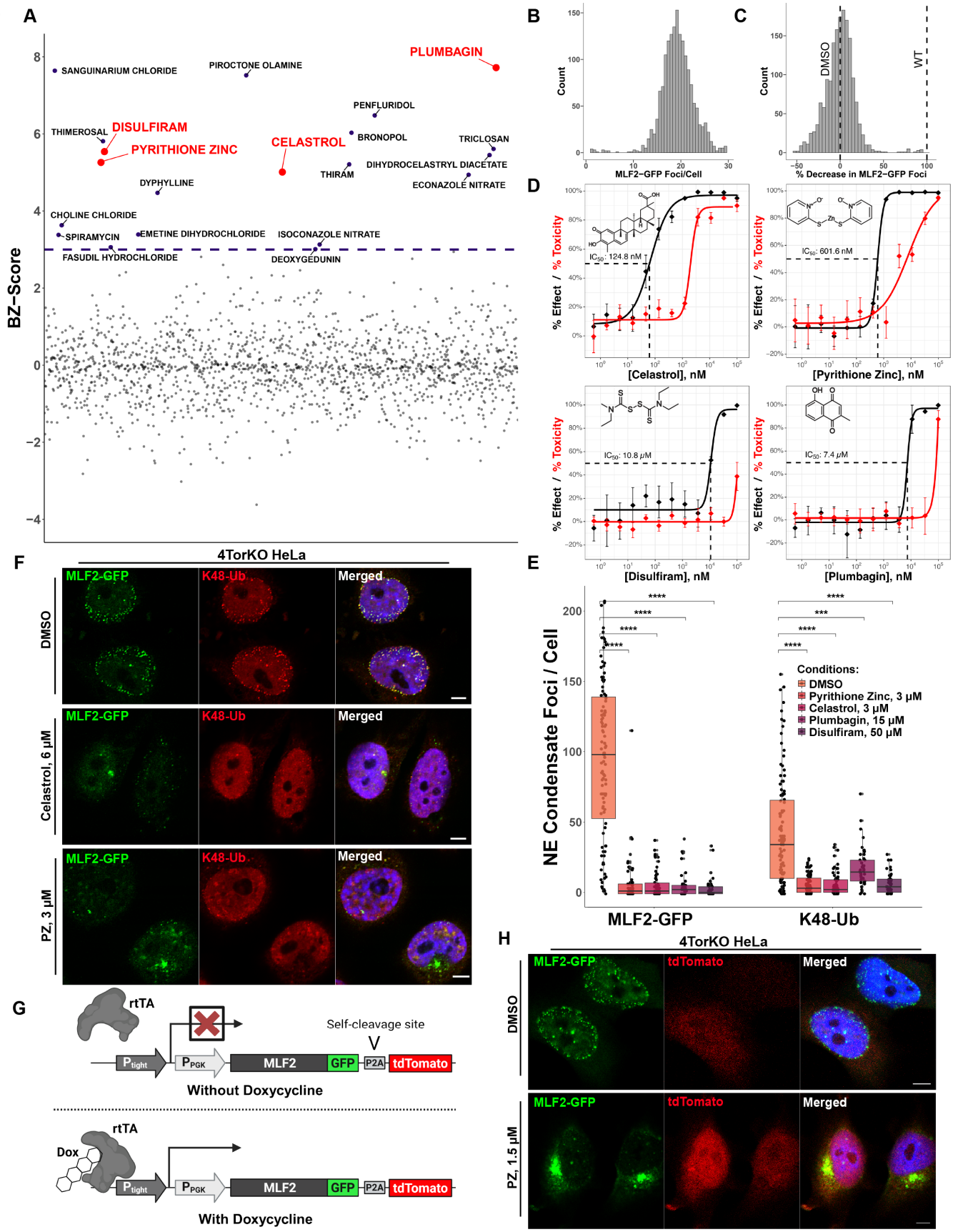

Figure 2, *Poch et al.*

Fig. 2. Chemical screen identifies drugs that modulate nuclear envelope condensates. (A) Drugs from Pharmakon 1760 screen decreases nuclear MLF2-GFP accumulation in 4TorKO HeLa cells. Compounds (circles) passing significance threshold (dotted-line, BZ-score > 3) are labeled purple, while red-labeled dots denote drugs selected for downstream validation. (**B**) Distribution in the average number of MLF2-GFP foci/cell and (**C**) percent decrease of MLF2-GFP foci per cell, normalized to the screening population. (**D**) Hit compounds are dose dependent. Determined IC_{50} from 12-point dose response (four orders of magnitude, n = 4). Black line represents % effect (decrease in MLF2-GFP foci), red line represents % toxicity relative to DMSO (four-parameter-log-logistic model). (**E**) Drugs modulate accumulation of K48-Ub proteins. Dots depict the number of MLF2-GFP (left, 6h Dox) or α-K48-ubiquitin (right) foci/cell after treatment with drug or DMSO. Asterisks indicate Bonferroni-corrected p-values calculated by Mann-Whitney U testing (***: p <= 0.001 ****: p <= 0.0001). (**F**) Representative immunofluorescence images of MLF2-GFP (green) and α-K48-ubiquitin (red) in torsin-deficient cells treated with DMSO compared to either PZ (3 μM) or celastrol (6 μM). (**G**) Schematic of inducible MLF2-GFP-P2A construct without (top) or with (bottom) Dox. (**H**) Confocal images of torsin-deficient cells expressing MLF2-GFP-P2A-tdTomato treated with either DMSO or 1.5 μM PZ treatments (6h). Scale bars, 5 μm.

Pyrithione Zinc Disrupts K48-Ubiquitinylated Protein Accumulation within Nuclear Condensates

To determine the optimal concentration, we tested each of the four candidate compounds—celastrol, PZ, plumbagin, and disulfiram—at four concentrations, including the IC_{50} itself, at 2x above/below the IC_{50}, and 8x above the IC_{50} concentrations, using a 3 h compound treatment. The resulting boxplot reflects each compound's effect at the single-cell level (Figure S1C). Consistent with the significant cell-cycle dependency of NE condensate formation previously demonstrated,(*21*) we observed significant variability in the number of MLF2-GFP foci/cells. However, all compounds significantly reduced perinuclear MLF2-GFP accumulation at 2x and 8x their respective IC_{50} (Figure S1C).

To ensure consistent reduction across different condensate markers, we used K48-linked ubiquitin (K48) as an independent and endogenous biomarker since 4TorKO cells have reduced protein turnover resulting in the accumulation of polyubiquitinylated proteins within nuclear envelope condensates.(*1*, *21*) Using immunofluorescence staining with α-K48 antibodies, all four compounds demonstrated a statistically significant decrease in endogenous, perinuclear K48 foci, replicating our results at reducing MLF2-GFP accumulation (Figure 2E). Given the optimal concentration profile of PZ and celastrol (Figure 2D), we confirmed that each compound results in the redistribution of both MLF2-GFP and endogenous K48 foci using confocal microscopy (Figure 2F). Importantly, this data validates that the mechanism of action either directly or indirectly mediates endogenous protein cargo within the nuclear envelope condensates of DYT1 dystonia.

A significant limitation to phenotype screening is the prevalence of identification of false positives that inhibit biomarker expression by non-specifically disrupting cellular

165 protein transcription or translation. To circumvent this, we constructed a HeLa cell line
166 with an inducible bicistronic MLF2-GFP-P2A-tdTomato reporter (Figure 2G). The P2A
167 site induces ribosomal skipping, allowing the translation of two separate polypeptides,
168 an N-terminal MLF2-GFP and a C-terminal tdTomato, both derived from the same
169 mRNA transcript.(*43*) We selected MLF2-GFP-P2A-tdTomato 4TorKO cells using Fluo-
170 rescence-Activated Cell Sorting (FACS) (Figure S1D), before inducing MLF2-GFP ex-
171 pression and treating cells with either PZ (0.75 μM or 1.5 μM, 6 h) or celastrol (1.5 μM
172 or 3 μM, 6 h) . We quantified the number of MLF2-GFP foci and tdTomato expression
173 between either uninduced cells treated with DMSO vehicle control, induced cells treated
174 with DMSO, and induced cells treated with either 100 μg/mL cycloheximide, PZ, or
175 celastrol. Concentrations of both celastrol and PZ at 1.5 μM had a significant decrease
176 in the number of MLF2-GFP foci/cell compared with the DMSO control (Figure S1E).
177 While PZ did not reduce the expression of secondary-tdTomato reporter (Figures 2H,
178 S1F), celastrol slightly, albeit at a non-statistically significant level, decreased secondary
179 reporter expression compared with DMSO-treated cells (Figure S1F).

180 To quantify the potential effects of a compound inhibiting reporter expression, we
181 developed an empirical metric, the Condensate Normalization Score (CNS). The CNS is
182 based on the differential fluorescence between the condensate biomarker and second-
183 ary reporter. A high CNS-value indicates a reduction in condensate biomarkers with
184 minimal impact on overall protein expression (see methods). Both celastrol and PZ-
185 treated cells had a significantly increased median CNS (Figure S1G) compared to all
186 controls, indicating a modulatory mechanism independent of reporter expression. Re-
187 maining 4TorKO-induced nuclear envelope condensates could be visualized using
188 transmission electron microscopy (TEM) after a 3 h treatment of 2 μM PZ (Figure S1H).
189 Further testing of additional concentrations and time points may reveal differential kinet-
190 ics between condensate resolution compared to modulation of condensate cargo
191 (MLF2/K48-Ub-proteins).

192 To identify the structural features responsible for the condensate-modifying activ-
193 ity of PZ, we examined whether this effect originates from its pyrithione moiety (2-Mer-
194 captopyridine N-oxide) or the zinc ion. To this end, we treated 4TorKO cells with either
195 DMSO vehicle control, PZ alone, or PZ with zinc chelator TPEN (N,N,N',N'- *tetrakis*-(2-

196 Pyridylmethyl)ethylenediamine) at a 3 molar equivalent.(*44*) After treating with TPEN to
197 chelate zinc, PZ failed to reduce MLF2-GFP condensate accumulation in 4TorKO cells
198 (Figures S2A, S2B). Furthermore, we tested whether zinc alone is sufficient to deplete
199 nuclear MLF2-GFP condensate foci. Due to the lower membrane permeability, 4TorKO
200 cells were treated with higher concentrations of zinc sulfate monohydrate ($ZnSO_4$).
201 Treatment with 2 mM $ZnSO_4$ alone decreased MLF2-GFP foci (Figure S2C, S2D) while
202 maintaining tdTomato expression (Figure S2E), resulting in a high CNS value (Figure
203 S2F). Thus, PZ modulates nuclear condensates through a zinc-dependent mechanism.
204

205 **Genome-Wide CRISPR Screening Uncovers Condensate Modulators Implicated in**
206 **Microcephaly and Neurodevelopmental Disorders**

207 Having established a robust screening pipeline, we next performed a genome-
208 wide CRISPR/Cas9 KO screen in WT HeLa cells to identify genes linked to phase mod-
209 ulation. Genetic hits that promote nuclear condensate formation may shed light on yet
210 unknown cellular activities as well as facilitating the identification of pharmacological tar-
211 gets. We screened an arrayed library of 19,195 human genes using the Horizon Discov-
212 ery Edit-R synthetic sgRNA collection. To increase KO efficiency, we designed and
213 FACS-sorted a WT HeLa cell line that constitutively expresses Cas9 with Dox-inducible
214 MLF2-GFP. Each well was reverse-transfected in a 384-well plate format using a pool of
215 3 sgRNAs targeting different regions of the same gene (Figure 3A). Cells were incu-
216 bated for 72 h, and MLF2-GFP was induced with Dox 16 h before fixation. Using *Con-*
217 *denScreen* to quantify the number of MLF2-GFP foci per nucleus, and then comparing
218 that number between the positive (TOR1A/B sgRNAs) and negative (non-targeting (NT)
219 sgRNA) controls, we calculated an average Z' value of 0.52. The calculated SB was 4.8,
220 and the SD and CV of the NT control were 3.1 and 9.7, respectively. We determined the
221 assay quality score to be sufficiently robust as indicated by similar scores in the litera-
222 ture.(*36*, *45*)
223
224

225

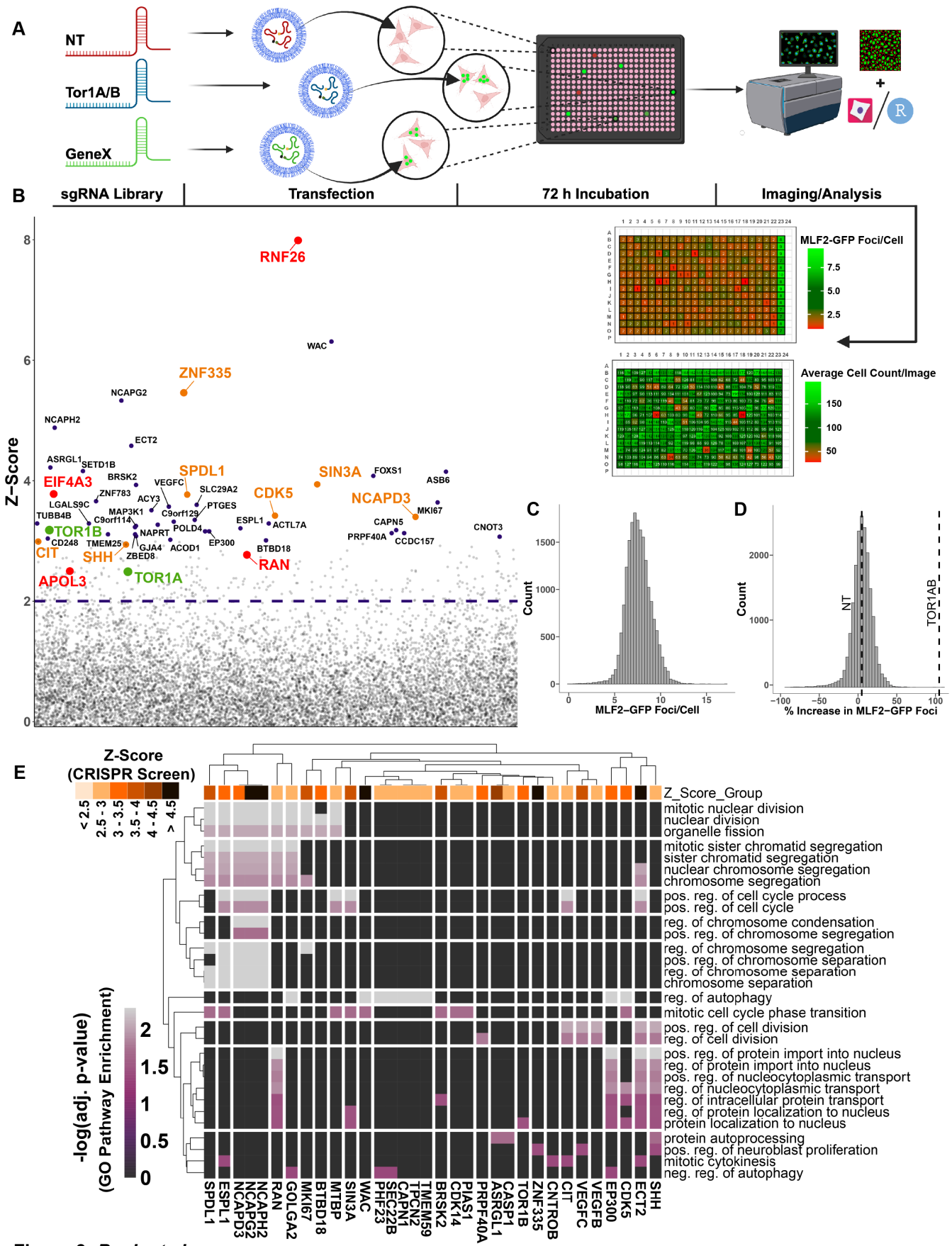

Figure 3, *Poch et al.*

Fig. 3. CRISPR KO Screen reveals genes that modulate nuclear condensates. (**A**) Schematic workflow of the genome-wide CRISPR KO screen using WT HeLa cells. Plates represent direct outputs provided by *CondenScreen*, indicating the number of MLF2-foci/cell (top plate) and cell count per image (bottom plate). (**B**) Genes hits from CRISPR KO screen associated with microcephaly modulate nuclear condensates. Dotted line represents significance threshold (Z-Score > 2) while highlighted and labeled genes have a Z-Score > 3. Green dots represent the TorsinA and TorsinB KO conditions; orange dots represent genes linked to microcephaly; and red dots indicate genes of additional interest. (**C**) Distribution of the average number of MLF2-GFP foci/cell and (**D**) percent increase in MLF2-GFP foci/cell, normalized to the screening population. (**E**) Heatmap of the 30 highest scoring GO biological processes (rows) and associated genes (columns) from the top 100 candidate CRISPR KO hits. Vertical scale legend indicates -log(10) Benjamini-Hochberg adjusted p-value, with lighter regions representing higher significance. Scaled boxes (top) indicate the Z-Score of each gene from the CRISPR Screen, with darker colors indicating a higher Z-score (horizontal scale bar). Euclidean distance hierarchical clustering for GO-enriched biological pathways and genes is shown with the left and top dendrogram, respectively.

We defined a candidate gene as any gene deletion that resulted in a Z-score > 2 (Figure 3B). The top 10-ranked candidate genes were: RNF26, WAC, ZNF335, NCAPG2, NCAPH2, ECT2, ASRGL1, SETD1B, ASB6, and FOXS1 (see Table S2 for comprehensive list). Overall, the screen resulted in a near-normal distribution of the number of MLF2-GFP foci per cell (Figures 3C, 3D). Notably, we observed several genes linked to neurodevelopmental disorders, with high enrichment of microcephaly-associated genes (orange labeled dots, Figure 3B).(*46*, *47*) Both homozygous and compound heterozygous mutations to one of our top hits, zinc finger protein 335 (ZNF335, also identified as MCPH10 or NIF1) have been documented to cause primary microcephaly.(*48*–*53*) ZNF335 has been tied to the regulation of neurogenesis through modulation of the REST complex.(*47*)

Additional top hits of interest include, RNF26, an E3-ubiquitin ligase that has been implicated in regulating the endosomal system, serving as a “gatekeeper” for vesicles in the perinuclear ER cloud.(*54*, *55*) Furthermore, RNF26 ubiquitinylates SQSTM1(*54*) and has recently been tied to ER-sheet morphology and maintenance.(*55*) Other hits include APOL3 (Z-Score = 2.5), a putative regulator of membrane dynamics with detergent-like bactericidal properties;(*56*, *57*) Ran (Z-Score = 2.8), a GTP-binding nuclear protein; and EIF4A3/DDX48 (Z-Score 3.8), a DEAD-box ATPase. DEAD-box ATPases have recently been linked to the regulation of phase-separated membranelles organelles,(*58*) indicating a potential role of EIF4A3 in modulating phase transitions. The condensate phenotypes defined by nuclear MLF2 accumulation observed across gene deletions reveal key entry points into previously unexplored biology and highlight promising avenues for therapeutic discovery.

To determine if candidate genes are clustered into enriched genetic pathways, we hierarchically clustered the top 100 gene hits with their corresponding 30 enriched biological processes (q-value < 0.05) using Gene Ontology (GO) analysis(*59*) (Figure

252 3E; interactive webpage available, see methods) and illustrated pathway interconnect-
253 edness (Figures S3A-S3C). We observed enrichment (10 out of 30 GO terms) of chro-
254 mosomal and chromatid condensation, separation, or segregation pathways (Figures
255 3E, S3A). GO pathways were manually clustered into the following five categories: Au-
256 tophagy, Cell Cycle, Chromosomal Regulation, Protein Transport, and Other (Figure
257 S3B). Consistent with these enriched pathways, loss of protein homeostasis,(*1*, *60*) a
258 cell-cycle dependency of condensates,(*21*) and disruption to nucleocytoplasmic
259 transport(*61*, *62*) have been documented in condensatopathies. Additionally, dysregula-
260 tion of both cell cycle checkpoints and chromosomal regulation has been linked to pri-
261 mary microcephaly.(*63*–*65*) Indeed, seven out of the top 50 gene hits have been linked
262 to microcephaly.(*66*, *67*) Beyond the enrichment of individual gene hits linked to micro-
263 cephaly, encephalopathies, and other neurodevelopmental disorders, additional en-
264 riched pathways of interest included organelle fission and positive regulation of neuro-
265 blast proliferation.
266 To validate the three highest candidate genes, we knocked out ZNF335, RNF26,
267 and the WW Domain Containing Adaptor with Coiled-Coil (WAC) and binned them ac-
268 cording to the number of nuclear MLF2-GFP foci/cell (5-15, 15-30, or >30 foci/cell) (Fig-
269 ure S3D). While WAC depletion had only a modest effect, RNF26 KO and ZNF335 KO
270 cells displayed pronounced nuclear condensate formation as indicated by MLF2-GFP
271 accumulation (Figure S3D), with ZNF335-deficient cells resulting in the greatest number
272 of condensates.
273
274 **Supervised Machine Learning Links Genetic Perturbations to Distinct Condensate**
275 **Phenotypes**
276 To capture subtle condensate phenotype variations between different gene KO condi-
277 tions that may be overlooked using either conventional quantification methods, we de-
278 veloped a supervised machine learning model. (Figure S4A) This is based on a deep re-
279 sidual convolutional neural network (ResNet18),(*68*) trained on high-content imaging
280 data from our CRISPR KO screen (Figure 4A). To initially train our model, we seg-
281 mented screening images into single-cell patches containing either RNF26 KO or
282 ZNF335 KO cells, due to their high significance and our previous validation (Figure

283 S3D), along with cells treated with NT sgRNAs-treated as a control. We then employed
284 the ResNet18 convolutional residual neural network backbone (ImageNet)(*69*) to learn a
285 regression mapping anchored by three reference points: RNF26 KO (-1), non-targeting
286 (0), and ZNF335 KO (+1). Based on phenotypic similarity to each anchor condition, indi-
287 vidual cells were assigned a continuous regression score from -1 to +1. This was de-
288 rived from 512-dimensional feature embeddings capturing both nuclear and condensate
289 morphology, along with the number and spatial distribution of nuclear condensates.
290 After optimization, the model was refined until it correctly predicted 70% of indi-
291 vidual cells being correctly classified as closest to their respective anchor condition (Fig-
292 ure 4B). The model achieved a robust Area Under the Receiver Operating Characteris-
293 tic curve (AUC ROC) of 0.94 (Figure 4C), a minimum loss of 0.14 (Figure S4B), and a
294 true vs predicted correlation of 0.73 (Figure S4C). While the similarity distribution of la-
295 beled cells (Figure 4D) demonstrated overlap between conditions, with only ~55% of
296 RNF26 cells being identified as true RNF26 KO cells (Figure 4E), this outcome was ex-
297 pected given the high variability in transfection efficiency. Fewer than 2% of RNF26 KO
298 cells were misclassified as ZNF335 KO cells and vice versa, strengthening our convic-
299 tion of the model's ability to discriminate and report biologically meaningful results
300 against technical noise (Figure 4E).
301 Once trained, our model produced embeddings extracted from the ResNet fea-
302 ture space and prediction scores for each CRISPR KO hit that achieved a Z-Score > 2
303 (358 genes). For each KO condition, the model selected up to 500 randomly sampled
304 single-cell embeddings, and similarity metrics were calculated against RNF26, ZNF335,
305 and NT control anchors. The centroid embedding vector for each gene KO condition
306 was visualized in 2D with UMAP, an unsupervised dimensionality reduction method,(*70*)
307 revealing three distinct clusters with KMeans clustering. A total of three clusters were
308 chosen based on a silhouette coefficient of 0.53 (Figures S4D, S4E). The resulting
309 UMAP plot separated NT controls (cluster 1), RNF26 KO cells (cluster 2), and ZNF335
310 KO cells (cluster 3), allowing genes to be classified by their phenotypic similarity to ei-
311 ther the ZNF335 and RNF26 KO conditions, or to be further scrutinized as potential
312 false positives (cluster 1) (Figure 4F).

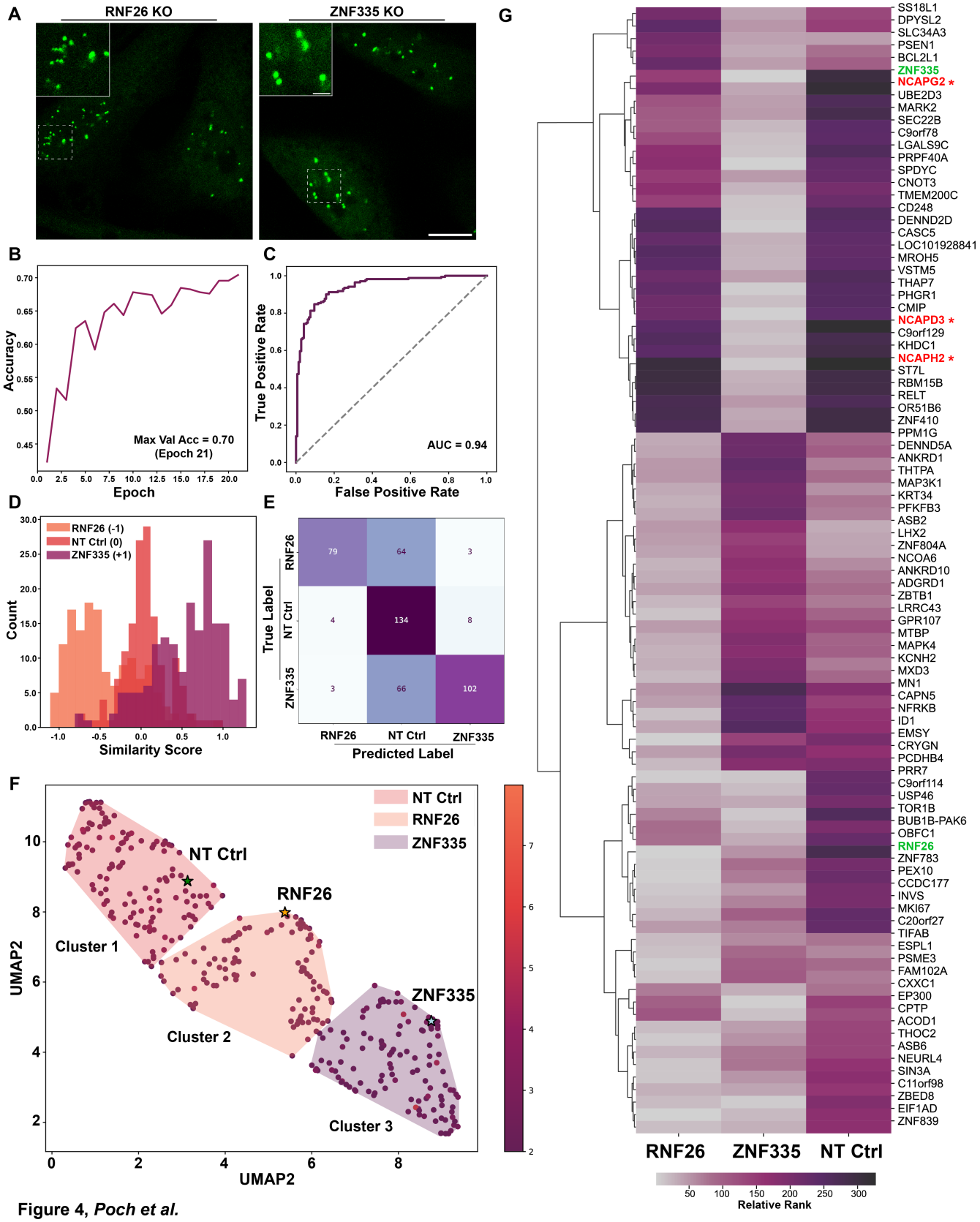

Figure 4, *Poch et al.*

313

Fig. 4. Deep learning model differentiates and clusters distinct condensate phenotypes. (**A**) RNF26 and ZNF335 have distinct MLF2-GFP condensate phenotypes. Representative confocal imaging of RNF26 KO (left) or ZNF335 KO (right) HeLa cells expressing MLF2-GFP (green). Scale bars, 10 µm. (**B**) Accuracy curve of deep residual convolutional neural network during training. Accuracy reflects the percentage of validation cells whose predicted class (RNF26 KO, ZNF335 KO, or NT control) matches the true label. Peak validation accuracy (0.70) was achieved at epoch 21. (**C**) Receiver operating characteristic curve of optimized model. Distinguishes true positives (y-axis) vs false positives (x-axis) with a calculated AUC of 0.94. (**D**) Correlation histogram indicating the similarity for each trained cell (x-axis) to one of three conditions (-1 = RNF26 KO, 0 = NT Ctrl, +1 = ZNF335 KO). Cells are colored by their true knockout condition (light orange = RNF26 KO; red = NT Ctrl; purple = ZNF335 KO). (**E**) Confusion matrix depicting the predicted label (x-axis) of each cell in the validation set compared to its true label (y-axis), based on rounding its similarity score to the nearest integer. (**F**) Top scoring CRISPR KO genes cluster into three distinct groups. Gene-level centroid embeddings of the top 358 genes (Z-Score > 2) were extracted before UMAP dimensionality reduction. K-means clustering partitions genes into three phenotypic clusters: cluster 1 (NT-like), cluster 2 (RNF26-like), and cluster 3 (ZNF335-like). (**G**) Hierarchical clustering heatmap showing phenotypic similarity of the top 50 gene knockouts to RNF26 or ZNF335 KO anchors. Lighter colors indicate greater similarity to the corresponding anchor (RNF26 KO, left; ZNF335 KO, middle; NT control, right). The RNF26 and ZNF335 KO conditions are labeled in green, while the non-SMC condensin II subunits are labeled in red with an asterisk.

We clustered gene-level similarity profiles using hierarchical clustering and constructed a heatmap of the top 50 genes sharing high phenotypic similarity to RNF26 KO and ZNF335 KO conditions (Figure 4G). Of interest, the condensate phenotype that arose from knocking out any of the non-SMC subunits of the condensin II complex (NCAPG2, NCAPD3, and NCAPH2)(*71*) ranked highly phenotypically similar to the ZNF335 KO condition (Figure 4G); knocking out the condensin II SMC2 subunit was also ranked more similar to the ZNF335 KO, albeit further down on the ranked list. Furthermore, NCAPG2 was the only gene to share a primary pairwise clade with ZNF335 (Figure 4G), indicating high phenotypic similarity.

RNF26 and ZNF335 Knockouts Exhibit Distinct Phenotypes but Share Accumulation of K48-Ubiquitin

To confirm our deep learning results suggesting that the condensates that arise from knocking out RNF26 and ZNF335 are distinct, we used confocal microscopy to determine if the condensates accumulate polyubiquitinylated proteins. Condensates in both RNF26 KO and ZNF335 KO cells displayed co-localization with K48-Ub using immunofluorescence imaging (Figure 5A). Next, we assessed whether both condensate conditions co-localize with nucleoporins by performing immunofluorescence staining using α-mAb414, which detects nucleoporins associated with the NPC. Similar to condensates observed in torsin-deficient cells with stalled NPCs (Figure 5B), both RNF26 KO, and to a lesser extent, ZNF335 KO cells demonstrated MLF2 co-localization with nucleoporins (Figure 5B).(*72*) Supporting the prediction that RNF26 and ZNF335 KO conditions promote different condensate phenotypes, confocal Z-stack analysis revealed distinct spatial-localization of condensates (Figure 5C). Cells deficient in RNF26 had

338 MLF2-GFP foci overlapping with nucleoporin staining (α-mAb414) that resembled invag-
339 inations of the nuclear membrane (Figure 5C). Consistent with immunofluorescence
340 data, TEM imaging demonstrated RNF26 KO cells contain condensates between the
341 nuclear envelope (Figure 5D, shaded yellow), with pronounced nuclear envelope invagi-
342 nations reminiscent of nucleoplasmic reticulum formation (Figure 5D).(*73*, *74*) While we
343 observed several nuclear envelope invaginations contiguous with the inner nuclear
344 membrane (INM), further electron tomographic studies are needed to determine if all in-
345 vaginations are contiguous with the nuclear envelope. As 1,6-hexanediol has been re-
346 ported to disrupt liquid-liquid phase separation events,(*75*) we treated RNF26, ZNF335,
347 or TOR1A/B depleted cells with 5% 1,6-hexanediol for five minutes. We found 1,6-hex-
348 anediol, but neither DMSO nor 2,5-hexanediol, dissolved all three types of condensates
349 (Figure S5). Together, these results indicate that both RNF26 and ZNF335 have a role
350 in safeguarding against aberrant condensate formation within the nucleus.(*21*, *30*)

351 **Condensates Induced from Genetic KO Conditions Are Modulated by the Chemi-**
352 **cal Hit Pyrithione Zinc**

353 Finally, we tested whether compounds identified in our chemical screen could
354 modulate the condensate phenotypes arising from the genetic perturbations, including
355 those linked to microcephaly. Due to the ability of PZ to potently modulate MLF2-GFP
356 condensate markers in torsin-deficient cells (Figures 2D, 2E), we tested whether it could
357 similarly modulate nuclear envelope condensates in WT HeLa cells with either RNF26
358 or ZNF335 knocked out. In both KO conditions, treatment with 2 μM PZ significantly re-
359 duced accumulation of nuclear MLF2-GFP condensate foci (Figures 5E, S6). Thus,
360 FDA-approved drugs, including PZ, can be repurposed to modulate multiple aberrant
361 phase transitions across phenotypically distinct genetic perturbations.
362
363
364
365

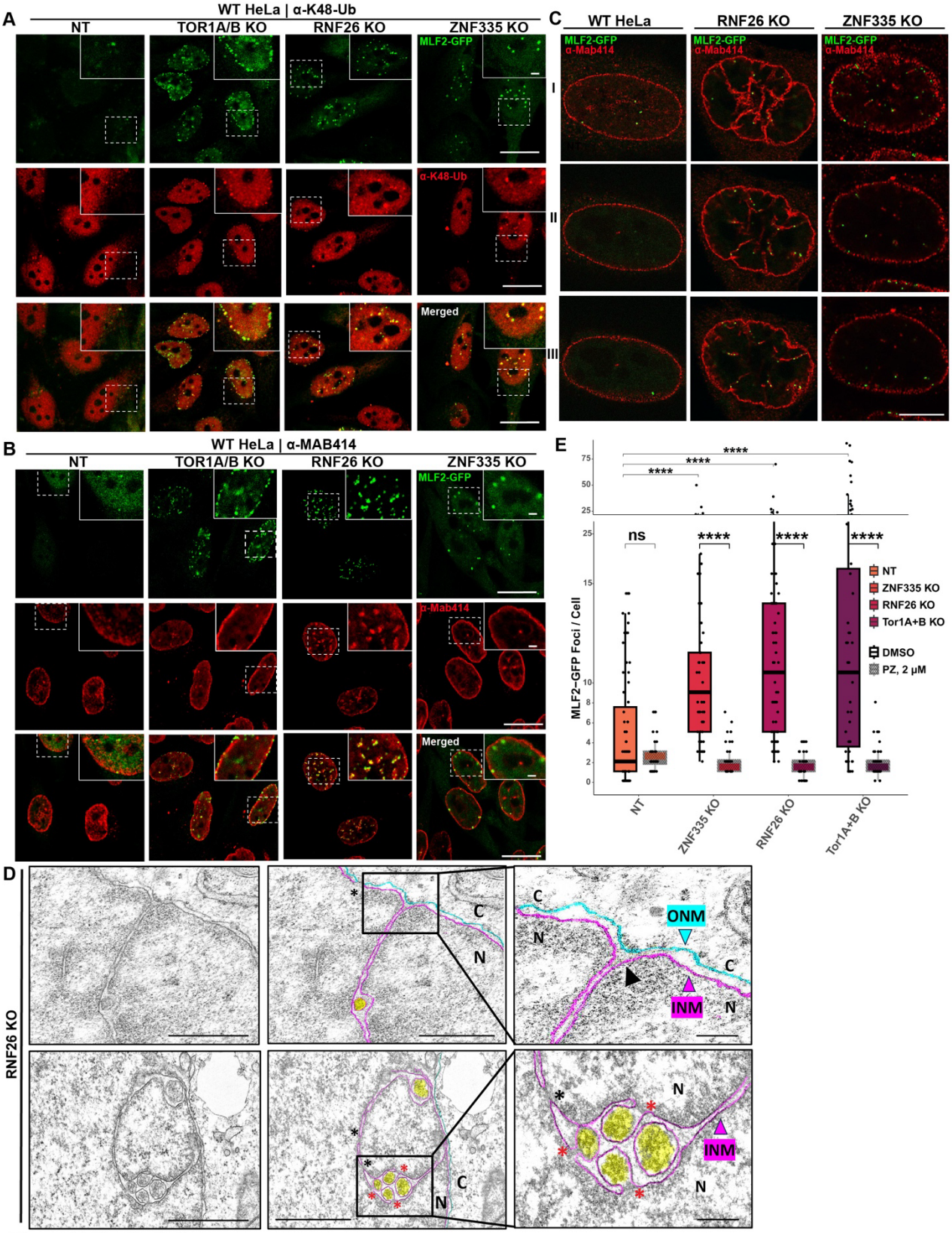

Figure 5, *Poch et al.*

366

Fig. 5. Knockout of RNF26 or ZNF335 induces proteotoxic nuclear condensate formation. (**A**) K48-ubiquitinylated proteins accumulate within nuclear condensates. Immunofluorescence confocal images comparing co-localization of MLF2-GFP (green) with α-K48-Ub staining (red) in Cas9-integrated HeLa cells transfected with either (from left to right) NT, TOR1A/B, RNF26 or ZNF335 sgRNAs. (**B**) Condensates of RNF26 KO cells co-localize with NPC markers. Both RNF26 and TOR1A/B KO conditions demonstrated strong co-localization (yellow) of MLF2-GFP (green) with α-Mab414 staining (red). (**C**) Z-stack confocal projections of cells treated with the same conditions as (B). (**D**) TEM images of RNF26 KO cells depicting pronounced nuclear invaginations. Condensates are shaded yellow, while the INM and ONM are outlined in magenta and cyan, respectively. Black arrows draw attention to the interconnectedness of finger-like nuclear protrusions with INM; black asterisks indicate putative NPCs; red asterisks indicate omega-shaped herniations N, nucleus; C, cytoplasm. (**E**) Quantification of MLF2-GFP foci in CRISPR KO cells after PZ treatment. HeLa expressing MLF2-GFP (green) were treated with NT, Tor1A/B, RNF26, or ZNF335 sgRNAs along with either 2 μM PZ or DMSO for 3 h. KO conditions treated with DMSO are shown on the left (black outline) while those treated with PZ are shown to the right (gray outline with shaded boxes). Asterisks indicate Bonferroni-corrected p-values by Mann-Whitney U testing (****: p <= 0.0001). IF scale bars at 10 μm (main) and 2 μm (inset); EM scale bars at 250 nm (main) and 100 nm (inset).

DISCUSSION

This study establishes a versatile high-throughput platform optimized for identifying chemical and genetic modulators of aberrant condensates that are increasingly recognized as contributors to neurological disorders.(*12*, *14*, *76*) Starting with our established pathological condensate readout using a DYT1 dystonia model,(*21*) we validate our platform through the identification of FDA-approved drugs capable of tuning aberrant phase transitions across genetic knockout conditions. Combining a genome-wide CRISPR KO screen with machine learning approaches, we identified key genes required to protect against distinct aberrant nuclear condensates. Notably, mutations in several of these genes have been identified in patients with microcephaly, dystonia, and related neurodevelopmental disorders. This highlights previously unappreciated commonalities in the cellular pathology underlying these conditions and suggests a potential novel role of aberrant nuclear condensate formation in microcephaly. Thus, our study advances a set of high-priority genetic candidates linking atypical phase transitions to neurodevelopmental disorders and provides insight into potential therapeutic targets.

Exploiting the known proclivity of MLF2 toward immersing into aberrant condensates,(*1*) we here establish MLF2 as a robust and disease-agnostic marker for aberrant condensates to perform both pharmacological and genomic screening. Not only is MLF2 a primary biomarker for DYT1 dystonia-associated condensates, but we discovered MLF2 is also recruited to stress granules associated with ALS/FTD (Figure 1B). Thus, our high-content imaging platform and *CondenScreen* analysis pipeline can be readily adapted to support a range of condensate screening campaigns, even when the affected disease allele product is not itself a condensate constituent. This is exemplified by TOR1A/B or ZNF335 KO conditions, which induce aberrant condensates as

392 indicated by MLF2 via indirect mechanisms. This represents a key strength, since con-
393 densates are typically studied using overexpression of fluorophore-derivatized disease
394 alleles, which can disrupt physiologically relevant equilibria. Moreover, the sensitivity of
395 MLF2 enables visualization of condensates from single torsin deletions that remain un-
396 detectable by conventional markers such as EM or K48-Ub staining.(*30*)

397 Inspired by recent condensate-modifying (c-mod) campaigns,(*36*, *37*, *77*, *78*) we
398 screened drugs that could modulate condensates either directly or indirectly. Several of
399 our identified drug hits caused robust re-localization of both our MLF2-GFP biomarker
400 and endogenous (K48-Ub) nuclear condensate markers (Figures 2E, 2F), without signif-
401 icantly altering reporter expression. Additionally, we found that pyrithione zinc acts
402 through a divalent zinc-dependent mechanism (Figure S2A-S2F), further underscoring
403 previous research on the importance of tight intracellular zinc regulation in phase transi-
404 tions.(*79*, *80*)

405 To determine the ability to repurpose our chemical hits for other condensato-
406 pathies, we compared our top 20 drug hits with two recent screens that identified modu-
407 lators of stress granules.(*36*, *37*) A total of 50% of our drug hits (10 out of 20), including
408 PZ, celastrol, and disulfiram, were all identified between the top 100 hits of these stress
409 granule screens. Moreover, our hits penfluridol, thimerosal, and disulfiram were identi-
410 fied across all three screens.(*36*, *37*) The ability of identical chemical matter modulating
411 disparate condensates in distinct disease phenotypes underscores the potential of
412 shared etiological principles across aberrant phase transitions.

413 Two major obstacles toward developing therapeutics for condensatopathies are
414 the lack of well-defined, druggable condensate targets and the incomplete understand-
415 ing of the mechanisms driving condensate formation. To address these limitations, we
416 performed a genome-wide CRISPR KO screen in WT HeLa cells expressing MLF2-GFP
417 to identify genes whose loss promotes nuclear condensate formation. Using a deep re-
418 sidual convolutional neural network to probe subtle changes to both nuclear and con-
419 densate morphology of the top genetic screen hits, we identified two distinct phenotypic
420 clusters around either the RNF26 KO or ZNF335 KO condition (Figure 4F). To distin-
421 guish these clusters, we performed downstream confocal and electron microscopy on
422 the RNF26 KO and ZNF335 KO conditions. Previous work has implicated RNF26 in ER

423 positioning, endosome retention, and Sec62-mediated recovERphagy.(*54*, *55*, *81*) While
424 condensates from both knockout conditions accumulated K48-Ub (Figure 5A), indicating
425 cellular protein quality control dysregulation, cells deficient in RNF26 demonstrated par-
426 ticularly strong condensate colocalization with nucleoporins (Figure 5B, 5C) and abnor-
427 mal nuclear envelope invaginations (Figures 5C, 5D). These results intersect with previ-
428 ous DYT1 dystonia research demonstrating that torsin deficiency promotes nuclear con-
429 densates that co-localize with nucleoporins,(*1*, *21*) and abnormal nuclear envelope in-
430 vaginations.(*61*)

431 Depletion of ZNF335, a C2H2-type zinc finger and component of the trithorax
432 H3K4-methylation complex,(*47*) resulted in the most pronounced condensate pheno-
433 type, with the average number of condensate foci exceeding that found in either the
434 RNF26 or TOR1A KO conditions (Figure S3D). Mutations or aberrant splice variants of
435 ZNF335 results in severe microcephaly and, often, reduced life expectancy.(*47*–*52*)
436 Moreover, ZNF335 has been reported to regulate the REST/NRSF complex with its
437 knockdown inhibiting neuronal differentiation (*47*)

438 In addition to ZNF335, pathogenic allelic variants of 15 out of the top 50 gene
439 hits are associated with neurodevelopmental disorders according to human test results
440 from OMIM(*66*) and the NIH Genetic Testing Registry.(*67*) (Figure 6A). Strikingly, 7/15
441 of these genes (ZNF335 [MCPH10], CIT [MCPH17], NCAPD3 [MCPH22], CDK5,
442 SIN3A, SHH, and SPDL1) are linked to microcephaly, holoprosencephaly, or lissen-
443 cephaly (OMIM, GTR), along with additional findings implicating NCAPG2, NCAPH2,
444 and EP300 with microcephaly.(*47*, *63*, *82*–*85*) Our high-content screening and deep
445 learning approaches uncovered that knocking out either the established microcephaly
446 gene CASC5 (MCPH4),(*86*) or CNTROB, a putative microcephaly gene not yet associ-
447 ated with humans,(*65*, *87*) produces nuclear condensates that are highly phenotypically
448 similar to those observed in ZNF335 KO cells. Given that whole-exome sequencing fails
449 to identify causal disease-causing mutations in more than 70% of microcephaly pa-
450 tients,(*88*) our curated list of MLF2 condensate-associated genes provides a valuable
451 resource for prioritizing candidates in future patient genetic studies.

452 Several of the microcephaly-associated gene hits identified in our screen are also
453 involved in chromosomal condensation, including several independently identified

454 subunits of the condensin II complex. Microcephaly has previously been labeled as a
455 "condensinopathy",(*63*) not to be confused with condensatopathies, reflecting prior work
456 that established a relationship between chromosomal condensation defects and re-
457 duced cerebral cortex size.(*63*) Our deep learning model uncovered a unusually similar
458 nuclear condensate phenotype between cells lacking ZNF335 and those deficient in all
459 non-SMC subunits of the condensin II complex, NCAPG2, NCAPH2, and NCAPD3.
460 Since nuclear MLF2-containing condensates accumulate upon disruption of four out of
461 five condensin II subunits (Figures 3B, 4G), it is reasonable to suggest a potential func-
462 tional link between MLF2 and the condensin II machinery. Interestingly, MLF2 was re-
463 cently suggested to act as a modulator of SWI/SNF (BAF) complex assembly.(*89*)
464 Therefore, our work may inspire future functional studies at the intersection of MLF2,
465 condensin biology, and phase separation.
466 While our study highlights several chemical and genetic modulators of disease-
467 linked condensates, opportunities exist for further refinement. Pyrithione zinc markedly
468 resolved the proteotoxic condensate cargo in torsin-deficient cells, while complete reso-
469 lution of the underlying condensates was not observed at the tested concentration (Fig-
470 ure S1H). Extended incubation times may result in condensate resolution, but increased
471 cytotoxicity limited this approach. Similarly, while our genetic screen uncovered an en-
472 richment of neurodevelopmentally-linked genes, highly redundant pathways, or those
473 strictly requiring a neuronal context, are likely to be underrepresented and we filtered
474 out conditions that resulted in severe cytotoxicity. Future large-scale screening of opti-
475 mized compound libraries and testing genetic redundancy through combinatorial pertur-
476 bations will likely identify additional condensate modulators. Finally, while we discov-
477 ered MLF2 as a hallmark of condensates across multiple disease contexts, including
478 nuclear envelope condensates in DYT1 dystonia,(*21*) cytoplasmic stress granules in
479 ALS and FTD, and nucleoplasmic condensates linked to microcephaly genes, further
480 studies are warranted to distinguish reversible from aggregated states and to determine
481 whether additional disease contexts involve condensates enriched for MLF2.In conclu-
482 sion, our study lays the foundation for exploring what cellular role biomolecular conden-
483 sates have in microcephaly or related neurodevelopmental disorders (Figure 6B). Fur-
484 thermore, we overcome the scarcity of generic phase-separation markers by

485 establishing MLF2 as a versatile and disease-agnostic condensate biomarker and re-
486 port several chemical modulators of aberrant phase transitions. This reinforces similar
487 chemical-modulator screening as a general strategy to identify therapeutic interventions
488 across condensatopathies. Our screen and analysis pipeline are readily adaptable to
489 explore diverse pathological phase transitions, including stress granules or other dis-
490 ease-relevant condensates. This study uncovers a connection between aberrant con-
491 densates and neurodevelopmental disorders, inspiring pharmacological strategies to
492 correct aberrant phase transitions.

A

Gene	Z-Score	Disorder	Gene	Z-Score	Disorder
RNF26	7.99	—	LGALS9C	3.29	—
WAC	6.31	Desanto-Shinawi Syndrome	NAPRT	3.27	—
ZNF335	5.46	Microcephaly	MAP3K1	3.25	46, XY sex reversal 6
NCAPG2	5.33	Khan-Khan-Katsanis Syndrome	C9orf114	3.23	Neurodevelopmental disorder with poor growth, seizures, and brain abnormalities
NCAPH2	4.88	Cardioencephalomyopathy, Myopia	ESPL1	3.21	—
ECT2	4.58	—	CAPN5	3.18	Vitreoretinopathy, neovascular inflammatory
ASRGL1	4.22	—	TOR1B	3.18	—
SETD1B	4.16	Intellectual Developmental Disorder with Seizures and Language Delay	POLD4	3.16	—
ASB6	4.15	—	EP300	3.16	Rubinstein-Taybi Syndrome 2, Menke-Hennekam Syndrome 2, Colorectal Cancer
FOXS1	4.08	—	PRPF40A	3.13	—
SIN3A	3.94	Witteveen-Kolk Syndrome (WITKOS)	CCDC157	3.13	—
BRSK2	3.93	Idiopathic Pulmonary Fibrosis Susceptibility*	ZBED8	3.11	—
EIF4A3	3.78	Robin Sequence with Cleft Mandible and Limb Anomalies	TMEM25	3.11	—
SPDL1	3.77	Severe Primary Microcephaly, Neonatal Death	GJA4	3.08	—
ZNF783	3.66	—	CNOT3	3.07	Intellectual developmental disorder with speech delay, autism, dysmorphic facies
MKI67	3.64	Long QT Syndrome	CD248	3.04	
SLC29A2	3.6	Bipolar disorder susceptibility*	ACOD1	3.02	Lymphatic Malformation
VEGFC	3.57	Lymphatic malformation	BTBD18	3.01	—
ACY3	3.51	—	CIT	2.99	Primary Microcephaly
CDK5	3.42	Isolated Microcephaly, Nonsyndromic Microcephaly, Lissencephaly	PHF23	2.99	—
NCAPD3	3.4	Primary Microcephaly	AK2	2.98	Reticular Dysgenesis
PTGES	3.35	—	TCEA2	2.98	—
C9orf129	3.32	—	C5orf60	2.95	—
TUBB4B	3.29	Leber Congenital Amaurosis with Early-Onset Deafness	SHH	2.94	Holoprosencephaly, Schizencephaly, Solitary Median Maxillary Central Incisor Syndrome, Microphthalmia
ACTL7A	3.29	Spermatogenic failure; breast cancer susceptibility*	LDLRAD1	2.93	—
Microcephaly/Congenital Brain Malformations			Other Neurodevelopmental Disorders		

B

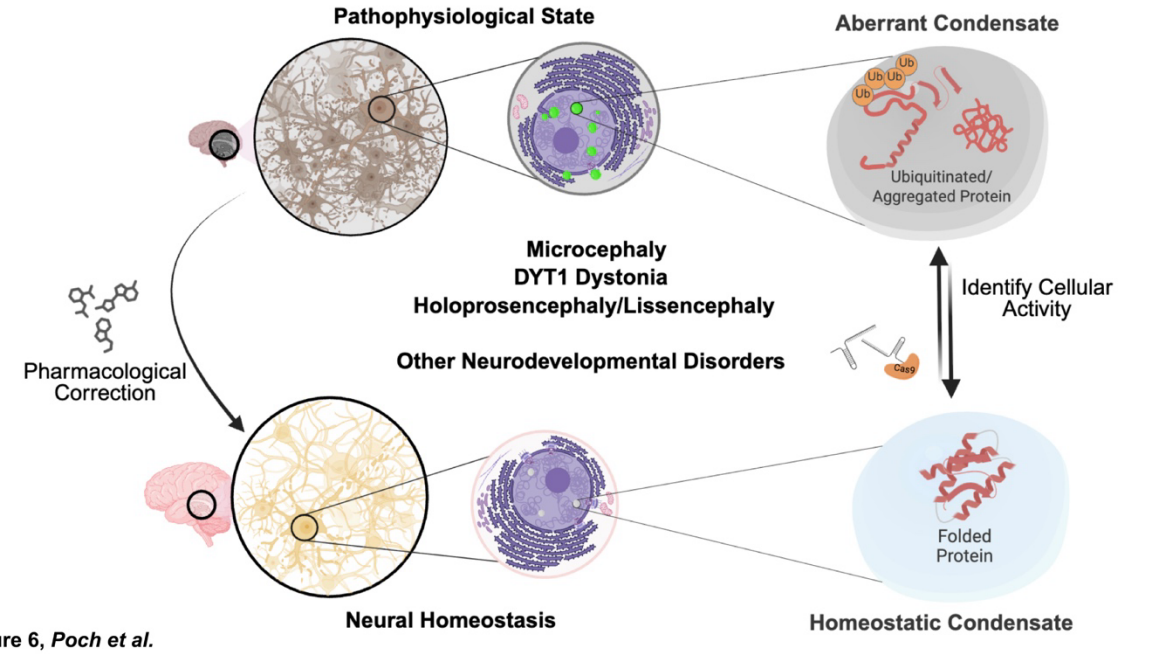

Figure 6, *Poch et al.*

Fig. 6. Targeting aberrant condensates as a therapeutic approach for neurodevelopmental disorders. (**A**) CRISPR KO hits enriched for genes implicated in neurodevelopmental disorders. Top 50 candidate genes (left columns), their significance score (Z-Score, middle columns), and associated conditions (right columns) are color-coded based on disorders. Rows labeled purple indicate an established link to microcephaly/congenital brain malformations, with dark gray indicating additional neurodevelopmental disorders. Genes denoted with an asterisk indicate evidence from genome-wide association studies. Gene relevance to disease was determined based on allelic variants obtained from OMIM and GTR. (**B**) Chemical and genetic modulation of aberrant condensates in neurodevelopmental disorders. Genes involved in the etiology of primary microcephaly and encephalopathies were highly enriched in our screen. Condensates can be modulated using small molecules or genetic perturbations. Targeting condensates could potentially lead to preserving neuronal homeostasis. Arrows indicate possible intervention points and potential outcomes.

495

496 **SUPPLEMENTARY FIGURES**

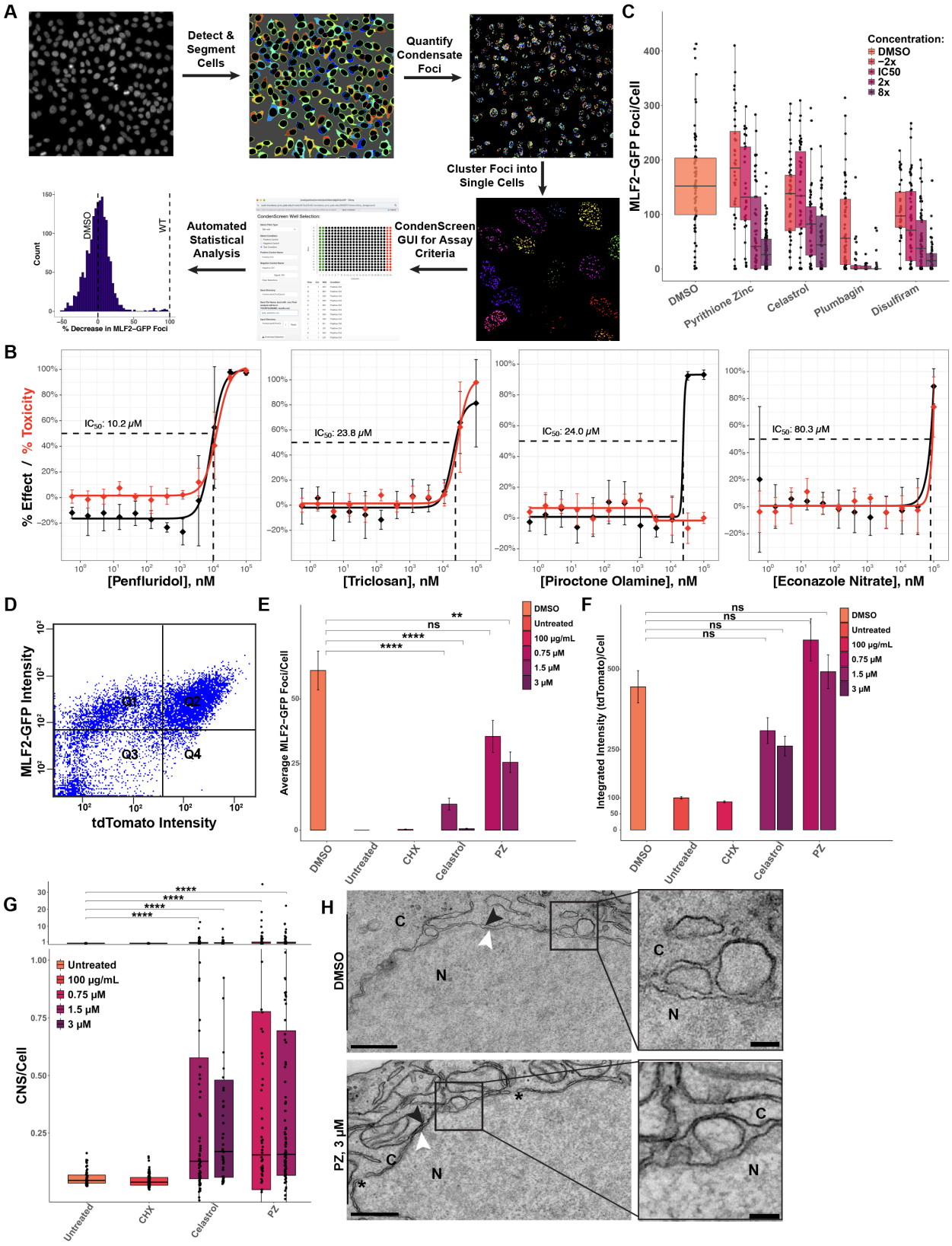

Figure S1, *Poch et al.*

Fig. S1. Validation of chemical hits from the Pharmakon 1760 screen. (**A**) Schematic of the computational *CondenScreen* Initial cell segmentation and foci detection uses custom *CellProfiler* script and while downstream statistical analysis is completed with *R*. (**B**) Calculated IC_{50} values of additional compounds with dose-dependent relationship (four-parameter-log-logistic). The black and red curves represent % effect and % toxicity, respectively, relative to the DMSO control. (**C**) Drug treatment decreases nuclear MLF2-GFP accumulation. Torsin-deficient cells expressing MLF2-GFP (8 h) were treated at four different concentrations (2x below IC_{50}, at IC_{50}, 2x above IC_{50}, and 8x above IC_{50}) for 3 h with either PZ, celastrol, plumbagin, or disulfiram. Nuclear MLF2-GFP foci were compared between treated and DMSO (orange boxplot) conditions. Each dot represents the number of MLF2-GFP foci for a single cell, while the horizontal lines indicate the median foci/condition. (**D**) Distribution of FACS results for 4TorKO MLF2-GFP-P2A-tdTomato HeLa cell sorting. Cells within quadrant 2 were sorted. Quantification of the number of MLF2-GFP foci/cell (**E**) and tdTomato integrated intensity/cell (**F**). Cells (4TorKO) were treated with either only Dox (untreated) or Dox with a combination of DMSO, CHX, celastrol, or PZ. (**G**) Average CNS-value/cell across treatment conditions normalized to the DMSO control. Asterisks indicate Bonferroni-corrected p-values determined by Mann-Whitney U testing (**: p <= 0.01; ****: p <= 0.0001). (**H**) PZ modulates condensate cargo but lacks full nuclear envelope condensate resolution. TEM images of 4TorKO HeLa cells treated with either 0.1% DMSO or 2 μM PZ for 3 h. Insets show nuclear envelope condensates between the inner and outer nuclear envelopes. Scale bars at 500 nm; inset scale bars at 100 nm. Asterisk, putative nuclear pores; black arrowhead, outer nuclear envelope; white arrow, inner nuclear envelope; C, cytoplasm; N, nucleus.

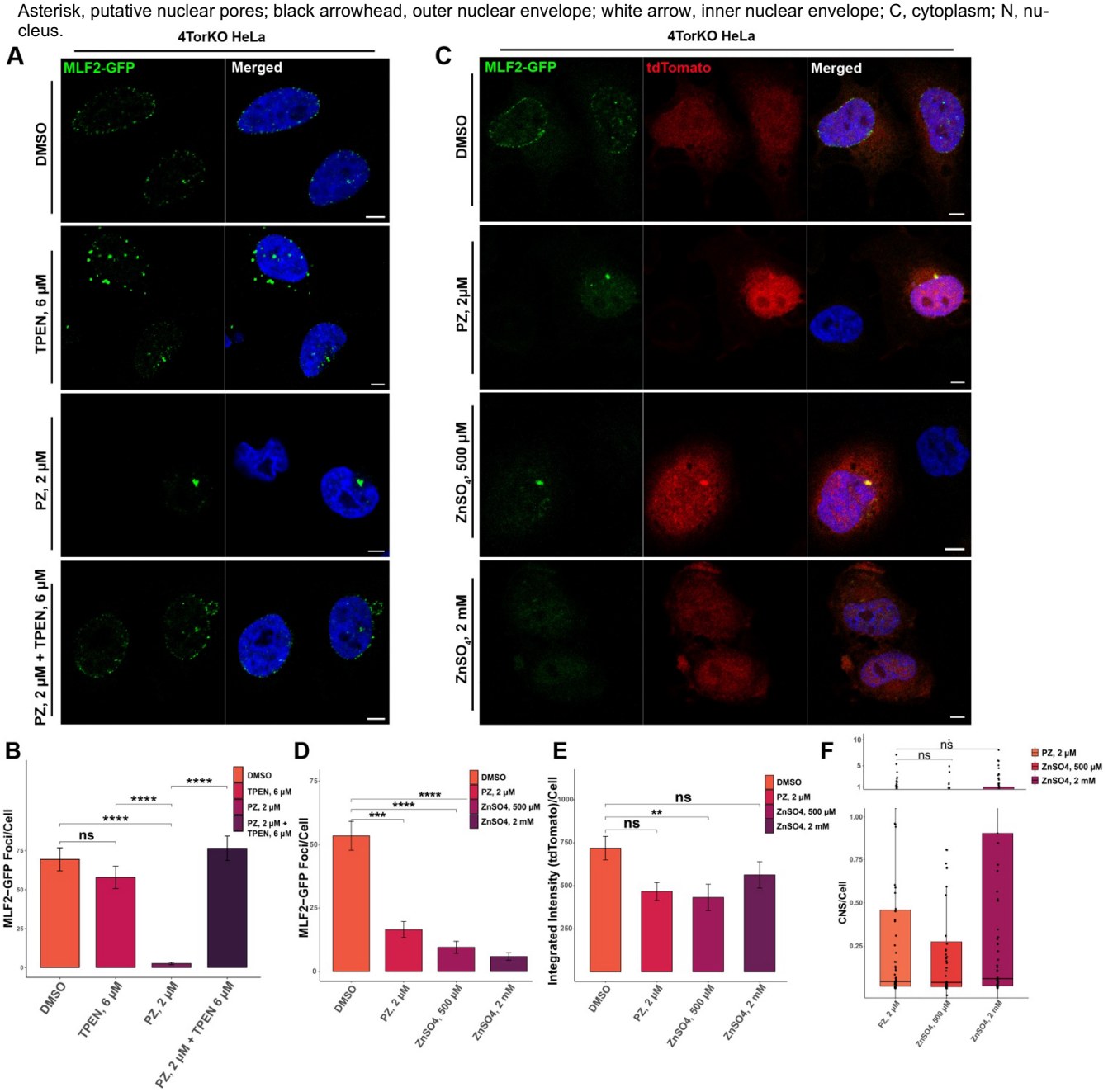

Figure S2, *Poch et al.*

497

Fig. S2. Zinc is sufficient for modulating nuclear condensates in 4TorKO cells. (**A**) TPEN negates the effect of PZ treatment. Confocal images of 4TorKO cells expressing MLF2-GFP (green) and stained with Hoechst (nucleus, blue). Cells were treated with either DMSO, 6 μM TPEN, 2 μM PZ, or 6 μM TPEN + 2 μM PZ (3 h). (**B**) Barplot quantification of (A) showing the average number of MLF2-GFP foci per cell for each condition. (**C**) Zinc sulfate is sufficient to modulate nuclear condensates. Dox-induced 4TorKO were treated with either DMSO, 2 μM PZ, 500 μM ZnSO4, or 2 mM ZnSO4. Representative confocal images demonstrating distribution of MLF2-GFP, tdTomato, and DAPI are shown by the green, red, and blue channels, respectively. The corresponding barplot quantifications of each condition are shown representing (**D**) the average number of MLF2-GFP foci per cell (**E**) the average tdTomato integrated intensity per cell, and (**F**) the average CNS-score per condition. Scale bars, 5 μm. Asterisks indicate Bonferroni corrected p-values using a Mann-Whitney U test (*: p <= 0.05; **: p <= 0.01; ***: p <= 0.001; ****: p <= 0.0001).

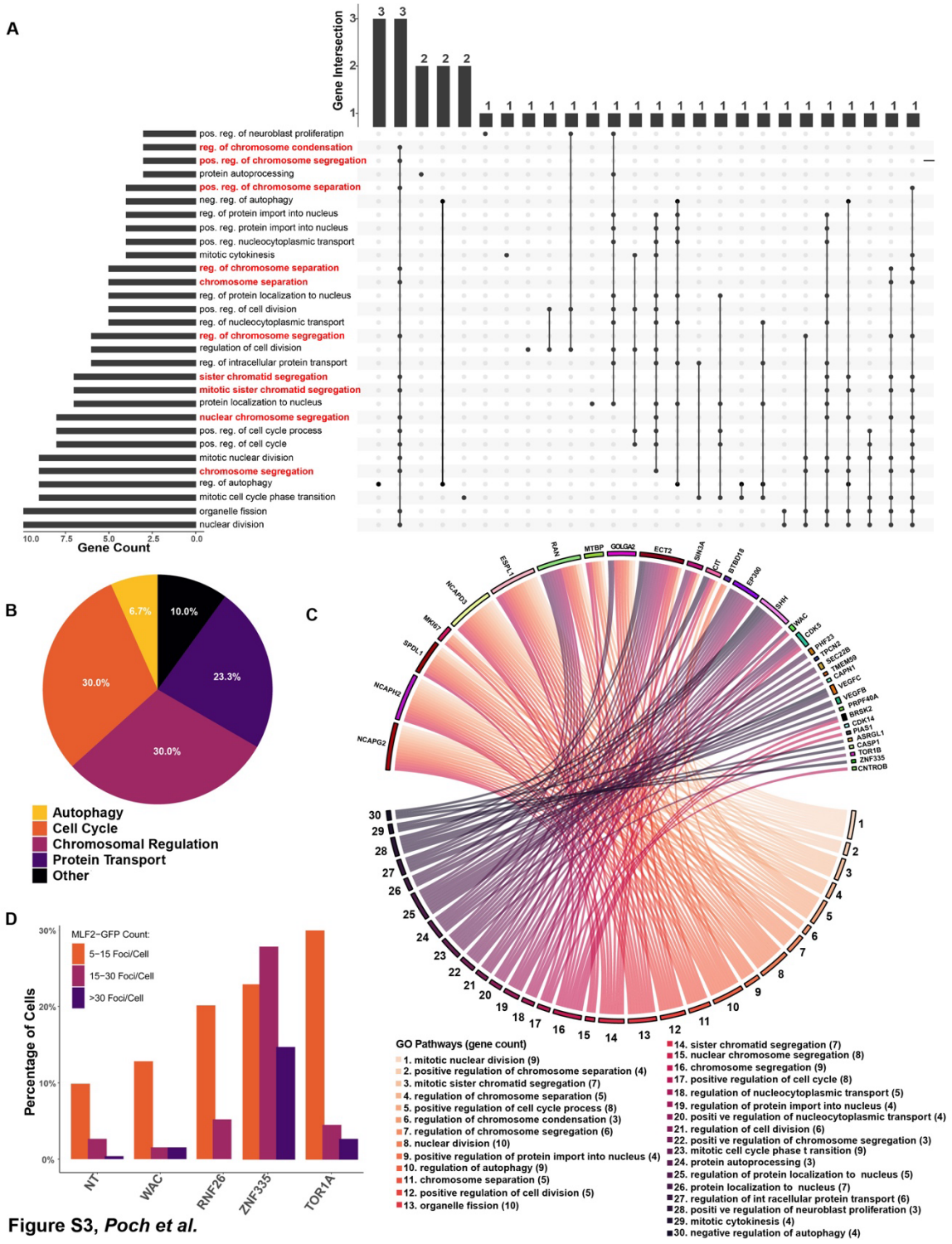

Figure S3, *Poch et al.*

498

Fig. S3. Validation and enriched gene pathways of CRISPR KO hits that promote condensate formation. (A) Enrichment of pathways related to chromosomal and chromatid regulation. Upset plot depicting relationships between the top 30 GO-enriched biological process pathways. Horizontal bars (left) indicate number of genes for each GO pathway, with terms ranked by gene count. Vertical bars (top) indicate size of each intersecting gene set of genes sharing multiple GO terms. The dot matrix below indicates shared genes among the corresponding GO terms (black dots connected by black line); single dots represent GO terms with unique gene sets. Red dots highlight pathways related to chromosomal or chromatid-related activities, (**B**) Manual clustering of GO biological pathways from the genome-wide KO screen in WT HeLa. The top 30 enriched pathways for the 100 highest-scoring candidate genes were manually clustered within five categories. Percentage represents the fraction of GO pathways within each cluster. (**C**) Chord diagram mapping each gene to its specific GO terms. Each gene (outer labels, top) is linked to its corresponding GO terms (bottom labels), with colored ribbons denoting relationships. (**D**) Binned quantification of the MLF2-GFP foci count after CRISPR KO of select candidate genes. RNF26 KO, ZNF335 KO, and WAC KO HeLa cells were compared against NT control. Bars indicate percent of cells with 5-15 foci MLF2-GFP foci/cell (orange), 15-30 foci (violet), or >30 foci (dark purple).

499

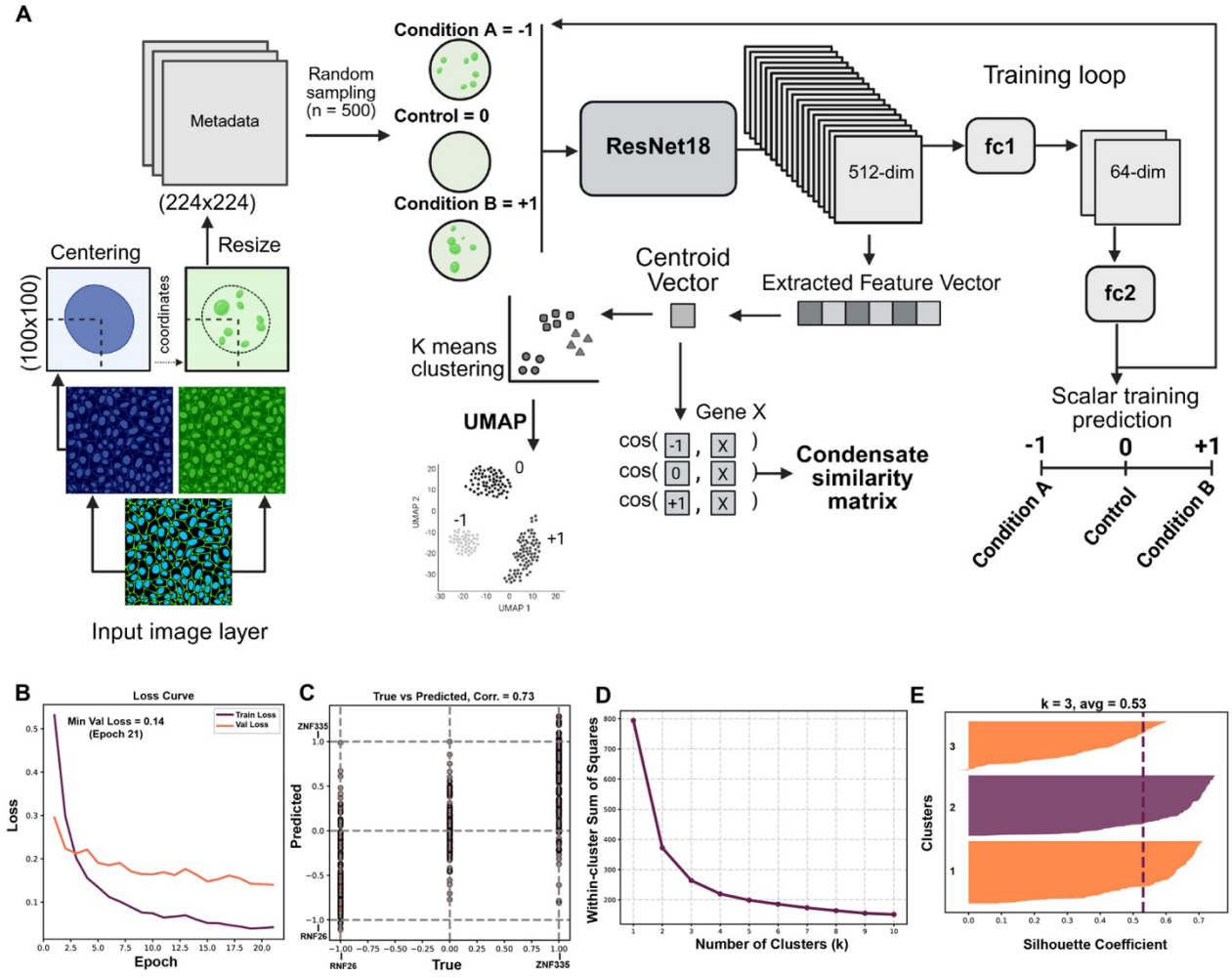

Figure S4, ***Poch et al.***

Fig. S4. Validation of deep learning model to detect subtle variations in condensate phenotypes. (**A**) Schematic workflow of the deep learning-based condensate phenotype analysis pipeline. The model was trained on raw images from the CRISPR screen, which were resized to single-cell patches and labeled. A residual convolutional neural network (ResNet18) was optimized to distinguish variations in condensate morphology. Centroid feature vectors from the top CRISPR hits (Z-score > 2) were extracted and mapped using UMAP and k-means clustering. (**B**) Training and validation loss curves based on Huber loss (nn.SmoothL1Loss). Loss of each epoch for both the training set (purple) and validation set (orange) are reported with a minimum loss (0.14) occurring on epoch 21. (**C**) Correlation scatter plot of validation set. Each dot represents the similarity score (-1 = RNF26 KO; 0 = NT Ctrl; +1 = ZNF335 KO) of a single cell from the validation set. The predicted value (y-axis) is compared against the true value (x-axis), with a Pearson correlation score of 0.73. (**D**) Elbow (scree) plot showing the within-cluster sum of squares across cluster numbers (k = 1-10). The inflection at k = 3 indicates the optimal number of clusters. (**E**) Silhouette plot confirming three clusters (k = 3) as an optimal number identified in (D). The average silhouette score (0.53) is indicated by the dashed vertical line.

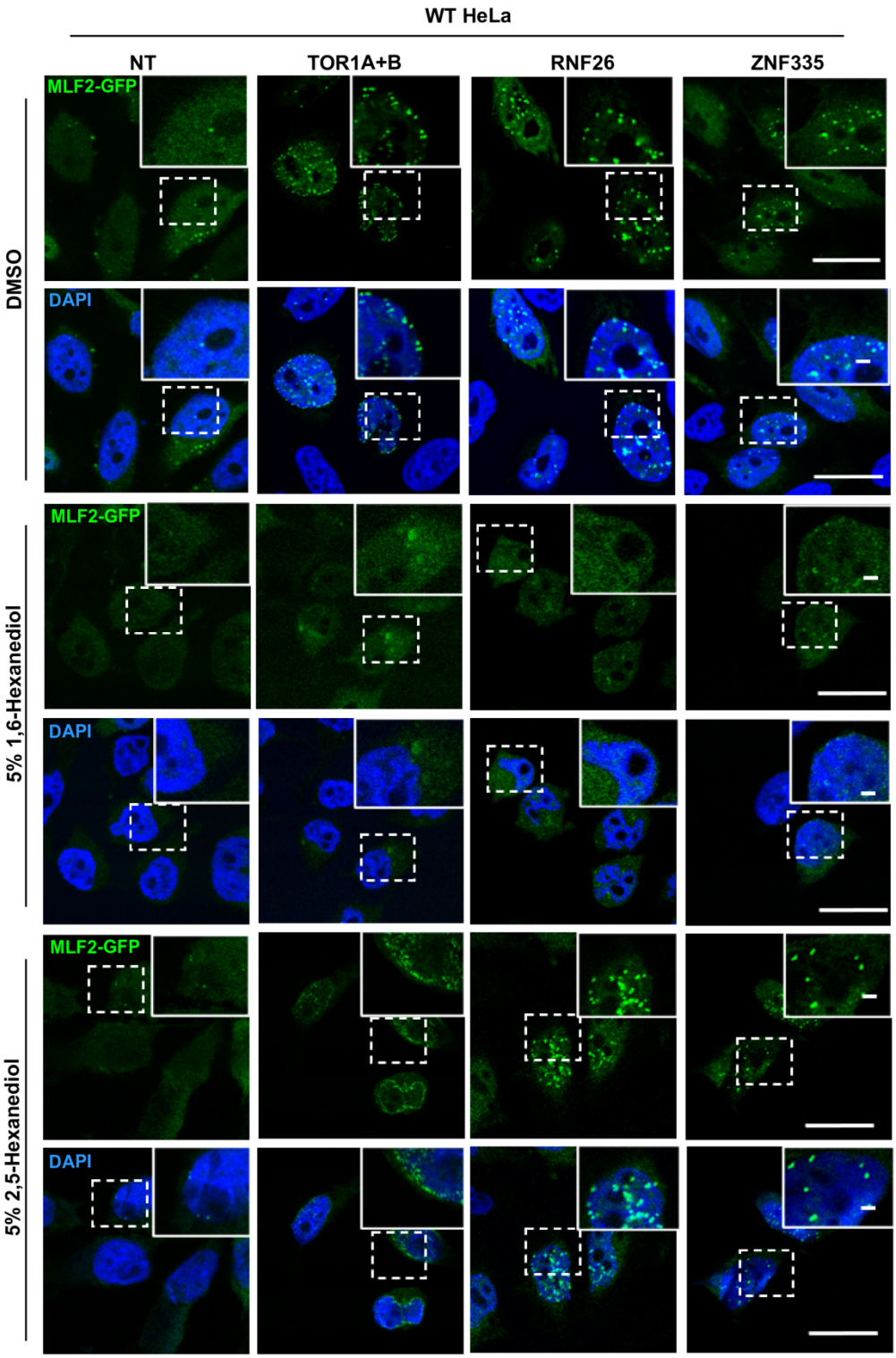

Figure S5, ***Poch et al.***

Fig. S5. Condensates resulting from CRISPR KO can be dissolved by 1,6-Hexanediol. (A) Fluorescent microscopy images of Cas9-integrated WT HeLa cells treated with sgRNAs against either NT control, TOR1A/B, RNF26, or ZNF335 (from left to right) for 72 h. Cells were then exposed to either DMSO (top), 5% 1,6-hexanediol (middle), or 5% 2,5-hexanediol (bottom) for 5 min. WT cells with ZNF335, RNF26, or TOR1A/TOR1B knocked out demonstrated dissolution of condensates marked by MLF2-GFP (green) only after exposure to 5% 1,6-hexanediol (middle), but not DMSO (top) or 5% 2,5-hexanediol (bottom) controls.

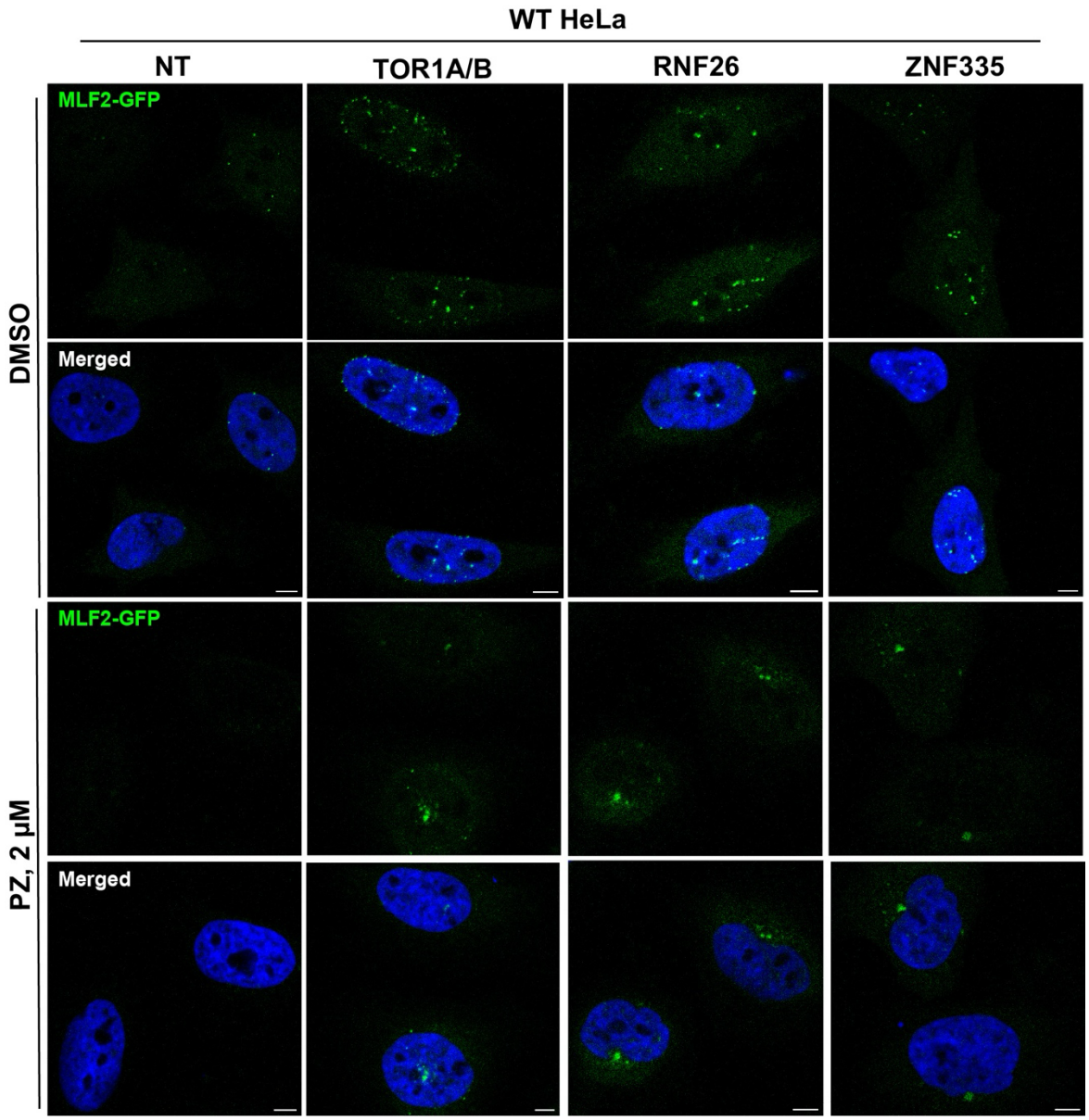

Figure S6, *Poch et al.*

Fig. S6. Pyrithione zinc modulates aberrant nuclear condensates across genetic perturbations. Cas9-integrated WT HeLa cells treated with sgRNAs against either NT control, TOR1A/B, RNF26, or ZNF335 (from left to right) for 72 h. Cells were then induced with MLF2-GFP (6 h Dox) and treated with either 0.1% DMSO (top panel) or 2 μM PZ (bottom panel) for 3 h. Representative confocal images show MLF2-GFP (green) and DAPI (blue). Scale bars, 5 μm.

512 **MATERIALS AND METHODS**

513 **Tissue culture**

514 Adherent cultured cell lines (HeLa and HEK293T) were maintained in Dulbecco's modi-
515 fied Eagle's medium (DMEM) supplemented with 10% (vol/vol) heat-inactivated fetal bo-
516 vine serum (Thermo Fisher Scientific) and 100 units (U) /mL penicillin-streptomycin
517 (Thermo Fisher Scientific). Cells were maintained in either T75 flasks or 10 cm dishes at
518 37℃ (95% humidity; 5% CO_2) and passaged every third day. For splitting, the culture
519 medium was removed, washed once with 1xPBS (Gibco), trypsinized with 0.25% tryp-
520 sin-EDTA (Gibco), and transferred to new dishes (diluted 1:5 or 1:10).

521

522 **Plasmid Constructs**

523 Unless otherwise stated, all plasmid constructs were subcloned into a pcDNA™3.1 (+)
524 Mammalian Expression Vector (Invitrogen, V79020) using standard Gibson Assembly
525 protocols. The cDNA for the MLF2-GFP-P2A-tdTomato construct (MLF2 GenBank ac-
526 cession no. BC000898) was subcloned into the pRetroX-Tight-Pur-GOI (gene of inter-
527 est) vector (Takara Bio) and co-transfected with a pRetroX-Tet-On advanced inducible
528 expression construct (Takara Bio). Viral constructs were co-transfected with both a
529 VSV-G viral envelope plasmid (Addgene #12259) and an MMLV gag/pol viral packaging
530 plasmid.For generation of HeLa cells expressing Cas9-mCherry, an mCherry tag was
531 inserted into the lentiviral plasmid containing Cas9-Flag (Addgene plasmid #52962) im-
532 mediately after the Flag tag into the BamHI cloning site.

533

534 **Viral Production**

535 Lentivirus production was carried out as previously described.(*90*, *91*) Briefly, to gener-
536 ate Cas9-flag lentivirus, low-passage HEK293T cells were plated in a 10 cm dish with a
537 density of 3.5×10^6 cells per dish. Medium was replaced the next day with 6 mL fresh
538 (antibiotic-free) DMEM and transfected with three plasmids encoding: psPAX2
539 (Addgene# 12260) (2 µg), viral envelope protein VSV-G (1 µg), and lenti-Cas9-Flag-
540 mCherry (6 µg) using Lipofectamine 2000 (Thermo Fisher Scientific) at a 3:1 ratio of
541 DNA to Lipofectamine. After 72 h, supernatants containing the virus encoding with
542 Cas9-mCherry were harvested and filtered through a 0.45 µm Steriflip filter unit

543 (MilliporeSigma), aliquoted into screw-cap tubes, and stored at -80°C or used for trans-
544 duction immediately. For the Dox-inducible MLF2-GFP-P2A-tdTomato system, two sep-
545 arate retroviruses were produced using three plasmids encoding: MMLV gag/pol (2 µg),
546 viral envelope protein VSV-G (1 µg), and either pRetroX-Tight-MLF2-GFP-P2A-
547 tdTomato or pRetroX-Tet-On (6 µg) using Lipofectamine 3000 (Thermo Fisher Scientific)
548 at a 2:1.5 ratio supplemented with Lipofectamine P3000 (2.5:1). After 72 h, superna-
549 tants were collected, stored, or used, as described.
550
551 **Generation of stable HeLa cell lines**
552 The detailed protocol regarding the Tor1A/1B/2A/3A KO and integration of a Dox-induci-
553 ble MLF2-GFP reporter in HeLa cells has been extensively described by our lab
554 (Laudermilch et al. 2016; Rampello et al. 2020).(*21*, *30*) The Cas9-flag-mCherry HeLa
555 and Dox-inducible MLF2-GFP-P2A-tdTomato 4TorKO HeLa cell lines were generated by
556 seeding Dox-inducible MLF2-GFP WT HeLa or 4TorKO cells in six-well plates. The next
557 day, media was replaced with DMEM (antibiotic-free) supplemented with 4 µg/mL
558 polybrene (Sigma-Aldrich), and 100 µl of either the Cas9-Flag-mCherry virus (WT HeLa)
559 or both 100 µL of the MLF2-GFP-P2A-tdTomato and pRetro-Tet-On viruses (4TorKO)
560 added dropwise to wells. Medium was replaced 48 hours after transduction (complete
561 DMEM) and grown for another 24 hours. Cells were collected in sorting buffer (ice-cold
562 1xPBS, Gibco) and filtered through a 40 µm nylon mesh filter. Using a BD FACS Aria III,
563 cells were sorted as single cells by selecting those high in either mCherry (WT HeLa) or
564 MLF2-GFP/tdTomato (4TorKO, 16 h after 0.5ug/mL Dox induction) fluorescence inten-
565 sity. A stringent gating strategy was applied to enrich for cells positive for the respective
566 fluorophores while excluding those with extremely high signal intensity, to obtain a pure
567 yet physiologically representative population.
568
569 **High-content phenotypic screening strategy**
570 All high-content screening was performed at Yale's Center for Molecular Discovery
571 (YCMD). The Horizon Discovery Edit-R synthetic sgRNA library was used in the whole
572 genome CRISPR KO screen, targeting 19,195 human genes, and comprising of the
573 Drug Targets library (3,683 genes, GA-004650-E2, Lot 20126), the Druggable Genome

574 library (4,643 genes, GA-004670-E2, Lot 20120), and the Human Genome library
575 (10,869 genes, GA-006500-E2, Lot 20130). Each screening plate included 16 replicates
576 of the negative control (NT sgRNA) and 16 replicates of the positive control (combined
577 Tor1A and Tor1B sgRNAs). To ensure robust KO, a pool of three sgRNAs targeting the
578 same gene at different regions was pre-plated into each arrayed well of 384-well glass-
579 bottom plates (Cellvis, #384-1.5H-N). Addition of sgRNA to screening plates was done
580 using the Labcyte Echo 550 Acoustic Dispenser (Labcyte, now Beckman) at a final con-
581 centration of 50 nM.

582 Reverse transfection was performed using DharmaFECT1 Transfection Reagent
583 (Horizon/Dharmacon), which was diluted at 1:500 in Opti-MEM (Thermo Fisher Scien-
584 tific) and mixed with sgRNA. The mixture of sgRNA and transfection reagent was incu-
585 bated for 30 minutes, following which, 2,000 cells/well (WT HeLa integrated with
586 Cas9/MLF2-GFP) in antibiotic-free DMEM were added using a Thermo MultiDrop
587 Combi (Thermo Fisher Scientific). The plates were then incubated for 72 hours at 37°C.
588 To induce MLF2-GFP reporter expression, 0.5 µg/mL Dox was added to each well 16 h
589 before fixation. Media was removed and cells washed once with PBS and fixed with 4%
590 paraformaldehyde (PFA) for 20 minutes. Cells were washed twice with PBS, permea-
591 bilized with 0.1% Triton X-100 for 5 minutes, and washed again with PBS. Cells were
592 stained with Hoechst 33342 dye (1:1000) (Invitrogen, H1399) and HCS CellMask Deep
593 Red Stain (Invitrogen, H32721) (1:5000 of 10 mg/mL, Thermo Fisher Scientific) diluted
594 in PBS for 30 minutes, and washed a final time in PBS. The Molecular Devices ImageX-
595 press 4.0 high-throughput fluorescent microscope (40x magnification) was used to ob-
596 tain 9-image sets (each consisting of DAPI, MLF2-GFP, and Deep Red CellMask im-
597 ages) per well, resulting in 10,368 TIFF images/plate. Images were uploaded to Yale's
598 McCleary high-performance computing cluster (Yale Center for Research Computing)
599 for analysis.

600 For chemical screening, the Pharmakon 1760 drug library (MicroSource Discov-
601 ery Systems, Inc.) was used, consisting of 1360 FDA-approved drugs and 400 Interna-
602 tional drugs on six 384-well plates. Each screening plate consisted of 16 replicates of
603 the negative control (4TorKO HeLa cells treated with DMSO [0.1% v/v]) and 16 repli-
604 cates of the positive control (WT HeLa cells treated with DMSO [0.1% v/v]). Plates were

605 seeded with 4,000 HeLa cells/well (Multidrop Combi) in 20 µL of DMEM before incubat-
606 ing at 37°C for 24 h. MLF2-GFP expression was induced by adding 0.5 µg/mL of Dox
607 into each well 8 h prior to fixation. Compounds were resuspended in DMSO and added
608 to each well 2 h prior to fixation using a Labcyte Echo 550 Acoustic Dispenser (Labcyte,
609 now Beckman), at a final concentration of 10 µM. Cells were washed, fixed, stained,
610 and imaged using the methods outlined above.
611
612 **Image segmentation of phenotypic screen data**
613 We created a custom analysis pipeline to segment condensate foci into single cells us-
614 ing the open-source software *CellProfiler 4.2.6.*(*40*, *41*) and downstream statistical anal-
615 ysis using *R Studio* (Version 2023.12.1+402).(*92*) Individual nuclei were distinguished
616 by shape for objects between 30-100 (+4) pixels using an adaptive (40 pixels) Minimum
617 Cross-Entropy thresholding (0.01-1.0 threshold) in *CellProfiler 4*. The cytoplasm of each
618 nucleus was delineated as a secondary object of each nucleus using a watershed algo-
619 rithm with global Otsu thresholding (three classes, foreground; 0-1.0 threshold, thresh-
620 old correction factor = 0.7).(*41*, *93*) Each GFP-channel image was enhanced using a
621 white top-hat transform (speckles; maximum 10 pixels), allowing each condensate foci
622 between 1-10 pixels to be detected using adaptive (50 pixels) Otsu thresholding (three
623 classes, foreground; 0.0175-1.0 threshold, threshold correction factor = 1.2). Groups of
624 condensate foci were then distinguished using local intensity and related to their corre-
625 sponding cytoplasm/nuclei. The same image analysis pipeline was applied across all
626 imaging experiments, with only minor adjustments (e.g., cell size or foci intensity thresh-
627 olds). Any parameter changes were applied uniformly across all conditions and controls
628 within a given experiment.
629 For parallel analysis, a Bash script was developed allowing the execution of mul-
630 tiple headless CellProfiler instances on a SLURM-managed compute cluster. Specifi-
631 cally, after constructing the analysis pipeline in the CellProfiler GUI, the CreateBatch-
632 Files module is used to generate a batch configuration file that defines image metadata
633 and processing instructions. This batch file is then imported to the cluster environment,
634 and all images are registered within a single CellProfiler project in array format, with im-
635 ages from each plate corresponding to a separate SLURM job. This approach enables

636 autonomous analysis of over 1 million high-content images.TIFF images—amounting to
637 multiple terabytes of data—within a matter of hours. Both the CellProfiler pipeline and
638 Bash scripts are open source and available publicly on the SchliekerLab GitHub page
639 as part of the CondenScreen pipeline (https://github.com/SchliekerLab/Condenscreen-
640 v1.1.git).
641

642 **Statistical analysis of screen data**

643 We developed a condensate screening using a combined *CellProfiler* and *R* pipeline,
644 *CondenScreen*, to statistically analyze all screen data. The *R*-script features an interac-
645 tive GUI allowing the user to select plate size (6-well to 384-well format), plate layout,
646 and input whether the data was taken from a signal "ON" or a signal "OFF" screen. With
647 limited user input, it then automatically iterates through images from each plate and as-
648 signs groups of image sets into their corresponding wells using string parsing. To con-
649 trol for variations in transfection efficiencies that could inadvertently impact downstream
650 scoring, the program has an optional filter to select only the top 25% of cells showing
651 the highest (or lowest) signal using the modest assumption that at least 1 out of 4 cells,
652 on average, will have received the sgRNAs of interest. Additionally, for both screens,
653 the image with the highest score was filtered out, as there were several conditions
654 where one image-set had screening aberrations skewing the condition toward a false
655 positive result. For each plate, aggregate statistics of the respective positive and nega-
656 tive controls are compared, including signal-to-background (SB), standard deviation
657 (SD) of the controls, and the coefficient of variation (CV) of the controls, among others.
658 To determine separation between positive and negative controls, a Z' assay quality
659 score was calculated: $Z' = 1 - (3(\sigma_{PC} + \sigma_{NC})/|\mu_{PC} - \mu_{NC}|)$,(*42*) where σ_{PC} and σ_{NC} is
660 the SD and μ_{PC} and μ_{NC} is the condensate mean of the positive and negative controls,
661 respectively.

662 The percent effect of each condition was quantified by either $E\% = \left(\frac{\mu - \overline{x}}{\mu}\right) \cdot 100$
663 (for signal "OFF" selection) or $E\% = \left(\frac{\overline{x} - \mu}{\mu}\right) \cdot 100$ (for signal "ON" selection) where $\overline{x}$ is
664 the is the average number of condensate foci/cell per tested condition, and μ is the av-
665 erage number of condensate foci/cell across the screening population of the plate. The

significance scoring function is automatically changed depending on selection of a signal "ON" vs a signal "OFF" screen, although this can be adjusted.

For signal "ON" screens, the significance of each condition was calculated using the Z-Score to normalize for intraplate variations: $Z = (E\% - \mu_{E\%})/\ \sigma_{E\%}$, where $\mu_{E\%}$ and $\sigma_{E\%}$ represent the population average percent effect and standard deviation, respectively. Since signal "OFF" screens can often have increased variability, the significance of each well was quantified by calculating the BZ-Score. First, B-Scores were computed by applying Tukey's median polish to each plate to remove row, column, and edge effects. The residuals from the median polish were then scaled by the median absolute deviation (MAD) to produce a robust per-well B-score. Then, the BZ-Score was calculated using the following formula: $BZ = (BScore - \mu_{BScore})/\ \sigma_{BScore}$, where μ_{BScore} and σ_{BScore} are the mean and standard deviation of the B-Scores over the entire plate.

To determine how chemical or genetic perturbations alter cell viability, we derived a function to extrapolate differential cellular health by morphological characterization of the nucleus. Specifically, we added several modules to *CellProfiler* to measure the size, shape, intensity, and texture of each nucleus.(*40*) This is crucial as several compound concentrations had a non-significant effect in changing the total cell count, but greatly disrupted nuclear morphology based on visual inspection. After testing the effects of multiple parameters, we derived the following empirical formula which best represented the data to estimate nuclear health: $NucHealth = \Delta_{CC}\ + ((100\ +\ \Delta_{CC}\ )/100) \cdot \sum(\Delta_{Area} +\ \Delta_{FF}\ )/2$, where Δ_{CC} , Δ_{Area}, and Δ_{FF} are the normalized difference in cell count, nuclear area, and nuclear FormFactor (measurement of nuclear irregularity) compared to the screening population (or DMSO for dose-response), respectively. Normalization was calculated according to: $\Delta_v = ((V - \mu_V)/\mu_V) \cdot 100$, in which *V* is the variable of interest. The equation was derived with the following assumptions and limitations. (1) Changes to the area and FormFactor should be weighted equally, and (2) cell death should be given a higher impact weight over morphological changes.

CondenScreen provides additional downstream inter-plate and intra-plate visualizations using open-source libraries, including *ggplot2* and *ggplate*. The output includes a ranked list of hits, plots of the screening distributions (raw condensate foci/cell and %effect), a Z/BZ-score plot, and individual plate heatmaps for each 384-well plate.

697 Finally, top hits in both screens were inspected manually to ensure quality control and
698 rule out any additional potential screening aberrations. Notably, manual inspection was
699 rarely a reliable measure of hit robustness. In several cases, hits that were initially over-
700 looked during first-pass screening (false negatives) were later validated and correctly
701 identified by the program. The *CondenScreen* platform is publicly available on the
702 SchliekerLab GitHub page: https://github.com/SchliekerLab/Condenscreen-v1.1.git
703

704 **Development of phenotypic machine learning model to distinguish condensate**

705 We developed a supervised machine learning model to differentiate between subtle
706 phenotypic differences in nuclear condensates. The model was built around the Res-
707 Net18 architecture(*68*) with pretrained initial starting weights borrowed from ImageNet1k
708 IMAGENET1K_V1(*69*) using torchvision(*94*) on Python 3.11.8 (conda-forge). Using im-
709 ages directly from the CRISPR KO screen, 4,523 single cells were extracted, cropped
710 (100x100) and resized (224x224) before centering (DAPI-channel) each cell. Single
711 cells were labeled in accordance with their file names, with NT Ctrl cells = 0, RNF26 KO
712 cells = -1, and ZNF335 KO cells = +1. To ensure even sampling numbers across condi-
713 tions, cells from the overrepresented NT Ctrl conditioned were downsampled to match
714 KO conditions, with only the MLF2-GFP channel selected for analysis. Data were split
715 80:20 into train and validation sets (sklearn.model_selection: train_test_split)(*95*). Ten-
716 sors were packaged (torch.utils.data) before employing the following parameters during
717 model training (PyTorch(*96*)): AdamW (lr = 1e-4, weight decay = 5e-2), dropout = 0.5,
718 hidden dimension = 64, SmoothL1Loss, CosineAnnealingLR, 100 max epochs, and
719 early stopping (patience = 25). Model performance was evaluated using several mod-
720 ules from sklearn.metrics and scipy.stats,(*95*, *97*) including: mean_absolute_error();
721 r2_score(); pearsonr(); accuracy_score(); confusion_matrix(); roc_curve();
722 roc_auc_score(); and ConfusionMatrixDisplay().
723 After optimization, all CRISPR KO gene hits with a Z-Score > 2 were cropped,
724 resized, and centered before randomly downsampled to 500 cells per condition (ac-
725 counting for differences in cell viability). Cells were then analyzed using the optimized
726 ResNet18 model and regression scores trained to align with condition labels (–1, 0, +1)
727 along with latent representation embeddings (512-dimensional latent embeddings per

728 cell, e.g., condensate morphology, size, intensity, etc.) were extracted in separate ar-
729 rays and feature vectors, respectively. Relative centroid vector scores were computed
730 based on the cosine similarity to each anchored centroid (RNF26/NT Ctrl/ZNF335) for
731 each condition and exported to an .xlsx file (see Table S3).
732 For visualization, dimensionality reduction was performed with UMAP (n_neigh-
733 bors = 15, min_dist = 0.1). The optimal cluster number was selected using the elbow
734 method (KMeans, k = 1–10) and silhouette analysis (sklearn.metrics: silhouette_sam-
735 ples | silhouette_score), yielding three distinct clusters. Embeddings were reduced to
736 two dimensions using UMAP (n_neighbors = 15, min_dist = 0.1, cosine metric), and
737 KMeans clustering with *k* = 3 was applied on the UMAP coordinates to assign genes to
738 discrete clusters. Finally, a hierarchical heatmap was constructed using pandas, sea-
739 born.clustermap, and matplotlib.pyplot. For a full list of modules imported and model
740 training parameters, please refer to our source script: https://github.com/Schliek-
741 erLab/Condenscreen-v1.1.git
742
743 **Solubilization of hit compounds**
744 All putative hit compounds were resuspended up to 100 mM in DMSO. The following
745 compounds were solubilized: pyrithione zinc (PZ) (CAS 13463-41-7), celastrol (CAS
746 34157-83-0), nerol (CAS 106-25-2), spiramycin (CAS 8025-81-8), choline chloride (CAS
747 67-48-1), disulfiram (CAS 97-77-8), bronopol (CAS 52-51-7), penfluridol (CAS 26864-
748 56-2), plumbagin (CAS 481-42-5), dyphylline (CAS 479-18-5), econazole nitrate (CAS
749 24169-02-6), ethacrynic acid (CAS 58-54-8), piroctone olamine (CAS 68890-66-4), dihy-
750 drocelastryl (CAS 80-97-7), fenipentol (CAS 583-03-9), triclosan (CAS 3380-34-5),
751 eprodisate disodium (CAS 36589-58-9), L-phenylalanine (CAS 63-91-2), nabumetone
752 (CAS 42924-53-8), edrophonium chloride (CAS 116-38-1), josamycin (CAS 16846-24-
753 5), mevastatin (CAS 73573-88-3), clofibric acid (CAS 882-09-7), docosahexaenoic acid
754 (CAS 6217-54-5).
755
756 **Dose-dependent validation of small molecules**
757 Based on the change in MLF2-GFP foci, viability, and structural diversity, the 22 top-
758 scoring putative hit compounds were resuspended to 100 mM in DMSO. Using

759 analogous methods described in the "High-content screening strategy" section, a 12-
760 point dose response was constructed with concentrations ranging from < 100 nM to 100
761 µM of each compound, with a final working DMSO concentration of 0.1%. All com-
762 pounds were tested using four replicates, and 12-positive (WT HeLa) and 12-negative
763 (DMSO) controls were included on each screening plate to ensure assay robustness.
764 MLF2-GFP was induced with 0.5 µg/mL of Dox 8 h before fixation, and compounds
765 were added 3 h prior to fixation. Cells were washed, fixed, stained, and imaged as de-
766 scribed above. All dose-response images were analyzed using an analogous *CellPro-*
767 *filer* + *R*-script pipeline.
768 The effect of each concentration was interpolated (*R*, *drc* & *ggplot*) using a four-
769 parameter log-logistic function for each compound: $f(x) = \min + (\max - \min)/(1 +$
770 $e^{h(\log(x) - \log(IC_{50}))})$, where x is the concentration, min represents the lower asymptote,
771 max represents the higher asymptote, IC_{50} is the concentration at which response is
772 50% of the way between min and max, and h is the hill coefficient.(*98*, *99*) The percent
773 toxicity was calculated by fitting the additive inverse of the nuclear health function val-
774 ues to the same model.
775
776 **Pathway enrichment analysis**
777 Gene Ontology (GO) term enrichment analysis was used for the top 100 CRISPR gene
778 hits.(*59*) The gene/chemical hits were separately analyzed in *R* to determine enriched
779 biological process pathways using the enrichGO function (*clusterProfiler*).(*100*) En-
780 riched pathways were compared against an annotation database (org.Hs.eg.db). The
781 top 30 pathways were then plotted using the pheatmap function and additionally plotted
782 to an interactive .html graphic using heatmaply.(*101*) Pathways were grouped by hierar-
783 chical clustering using Euclidean distance with complete linkage. To explore the interac-
784 tive heatmap, please visit the Schlieker Lab GitHub page and download the correspond-
785 ing .html files under the "Interactive Figures" folder.
786
787 **Immunofluorescence and confocal microscopy:**
788 HeLa cells were grown on 13 mm round coverslips and fixed with 4% (v/v) paraformal-
789 dehyde (Thermo Fisher Scientific), washed with 1x PBS (Gibco), permeabilized with

790 0.1% (v/v) Triton X-100 (Sigma-Aldrich) for 5 minutes, and blocked with a PBS mixture
791 containing 2% (w/v) BSA and 2% (w/v) glycine. Primary antibodies were diluted in
792 blocking buffer and added dropwise to coverslips for a one-hour incubation. After wash-
793 ing 3x with blocking buffer, fluorescent secondary antibodies diluted in blocking buffer
794 were added for an additional hour, followed by a 1x PBS wash. Coverslips were then
795 stained using a 1:2000 dilution of Hoechst 33342 (Life Technologies). Cells were then
796 washed 3x with PBS before mounting onto slides (Fluoromount-G, Southern Biotech).
797 Primary antibodies were diluted 1:500: anti-HA produced in rat (11867423001, Roche),
798 anti-K48 produced in rabbit (ZRB2150, Millipore Sigma), and anti-Mab414 produced in
799 mouse. Secondary antibodies were conjugated to Alexa Fluor 568 (Life Technologies).
800 Epifluorescence images were captured using a Zeiss AXIO Observer Z1 inverted
801 fluorescence microscope using a ×63/1.40 oil objective. All confocal images were ac-
802 quired at Yale's Imaging Facility on Science Hill using a Zeiss LSM 880 laser scanning
803 confocal microscope with a C Plan-Apochromat ×63/1.40 oil objective. Images were an-
804 alyzed with *CondenScreen* and compiled for publication using Zeiss v3.4 and Image J
805 (Fiji) software.
806
807 **Validation of hit compounds**
808 Based on the dose-response results, we focused validation efforts on four compounds:
809 PZ (IC_{50} = 601.6 nM), celastrol (IC_{50} = 124.3 nM), disulfiram (IC_{50} = 10.8 μM), and plum-
810 bagin (IC_{50} = 7.41 μM). To determine the optimal concentrations using a tighter set of
811 concentrations, MLF2-GFP expression was induced in 4TorKO HeLa cells (0.5 μg/mL
812 Dox, 8 h) and treated with either DMSO or one of four different concentrations/com-
813 pounds: 2x below the IC_{50}, IC_{50}, 2x above the IC_{50}, or 8x above the IC_{50}. Compounds
814 were added 3 h before fixation. For each condition, images (n > 50 cells/condition) were
815 captured on a Zeiss AXIO Observer Z1 inverted fluorescence microscope and analyzed
816 using our *CellProfiler* pipeline. *R Studio* was used to cluster MLF2-GFP foci into single
817 cells grouped by treatment condition and construct jittered boxplots. To determine the
818 effect of compounds on K48-ubiquitinylated proteins accumulation within condensates,
819 we incubated cells with compounds for 3h (PZ, 3 μM; celastrol, 3 μM; plumbagin, 15
820 μM; and disulfiram, 50 μM), followed by immunofluorescent labeling with α-K48 primary

821 antibody (1:500). Images were captured and analyzed as described above. High-resolu-
822 tion representative images were acquired using confocal microscopy (Zeiss LSM 880).
823 To determine differential protein expression, we constructed a 4TorKO HeLa cell
824 line with a Dox inducible MLF2-GFP-P2A-tdTomato marker, where tdTomato serves as
825 a diffuse secondary reporter for global protein synthesis. Cells were induced with 0.5
826 µg/mL Dox and treated with their respective compound for 6 h before fixation. Imaging
827 was carried out on the Zeiss AXIO and the Zeiss LSM 880 for quantification and high-
828 resolution, respectively. The integrated intensity of both GFP and tdTomato signal was
829 quantified (n > 50) with *CellProfiler* and *R*. To ascertain the relative influence of global
830 protein expression inhibition on the reduced number of MLF2-GFP foci, we derived the
831 following empirical equation to determine the condensate normalization score ($CNS =$
832 $(1 - (\overline{F}_x/\overline{F}_{DMSO})) \cdot (\overline{II}_x/\overline{II}_{DMSO})^2$, where $\overline{F}$ and $\overline{II}$ are the average number of NE con-
833 densate foci and the average integrated intensity of the secondary marker, respectively,
834 for either the tested condition (x) or the DMSO control. The scoring system is designed
835 to prioritize compounds that robustly reduce condensate foci without affecting tdTomato
836 reporter expression and penalizes those that have little effect on foci or substantially
837 suppress reporter expression. It penalizes modulators that have either little effect on foci
838 or those that substantially suppress reporter expression. The weighting is particularly
839 sensitive to changes in tdTomato intensity due to the squared exponent, thereby
840 strongly penalizing compounds which impair protein expression. A low CNS-score sug-
841 gests non-specific or off-target activities, such as decreased transcription, translation, or
842 increased biomarker turnover.
843
844 **Guide RNA-mediated KO experiments:**
845 Each gene KO consisted of a pool of 3 distinct, synthetic sgRNAs, each targeting a dif-
846 ferent portion of the same gene (Horizon Discovery). Gene knock-out was conducted
847 using either DharmaFECT or DharmaFECT Duo Transfection Reagent (Horizon discov-
848 ery) using Tor1A KO + Tor1B KO as a positive control and NT sgRNAs as a negative
849 control (Horizon discovery). The sgRNAs were transfected into stably-integrated Cas9
850 expressing HeLa cells and incubated for 72 h with a final sgRNA concentration of 50

851 nM. The cells were treated with 0.5 µg/mL Dox to induce the MLF2-GFP expression 6 h
852 before fixation.
853 The pooled sgRNA sequences used were: ZNF335 sgRNAs (SG-016173, Hori-
854 zon Discovery), GATGATGGTGGCGCCGCCCA, GTGCCAGGGTGATCTGCGGT, AC-
855 CAAATGGACCAGGTGCCA; RNF26 sgRNAs (SG-007060, Horizon Discovery),
856 TATCATGGAGGCAGTGTACC, GCATGACATCCTCTCGAAGA, CAGTGTACCTGG-
857 TAGTGAAT; TOR1A sgRNAs (SG-011023, Horizon Discovery)
858 TATGAACATGGCTTTCTGGT, TATGAAGATGGACCTCGCAC, AGTG-
859 CAATGTGGCCACAAAC; and TOR1B sgRNAs (SG-014203-01, Horizon Discovery),
860 AGTGCAGAGTCGATACAAAC, CAAGAGCCGTCTTAGTTATA, TTGGCGGGCCG-
861 GAAGAAAGA.
862
863 **Sodium arsenite and 1,6-hexanediol treatment**
864 To determine the co-localization of MLF2 into stress granules, Lipofectamine 2000 was
865 used to co-transfect MLF2 with either G3BP1-GFP (Addgene #135997) + MLF2-HA,
866 TDP-43-tdTomato-HA (Addgene #28205) + MLF2-GFP-Flag (Addgene #230986), or
867 pcDNA3.1 empty vector control. Next day (16 h post-transfection), HeLa cells were
868 treated with $NaAsO_2$ (0.5 mM) and incubated for 30 min. Cells were fixed for 20 min with
869 PFA, stained, and imaged as outlined. For hexanediol treatment, cells were incubated in
870 complete DMEM medium containing 5% (v/v) 1,6-hexanediol (CAS 629-11-8) or 5%
871 (v/v) 2,5-hexanediol (CAS 2935-44-6) for 5 min before fixation and immunofluorescence
872 staining as described.
873
874 **Electron Microscopy:**
875 Electron microscopy was performed at the Center for Cellular and Molecular Imaging,
876 Yale School of Medicine, using our previously described workflow (Laudermilch et al.,
877 2016).(*30*) Briefly, cells were fixed for 1 h in 2.5% glutaraldehyde in 0.1 M sodium caco-
878 dylate buffer, pH 7.4. Cells were rinsed and scraped in 1% gelatin and centrifuged in a
879 2% agar solution. Chilled cell blocks were processed with osmium and thiocarbohydra-
880 zide-osmium ligand as described (West et al., 2010), and samples were embedded in
881 Durcupan ACM resin (Electron Microscopy Science). Polymerization was performed by

incubating samples at 60°C overnight. These blocks were cut into 60 nm sections with a Leica UltraCut UC7 and stained with 2% uranyl acetate and lead citrate on Formvar/carbon-coated grids. Samples were visualized with an FEI Tecnai Biotwin TEM at 80 kV, and pictures were taken with Morada CCD and iTEM (Olympus) software.

Statistical Analysis:

Genetic and chemical hits were ranked based on their Z-Score or BZ-Score, respectively, as described in detail within the *Statistical analysis of screen data* section. GO pathways were ranked using a p-value cutoff of 0.05, after controlling for false discovery rate using the Benjamini-Hochberg (BH) procedure. The significance of all chemical and genetic validation experiments was assessed using a Bonferroni multiple comparisons corrected Mann–Whitney U test with > 50 cells per condition individually analyzed. Unless otherwise stated, asterisks indicate Bonferroni corrected p-values using a Mann-Whitney U test where *: p <= 0.05; **: p <= 0.01; ***: p <= 0.001; and ****: p <= 0.0001.

AWKNOWLEDGEMENTS

Funding:

This work was funded by the DOD PR200788 (CS) awarded by the U.S. Department of Defense, Department of the Army, along with support from the Dystonia Medical Research Foundation. Additionally, this work was supported by the NIH T32 GM145469 (DP), the European Molecular Biology Organization (EMBO) ALTF 910-2022 (EFFK), a Bachmann-Strauss Dystonia Fellowship (SM), and the Dutch Research Council (NWO) 019.222EN.007 (EFEK) This work utilized ImageXpress Micro 4 high content imager that was purchased with funding from a National Institutes of Health SIG grant 1S10OD032384. We thank Morven Graham and the Center for Cellular and Molecular Imaging (CCMI) Electron Microscopy Facility for the collection of TEM images and Shawn Ferguson for sharing stress granule marker reagents. The authors thank Robert Niescier, Anthony J Koleske, and Matthew Simon for their expert advice and feedback. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions:

Conceptualization: DP, CM, EFEK, YVS, CS
Methodology: DP, CM, SM, EFEK, VT, ND
Investigation: DP, CM, SM, EFEK, VT, ND, CS
Computational Pipeline: DP, VT
Machine Learning: DP
Visualization: DP
Supervision: CS
Writing—original draft: DP
Writing—review & editing: DP, CS, EFEK, CM, SM, ND, YVS, CS

Competing interests:

Authors declare that they have no competing interests.

Data and materials availability:

All data are available in the main text or the supplementary materials. The computational *CondenScreen* pipeline and machine learning source code are available on the Schlieker Lab GitHub page: https://github.com/SchliekerLab/Condenscreen-v1.1.git. At the time of publication, all raw screening images will be uploaded to the EMBL-EBI BioImage Archive: https://www.ebi.ac.uk/bioimage-archive/.

947 **REFERENCES**

948 1. S. M. Prophet, A. J. Rampello, R. F. Niescier, J. E. Gentile, S. Mallik, A. J. Koleske,
949 C. Schlieker, Atypical nuclear envelope condensates linked to neurological disor-
950 ders reveal nucleoporin-directed chaperone activities. *Nat. Cell Biol.* **24**, 1630–
951 1641 (2022).

952 2. I. R. Mackenzie, A. M. Nicholson, M. Sarkar, J. Messing, M. D. Purice, C. Pottier, K.
953 Annu, M. Baker, R. B. Perkerson, A. Kurti, B. J. Matchett, T. Mittag, J. Temirov, G.-
954 Y. R. Hsiung, C. Krieger, M. E. Murray, M. Kato, J. D. Fryer, L. Petrucelli, L. Zin-
955 man, S. Weintraub, M. Mesulam, J. Keith, S. A. Zivkovic, V. Hirsch-Reinshagen, R.
956 P. Roos, S. Züchner, N. R. Graff-Radford, R. C. Petersen, R. J. Caselli, Z. K.
957 Wszolek, E. Finger, C. Lippa, D. Lacomis, H. Stewart, D. W. Dickson, H. J. Kim, E.
958 Rogaeva, E. Bigio, K. B. Boylan, J. P. Taylor, R. Rademakers, TIA1 Mutations in
959 Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Sep-
960 aration and Alter Stress Granule Dynamics. *Neuron* **95**, 808-816.e9 (2017).

961 3. S. Alberti, D. Dormann, Liquid-Liquid Phase Separation in Disease. *Annu. Rev.*
962 *Genet.* **53**, 171–194 (2019).

963 4. N. Kedersha, G. Stoecklin, M. Ayodele, P. Yacono, J. Lykke-Andersen, M. J. Fritzler,
964 D. Scheuner, R. J. Kaufman, D. E. Golan, P. Anderson, Stress granules and pro-
965 cessing bodies are dynamically linked sites of mRNP remodeling. *J. Cell Biol.* **169**,
966 871–884 (2005).

967 5. B. Portz, B. L. Lee, J. Shorter, FUS and TDP-43 Phases in Health and Disease.
968 *Trends Biochem. Sci.* **46**, 550–563 (2021).

969 6. S. Alberti, A. Gladfelter, T. Mittag, Considerations and Challenges in Studying Liq-
970 uid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419–434
971 (2019).

972 7. T. Mittag, R. V. Pappu, A conceptual framework for understanding phase separation
973 and addressing open questions and challenges. *Mol. Cell* **82**, 2201–2214 (2022).

974 8. C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gha-
975 rakhani, F. Jülicher, A. A. Hyman, Germline P granules are liquid droplets that lo-
976 calize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).

977 9. H. B. Schmidt, D. Görlich, Transport Selectivity of Nuclear Pores, Phase Separation,
978 and Membraneless Organelles. *Trends Biochem. Sci.* **41**, 46–61 (2016).

979 10. P. Li, S. Banjade, H.-C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hol-
980 lingsworth, D. S. King, S. F. Banani, P. S. Russo, Q.-X. Jiang, B. T. Nixon, M. K.
981 Rosen, Phase transitions in the assembly of multivalent signalling proteins. *Nature*
982 **483**, 336–340 (2012).

11. B. Wolozin, P. Ivanov, Stress granules and neurodegeneration. *Nat. Rev. Neurosci.* **20**, 649–666 (2019).

12. S. Alberti, A. A. Hyman, Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat. Rev. Mol. Cell Biol.* **22**, 196–213 (2021).

13. X. Yan, D. Kuster, P. Mohanty, J. Nijssen, K. Pombo-García, J. G. Morato, A. Rizuan, T. M. Franzmann, A. Sergeeva, A. M. Ly, F. Liu, P. M. Passos, L. George, S.-H. Wang, J. Shenoy, H. L. Danielson, B. Ozguney, A. Honigmann, Y. M. Ayala, N. L. Fawzi, D. W. Dickson, W. Rossoll, J. Mittal, S. Alberti, A. A. Hyman, Intra-condensate demixing of TDP-43 inside stress granules generates pathological aggregates. *Cell* **188** (2025).

14. M. Vendruscolo, M. Fuxreiter, Protein condensation diseases: therapeutic opportunities. *Nat. Commun.* **13**, 5550 (2022).

15. S. C. Ng, D. Görlich, A simple thermodynamic description of phase separation of Nup98 FG domains. *Nat. Commun.* **13**, 6172 (2022).

16. T. Kozai, J. Fernandez-Martinez, T. van Eeuwen, P. Gallardo, L. E. Kapinos, A. Mazur, W. Zhang, J. Tempkin, R. Panatala, M. Delgado-Izquierdo, B. Raveh, A. Sali, B. T. Chait, L. M. Veenhoff, M. P. Rout, R. Y. H. Lim, Dynamic molecular mechanism of the nuclear pore complex permeability barrier. *bioRxiv*, 2023.03.31.535055 (2023).

17. D. Cowburn, M. Rout, Improving the hole picture: towards a consensus on the mechanism of nuclear transport. *Biochem. Soc. Trans.* **51**, 871–886 (2023).

18. S. Frey, R. P. Richter, D. Görlich, FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* **314**, 815–817 (2006).

19. S. Mallik, D. Poch, S. Burick, C. Schlieker, Protein folding and quality control during nuclear transport. *Curr. Opin. Cell Biol.* **90**, 102407 (2024).

20. S. Kim, S. Phan, H. T. Tran, T. R. Shaw, S. H. Shahmoradian, M. H. Ellisman, S. L. Veatch, S. J. Barmada, S. S. Pappas, W. T. Dauer, TorsinA is essential for neuronal nuclear pore complex localization and maturation. *Nat. Cell Biol.* **26**, 1482–1495 (2024).

21. A. J. Rampello, E. Laudermilch, N. Vishnoi, S. M. Prophet, L. Shao, C. Zhao, C. P. Lusk, C. Schlieker, Torsin ATPase deficiency leads to defects in nuclear pore biogenesis and sequestration of MLF2. *J. Cell Biol.* **219** (2020).

22. A. N. Coyne, J. D. Rothstein, Nuclear pore complexes — a doorway to neural injury in neurodegeneration. *Nat. Rev. Neurol.* **18**, 348–362 (2022).

1019 23. Y. Fan, Z. Si, L. Wang, L. Zhang, DYT-TOR1A dystonia: an update on pathogene-
1020 sis and treatment. *Front. Neurosci.* **17**, 1216929 (2023).

1021 24. B. Balint, N. E. Mencacci, E. M. Valente, A. Pisani, J. Rothwell, J. Jankovic, M.
1022 Vidailhet, K. P. Bhatia, Dystonia. *Nat. Rev. Dis. Primer* **4**, 1–23 (2018).

1023 25. L. J. Ozelius, J. W. Hewett, C. E. Page, S. B. Bressman, P. L. Kramer, C. Shalish,
1024 D. de Leon, M. F. Brin, D. Raymond, D. P. Corey, S. Fahn, N. J. Risch, A. J. Buck-
1025 ler, J. F. Gusella, X. O. Breakefield, The early-onset torsion dystonia gene (DYT1)
1026 encodes an ATP-binding protein. *Nat. Genet.* **17**, 40–48 (1997).

1027 26. C.-C. Liang, L. M. Tanabe, S. Jou, F. Chi, W. T. Dauer, TorsinA hypofunction
1028 causes abnormal twisting movements and sensorimotor circuit neurodegeneration.
1029 *J. Clin. Invest.* **124**, 3080–3092 (2014).

1030 27. C. Zhao, R. S. H. Brown, A. R. Chase, M. R. Eisele, C. Schlieker, Regulation of
1031 Torsin ATPases by LAP1 and LULL1. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E1545–
1032 E1554 (2013).

1033 28. R. S. H. Brown, C. Zhao, A. R. Chase, J. Wang, C. Schlieker, The mechanism of
1034 Torsin ATPase activation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E4822–E4831
1035 (2014).

1036 29. F. E. Demircioglu, B. A. Sosa, J. Ingram, H. L. Ploegh, T. U. Schwartz, Structures
1037 of TorsinA and its disease-mutant complexed with an activator reveal the molecular
1038 basis for primary dystonia. *eLife* **5**, e17983 (2016).

1039 30. E. Laudermilch, P.-L. Tsai, M. Graham, E. Turner, C. Zhao, C. Schlieker, Dissect-
1040 ing Torsin/cofactor function at the nuclear envelope: a genetic study. *Mol. Biol. Cell*
1041 **27**, 3964–3971 (2016).

1042 31. R. E. Goodchild, C. E. Kim, W. T. Dauer, Loss of the Dystonia-Associated Protein
1043 TorsinA Selectively Disrupts the Neuronal Nuclear Envelope. *Neuron* **48**, 923–932
1044 (2005).

1045 32. E. F. E. Kuiper, P. Gallardo, T. Bergsma, M. Mari, M. Kolbe Musskopf, J. Kuipers,
1046 B. N. G. Giepmans, A. Steen, H. H. Kampinga, L. M. Veenhoff, S. Bergink, The
1047 chaperone DNAJB6 surveils FG-nucleoporins and is required for interphase nu-
1048 clear pore complex biogenesis. *Nat. Cell Biol.* **24**, 1584–1594 (2022).

1049 33. E. W. Martin, A. S. Holehouse, I. Peran, M. Farag, J. J. Incicco, A. Bremer, C. R.
1050 Grace, A. Soranno, R. V. Pappu, T. Mittag, Valence and patterning of aromatic res-
1051 idues determine the phase behavior of prion-like domains. *Science* **367**, 694–699
1052 (2020).

1053 34. C. Colombrita, E. Zennaro, C. Fallini, M. Weber, A. Sommacal, E. Buratti, V. Silani,
1054 A. Ratti, TDP-43 is recruited to stress granules in conditions of oxidative insult. *J.*
1055 *Neurochem.* **111**, 1051–1061 (2009).

1056 35. F. Palangi, S. M. Samuel, I. R. Thompson, C. R. Triggle, M. M. Emara, Effects of
1057 oxidative and thermal stresses on stress granule formation in human induced plu-
1058 ripotent stem cells. *PLOS ONE* **12**, e0182059 (2017).

1059 36. M. Y. Fang, S. Markmiller, A. Q. Vu, A. Javaherian, W. E. Dowdle, P. Jolivet, P. J.
1060 Bushway, N. A. Castello, A. Baral, M. Y. Chan, J. W. Linsley, D. Linsley, M. Mer-
1061 cola, S. Finkbeiner, E. Lecuyer, J. W. Lewcock, G. W. Yeo, Small-Molecule Modu-
1062 lation of TDP-43 Recruitment to Stress Granules Prevents Persistent TDP-43 Ac-
1063 cumulation in ALS/FTD. *Neuron* **103**, 802-819.e11 (2019).

1064 37. H. Uechi, S. Sridharan, J. Nijssen, J. Bilstein, J. M. Iglesias-Artola, S. Kishigami, V.
1065 Casablancas-Antras, I. Poser, E. J. Martinez, E. Boczek, M. Wagner, N. Tom-
1066 schke, A. M. de Jesus Domingues, A. Pal, T. Doeleman, S. Kour, E. N. Anderson,
1067 F. Stein, H. O. Lee, X. Zhang, A. W. Fritsch, M. Jahnel, J. Fürsch, A. C. Murthy, S.
1068 Alberti, M. Bickle, N. L. Fawzi, A. Nadler, D. C. David, U. B. Pandey, A. Hermann,
1069 F. Stengel, B. G. Davis, A. J. Baldwin, M. M. Savitski, A. A. Hyman, R. J. Wheeler,
1070 Small-molecule dissolution of stress granules by redox modulation benefits ALS
1071 models. *Nat. Chem. Biol.* **21**, 1–12 (2025).

1072 38. T. Bergsma, A. Steen, J. L. Kamenz, T. A. Otto, P. Gallardo, L. M. Veenhoff, Imag-
1073 ing-based quantitative assessment of biomolecular condensates in vitro and in
1074 cells. *J. Biol. Chem.* **301**, 108130 (2024).

1075 39. D. W. Baggett, A. Medyukhina, S. Tripathi, H. K. Shirnekhi, H. Wu, S. B. Pounds,
1076 K. Khairy, R. Kriwacki, An Image Analysis Pipeline for Quantifying the Features of
1077 Fluorescently-Labeled Biomolecular Condensates in Cells. *Front. Bioinforma.* **2**
1078 (2022).

1079 40. D. R. Stirling, M. J. Swain-Bowden, A. M. Lucas, A. E. Carpenter, B. A. Cimini, A.
1080 Goodman, CellProfiler 4: improvements in speed, utility and usability. *BMC Bioin-*
1081 *formatics* **22**, 433 (2021).

1082 41. A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D.
1083 A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, D. M. Sabatini, Cell-
1084 Profiler: image analysis software for identifying and quantifying cell phenotypes.
1085 *Genome Biol.* **7**, R100 (2006).

1086 42. J.-H. Zhang, T. D. Y. Chung, K. R. Oldenburg, A Simple Statistical Parameter for
1087 Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol.*
1088 *Screen.* **4**, 67–73 (1999).

1089 43. Z. Liu, O. Chen, J. B. J. Wall, M. Zheng, Y. Zhou, L. Wang, H. Ruth Vaseghi, L.
1090 Qian, J. Liu, Systematic comparison of 2A peptides for cloning multi-genes in a
1091 polycistronic vector. *Sci. Rep.* **7**, 2193 (2017).

1092 44. W.-M. Wang, Z. Liu, A.-J. Liu, Y.-X. Wang, H.-G. Wang, D. An, B. Heng, L.-H. Xie,
1093 J.-L. Duan, Y.-Q. Liu, The Zinc Ion Chelating Agent TPEN Attenuates Neuronal
1094 Death/apoptosis Caused by Hypoxia/ischemia Via Mediating the

1095 Pathophysiological Cascade Including Excitotoxicity, Oxidative Stress, and Inflam-
1096 mation. *CNS Neurosci. Ther.* **21**, 708–717 (2015).

1097 45. H. L. Martin, M. Adams, J. Higgins, J. Bond, E. E. Morrison, S. M. Bell, S. Warriner,
1098 A. Nelson, D. C. Tomlinson, High-Content, High-Throughput Screening for the
1099 Identification of Cytotoxic Compounds Based on Cell Morphology and Cell Prolifer-
1100 ation Markers. *PLOS ONE* **9**, e88338 (2014).

1101 46. D. D. Patel, K. W. Gripp, E. Wadman, I. Mishra, V. Kandula, Novel biallelic ZNF335
1102 variant causing primary microcephaly: A case report and radiological review. *Am. J.*
1103 *Med. Genet. A.* **194**, e63593 (2024).

1104 47. Y. J. Yang, A. E. Baltus, R. S. Mathew, E. A. Murphy, G. D. Evrony, D. M. Gonza-
1105 lez, E. P. Wang, C. A. Marshall-Walker, B. J. Barry, J. Murn, A. Tatarakis, M. A.
1106 Mahajan, H. H. Samuels, Y. Shi, J. A. Golden, M. Mahajnah, R. Shenhav, C. A.
1107 Walsh, Microcephaly Gene Links Trithorax and REST/NRSF to Control Neural
1108 Stem Cell Proliferation and Differentiation. *Cell* **151**, 1097–1112 (2012).

1109 48. D. Li, Z. Quan, J. Ni, H. Li, H. Qing, The many faces of the zinc finger protein 335
1110 in brain development and immune system. *Biomed. Pharmacother.* **165**, 115257
1111 (2023).

1112 49. K. Stouffs, A. B. Stergachis, T. Vanderhasselt, A. Dica, S. Janssens, L.
1113 Vandervore, A. Gheldof, O. Bodamer, K. Keymolen, S. Seneca, I. Liebaers, D.
1114 Jayaraman, H. E. Hill, J. N. Partlow, C. A. Walsh, A. C. Jansen, Expanding the clin-
1115 ical spectrum of biallelic ZNF335 variants. *Clin. Genet.* **94**, 246–251 (2018).

1116 50. R. Sato, J. Takanashi, Y. Tsuyusaki, M. Kato, H. Saitsu, N. Matsumoto, T.
1117 Takahashi, Association Between Invisible Basal Ganglia and ZNF335 Mutations: A
1118 Case Report. *Pediatrics* **138**, e20160897 (2016).

1119 51. A. O. Caglayan, K. Yaghouti, T. Kockaya, D. Kemer, T. Cankaya, N. Ameziane, O.
1120 Cogulu, M. Coker, C. Yalcinkaya, Biallelic ZNF335 mutations cause basal ganglia
1121 abnormality with progressive cerebral/cerebellar atrophy. *J. Neurogenet.* **35**, 23–28
1122 (2021).

1123 52. T. Nishida, T. Mihara, T. Horiki, K. Ka, Anaesthesia and orphan disease: primary
1124 autosomal recessive microcephaly-10 caused by a mutation in the ZNF335 gene.
1125 *Eur. J. Anaesthesiol.* **33**, 543–545 (2016).

1126 53. E. Theusch, F. Y. Ting, Y. Qin, K. Stevens, D. Naidoo, S. M. King, N. V. Yang, J.
1127 Orr, B. Y. Han, J. G. Cyster, Y.-D. I. Chen, J. I. Rotter, R. M. Krauss, M. W. Me-
1128 dina, Participant-derived cell line transcriptomic analyses and mouse studies reveal
1129 a role for ZNF335 in plasma cholesterol statin response. *Genome Med.* **16**, 93
1130 (2024).

1131 54. T. Cremer, M. L. M. Jongsma, F. Trulsson, A. C. O. Vertegaal, J. Neefjes, I. Berlin,
1132 The ER-embedded UBE2J1/RNF26 ubiquitylation complex exerts spatiotemporal
1133 control over the endolysosomal pathway. *Cell Rep.* **34** (2021).

1134 55. T. Cremer, L. M. Voortman, E. Bos, M. L. Jongsma, L. R. ter Haar, J. J. Akker-
1135 mans, C. M. Talavera Ormeño, R. H. Wijdeven, J. de Vries, R. Q. Kim, G. M.
1136 Janssen, P. A. van Veelen, R. I. Koning, J. Neefjes, I. Berlin, RNF26 binds peri-
1137 nuclear vimentin filaments to integrate ER and endolysosomal responses to proteo-
1138 toxic stress. *EMBO J.* **42**, e111252 (2023).

1139 56. L. Lecordier, P. Heo, J. H. Graversen, D. Hennig, M. K. Skytthe, A. Cornet d'Elzius,
1140 F. Pincet, D. Pérez-Morga, E. Pays, Apolipoproteins L1 and L3 control mitochon-
1141 drial membrane dynamics. *Cell Rep.* **42**, 113528 (2023).

1142 57. R. G. Gaudet, S. Zhu, A. Halder, B.-H. Kim, C. J. Bradfield, S. Huang, D. Xu, A.
1143 Mamiñska, T. N. Nguyen, M. Lazarou, E. Karatekin, K. Gupta, J. D. MacMicking, A
1144 human apolipoprotein L with detergent-like activity kills intracellular pathogens. *Sci-*
1145 *ence* **373** (2021).

1146 58. M. Hondele, R. Sachdev, S. Heinrich, J. Wang, P. Vallotton, B. M. A. Fontoura, K.
1147 Weis, DEAD-box ATPases are global regulators of phase-separated organelles.
1148 *Nature* **573**, 144–148 (2019).

1149 59. M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Da-
1150 vis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A.
1151 Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G.
1152 Sherlock, Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29
1153 (2000).

1154 60. K. Shroff, Z. F. Caffall, N. Calakos, DYT-TOR1A subcellular proteomics reveals se-
1155 lective vulnerability of the nuclear proteome to cell stress. *Neurobiol. Dis.* **158**,
1156 105464 (2021).

1157 61. B. Ding, Y. Tang, S. Ma, M. Akter, M.-L. Liu, T. Zang, C.-L. Zhang, Disease Model-
1158 ing with Human Neurons Reveals LMNB1 Dysregulation Underlying DYT1 Dysto-
1159 nia. *J. Neurosci.* **41**, 2024–2038 (2021).

1160 62. S. S. Pappas, C.-C. Liang, S. Kim, C. O. Rivera, W. T. Dauer, TorsinA dysfunction
1161 causes persistent neuronal nuclear pore defects. *Hum. Mol. Genet.* **27**, 407–420
1162 (2018).

1163 63. C.-A. Martin, J. E. Murray, P. Carroll, A. Leitch, K. J. Mackenzie, M. Halachev, A. E.
1164 Fetit, C. Keith, L. S. Bicknell, A. Fluteau, P. Gautier, E. A. Hall, S. Joss, G. Soares,
1165 J. Silva, M. B. Bober, A. Duker, C. A. Wise, A. J. Quigley, S. R. Phadke, T. D. D. D.
1166 Study, A. J. Wood, P. Vagnarelli, A. P. Jackson, Mutations in genes encoding con-
1167 densin complex proteins cause microcephaly through decatenation failure at mito-
1168 sis. *Genes Dev.* **30**, 2158–2172 (2016).

64. M. Trimborn, S. M. Bell, C. Felix, Y. Rashid, H. Jafri, P. D. Griffiths, L. M. Neumann, A. Krebs, A. Reis, K. Sperling, H. Neitzel, A. P. Jackson, Mutations in Microcephalin Cause Aberrant Regulation of Chromosome Condensation. *Am. J. Hum. Genet.* **75**, 261–266 (2004).

65. F. Jean, A. Stuart, M. Tarailo-Graovac, Dissecting the Genetic and Etiological Causes of Primary Microcephaly. *Front. Neurol.* **11** (2020).

66. J. S. Amberger, C. A. Bocchini, F. Schiettecatte, A. F. Scott, A. Hamosh, OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* **43**, D789–D798 (2015).

67. W. S. Rubinstein, D. R. Maglott, J. M. Lee, B. L. Kattman, A. J. Malheiro, M. Ovetsky, V. Hem, V. Gorelenkov, G. Song, C. Wallin, N. Husain, S. Chitipiralla, K. S. Katz, D. Hoffman, W. Jang, M. Johnson, F. Karmanov, A. Ukrainchik, M. Denisenko, C. Fomous, K. Hudson, J. M. Ostell, The NIH genetic testing registry: a new, centralized database of genetic tests to enable access to comprehensive information and improve transparency. *Nucleic Acids Res.* **41**, D925–D935 (2013).

68. K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition. *arXiv*, doi: 1512.03385 (2015).

69. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, “ImageNet: A large-scale hierarchical image database” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255.

70. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv arXiv:1802.03426 [Preprint] (2020). https://doi.org/10.48550/arXiv.1802.03426.

71. T. Hirano, Condensins: Organizing and Segregating the Genome. *Curr. Biol.* **15**, R265–R275 (2005).

72. L. I. Davis, G. Blobel, Identification and characterization of a nuclear pore complex protein. *Cell* **45**, 699–709 (1986).

73. M. M. Drozdz, D. J. Vaux, Shared mechanisms in physiological and pathological nucleoplasmic reticulum formation. *Nucleus* **8**, 34–45 (2016).

74. M. M. Drozdz, H. Jiang, L. Pytowski, C. Grovenor, D. J. Vaux, Formation of a nucleoplasmic reticulum requires de novo assembly of nascent phospholipids and shows preferential incorporation of nascent lamins. *Sci. Rep.* **7**, 7454 (2017).

75. S. Kroschwald, S. Maharana, A. Simon, Hexanediol: a chemical probe to investigate the material properties of membrane-less compartments. *Matters*, doi: 10.19185/matters.201702000010 (2017).

1204 76. C. Mathieu, R. V. Pappu, J. P. Taylor, Beyond aggregation: Pathological phase
1205 transitions in neurodegenerative disease. *Science* **370**, 56–60 (2020).

1206 77. D. M. Mitrea, M. Mittasch, B. F. Gomes, I. A. Klein, M. A. Murcko, Modulating bio-
1207 molecular condensates: a novel approach to drug discovery. *Nat. Rev. Drug Dis-*
1208 *cov.* **21**, 841–862 (2022).

1209 78. A. Patel, D. Mitrea, V. Namasivayam, M. A. Murcko, M. Wagner, I. A. Klein, Princi-
1210 ples and functions of condensate modifying drugs. *Front. Mol. Biosci.* **9** (2022).

1211 79. J. B. Rayman, K. A. Karl, E. R. Kandel, TIA-1 Self-Multimerization, Phase Separa-
1212 tion, and Recruitment into Stress Granules Are Dynamically Regulated by Zn2+.
1213 *Cell Rep.* **22**, 59–71 (2018).

1214 80. V. Singh, L. Xu, S. Boyko, K. Surewicz, W. K. Surewicz, Zinc promotes liquid–liquid
1215 phase separation of tau protein. *J. Biol. Chem.* **295**, 5850–5856 (2020).

1216 81. M. L. M. Jongsma, I. Berlin, R. H. M. Wijdeven, L. Janssen, G. M. C. Janssen, M.
1217 A. Garstka, H. Janssen, M. Mensink, P. A. van Veelen, R. M. Spaapen, J. Neefjes,
1218 An ER-Associated Pathway Defines Endosomal Architecture for Controlled Cargo
1219 Transport. *Cell* **166**, 152–166 (2016).

1220 82. R. Shaheen, S. Maddirevula, N. Ewida, S. Alsahli, G. M. H. Abdel-Salam, M. S.
1221 Zaki, S. A. Tala, A. Alhashem, A. Softah, M. Al-Owain, A. M. Alazami, B. Abadel,
1222 N. Patel, T. Al-Sheddi, R. Alomar, E. Alobeid, N. Ibrahim, M. Hashem, F. Abdul-
1223 wahab, M. Hamad, B. Tabarki, A. H. Alwadei, F. Alhazzani, F. A. Bashiri, A.
1224 Kentab, S. Şahintürk, E. Sherr, B. Fregeau, S. Sogati, S. A. M. Alshahwan, S.
1225 Alkhalifi, Z. Alhumaidi, S. Temtamy, M. Aglan, G. Otaify, K. M. Girisha, M. Tulbah,
1226 M. Z. Seidahmed, M. A. Salih, M. Abouelhoda, A. A. Momin, M. A. Saffar, J. N.
1227 Partlow, S. T. Arold, E. Faqeih, C. Walsh, F. S. Alkuraya, Genomic and phenotypic
1228 delineation of congenital microcephaly. *Genet. Med.* **21**, 545–552 (2019).

1229 83. L. Pavinato, S. Carestiato, S. Trajkova, L. Sorasio, G. Mantovani, L. De Sanctis, J.
1230 Kerkhof, H. McConkey, J. Rzasa, E. Todd, M. Balzo, S. Cardaropoli, A. Bruselles,
1231 S. De Rubeis, J. D. Buxbaum, M. Tartaglia, B. Sadikovic, G. B. Ferrero, A. Brusco,
1232 Skipping of Exon 20 in EP300: A Novel Variant Linked to Rubinstein–Taybi Syn-
1233 drome With Atypical and Severe Clinical Manifestations. *Clin. Genet.* **107**, 354–358
1234 (2025).

1235 84. E. Roessler, E. Belloni, K. Gaudenz, P. Jay, P. Berta, S. W. Scherer, L. C. Tsui, M.
1236 Muenke, Mutations in the human Sonic Hedgehog gene cause holoprosencephaly.
1237 *Nat. Genet.* **14**, 357–360 (1996).

1238 85. C. K. C. Loo, M. A. Pearen, G. A. Ramm, The Role of Sonic Hedgehog in Human
1239 Holoprosencephaly and Short-Rib Polydactyly Syndromes. *Int. J. Mol. Sci.* **22**,
1240 9854 (2021).

1241 86. A. Genin, J. Desir, N. Lambert, M. Biervliet, N. Van Der Aa, G. Pierquin, A. Killian,
1242 M. Tosi, M. Urbina, A. Lefort, F. Libert, I. Pirson, M. Abramowicz, Kinetochore KMN
1243 network gene CASC5 mutated in primary microcephaly. *Hum. Mol. Genet.* **21**,
1244 5306–5317 (2012).

1245 87. Y. A. Ogungbenro, T. C. Tena, D. Gaboriau, P. Lalor, P. Dockery, M. Philipp, C. G.
1246 Morrison, Centrobin controls primary ciliogenesis in vertebrates. *J. Cell Biol.* **217**,
1247 1205–1215 (2018).

1248 88. P. Rump, O. Jazayeri, K. K. van Dijk-Bos, L. F. Johansson, A. J. van Essen, J. B.
1249 G. M. Verheij, H. E. Veenstra-Knol, E. J. W. Redeker, M. M. A. M. Mannens, M. A.
1250 Swertz, B. Z. Alizadeh, C. M. A. van Ravenswaaij-Arts, R. J. Sinke, B. Sikkema-
1251 Raddatz, Whole-exome sequencing is a powerful approach for establishing the eti-
1252 ological diagnosis in patients with intellectual disability and microcephaly. *BMC*
1253 *Med. Genomics* **9**, 7 (2016).

1254 89. H. Schwaemmle, H. Soldati, N. M. R. Lykoskoufis, M. Docquier, A. Hainard, S. M.
1255 G. Braun, CRISPR screen decodes SWI/SNF chromatin remodeling complex as-
1256 sembly. *Nat. Commun.* **16**, 5011 (2025).

1257 90. P.-L. Tsai, C. J. F. Cameron, M. F. Forni, R. R. Wasko, B. S. Naughton, V. Horsley,
1258 M. B. Gerstein, C. Schlieker, Dynamic quality control machinery that operates
1259 across compartmental borders mediates the degradation of mammalian nuclear
1260 membrane proteins. *Cell Rep.* **41**, 111675 (2022).

1261 91. P.-L. Tsai, C. Zhao, C. Schlieker, Methodologies to monitor protein turnover at the
1262 inner nuclear membrane. *Methods Enzymol.* **619**, 47–69 (2019).

1263 92. R. Ihaka, R. Gentleman, R: A Language for Data Analysis and Graphics. *J. Com-*
1264 *put. Graph. Stat.* **5**, 299–314 (1996).

1265 93. L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on
1266 immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 583–598 (1991).

1267 94. S. Marcel, Y. Rodriguez, "Torchvision the machine-vision package of torch" in *Pro-*
1268 *ceedings of the 18th ACM International Conference on Multimedia* (ACM, Firenze
1269 Italy, 2010; https://dl.acm.org/doi/10.1145/1873951.1874254), pp. 1485–1488.

1270 95. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M.
1271 Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg,
1272 J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay,
1273 Scikit-learn: Machine Learning in Python. arXiv arXiv:1201.0490 [Preprint] (2018).
1274 https://doi.org/10.48550/arXiv.1201.0490.

1275 96. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin,
1276 N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A.
1277 Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, PyTorch: An

Imperative Style, High-Performance Deep Learning Library. arXiv arXiv:1912.01703 [Preprint] (2019). https://doi.org/10.48550/arXiv.1912.01703.

97. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, S. 1 0 Contributors, SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **17**, 261–272 (2020).

98. C. Ritz, F. Baty, J. C. Streibig, D. Gerhard, Dose-Response Analysis Using R. *PLOS ONE* **10**, e0146021 (2015).

99. H. Wickham, ggplot2. *WIREs Comput. Stat.* **3**, 180–185 (2011).

100. T. Wu, E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, X. Fu, S. Liu, X. Bo, G. Yu, clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2** (2021).

101. T. Galili, A. O'Callaghan, J. Sidi, C. Sievert, heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* **34**, 1600–1602 (2018).