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. Elsiena 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 neurological disorders including ALS, frontotemporal dementia, and DYT1 dystonia, yet approaches to systematically identify their modulators remain limited. Here we establish MLF2 as a versatile condensate biomarker and develop CondenScreen, an integrated high-content screening and bioinformatic pipeline enabling identification of condensate modulators across chemical and genetic space. Screening 1,760 FDA-approved compounds in a cellular DYT1 dystonia model, we identify drugs that alter aberrant condensate properties, validating the platform for condensate-targeted drug discovery. In parallel, a genome-wide CRISPR/Cas9 screen links condensate accumulation to microcephaly genes and more than 10 additional neurodevelopmental disorders. Machine learning and confocal imaging resolves distinct condensate phenotypes: loss of microcephaly-associated ZNF335 results in nucleoplasmic condensates, whereas RNF26 deletion produces nuclear envelope condensates that phenocopy hallmarks of torsin deficiency. Our study provides a scalable resource for identifying corrective modulators of condensates and establishes a link between nuclear condensate accumulation and neurodevelopmental disorders.

Keywords

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

Introduction

Dysregulation of physiological condensates is a driver for many neurological disorders, including Amyotrophic Lateral Sclerosis (ALS), Frontotemporal Dementia (FTD), and DYT1 dystonia.[1–5] Biomolecular condensates, often made from a milieu of hundreds of different proteins, consist of a thermodynamically favorable separation between a dense and dilute phase.[6,7] When properly regulated, biomolecular condensates allow efficient, fine-tuned microenvironments, 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, leading to protein aggregation or amyloid formation that is not easily reversible.[11–13] It is estimated that up to one-third of common diseases can be linked to condensate-forming proteins.[14] Determining mechanisms and cellular activities of these processes 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 selective phase, contributing to the selective barrier of the NPC.[15–18] Disruptions to NPC assembly or nucleocytoplasmic transport have been strongly implicated in neurological diseases, including DYT1 dystonia,[1,19,20] a debilitating movement disorder characterized by involuntary muscle contractions and painful sustained postures.[21,22]

DYT1 dystonia is caused by a mutation within the TOR1A gene that disrupts activation of the encoded AAA+ ATPase,[23–27] TorsinA, resulting in the formation of pathological nuclear envelope condensates.[1,19,20,28,29] As with many condensatopathies, there are only palliative interventions, with no curative or preventative treatments for dystonia.[21,22] These nuclear envelope condensates accumulate K48-linked polyubiquitinylated proteins that evade proteasomal degradation and sequester chaperone proteins, causing cellular proteotoxicity.[1,20,30] Thus, regulating aberrant

nuclear condensates by either direct dissolution or modulation of its proteotoxic properties represents a practical therapeutic approach. However, current 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 establish Myeloid leukemia factor 2 (MLF2) as a robust reporter that is enriched in proteotoxic nuclear condensates and stress granules, as evidenced by co-localization with TDP-43 and G3BP1. Using an MLF2 readout, we integrated complementary chemical and genetic high-content screening strategies to define the landscape underlying atypical nuclear condensates. First, a target-agonistic chemical screen in a cell-based DYT1 dystonia model identified several FDA-approved small molecules that modulate proteotoxic nuclear condensates, providing a tractable methodological framework for identifying therapeutics for condensatopathies. Second, a CRISPR/Cas9 knockout (KO) screen in wild-type cells, combined with deep learning approaches, identified RNF26 and ZNF335 as protective factors against distinct condensate phenotypes. Notably, 15 out of the top 50 genetic hits are associated with neurodevelopmental disorders, including seven directly linked to primary microcephaly or related congenital brain malformations. Together, these orthogonal approaches converge on a shared connection between nuclear condensate accumulation, perturbed nuclear proteostasis, and neurodevelopmental disorders, while establishing a scalable platform for therapeutic discovery.

Results

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

Having established the connection between impaired protein quality control and condensates in a torsin-deficient DYT1 dystonia HeLa cell model (4TorKO cells) (Figure 1A),[1,20] 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[20] and readily immerses into FG-nup condensates *in vitro*,[1] likely due to its highly enriched Methionine and Arginine residues. Since many condensates share aromatic side chains as part of

their molecular grammar,[31] we previously suggested that MLF2 might localize to additional aberrant condensates through π-cation interactions as MLF2 is highly enriched in Arg residues that cannot be functionally replaced by charge-conserving Lys residues.[1] Here, we have exploited these properties to visualize proteotoxic condensates by designing doxycycline (Dox) inducible fluorophore-derivatized variants of MLF2 as condensate biomarkers.

Dysregulation of stress granules (SGs) has been pathologically associated with several neurological disorders, including ALS and FTD.[11] To assess MLF2 localization to stress granules, we co-expressed the SG markers G3BP1 and TDP-43 in wild-type (WT) HeLa cells exposed to 0.5 mM sodium arsenite ($NaAsO_2$) for 30 minutes, an established condition for SG induction.[32–35] Upon $NaAsO_2$ exposure, MLF2 exhibited strong co-localization with both G3BP1-GFP and TDP-43-tdTomato (Figure 1B). These findings suggest that MLF2 serves as a robust reporter of condensates associated with distinct neurological disorders.

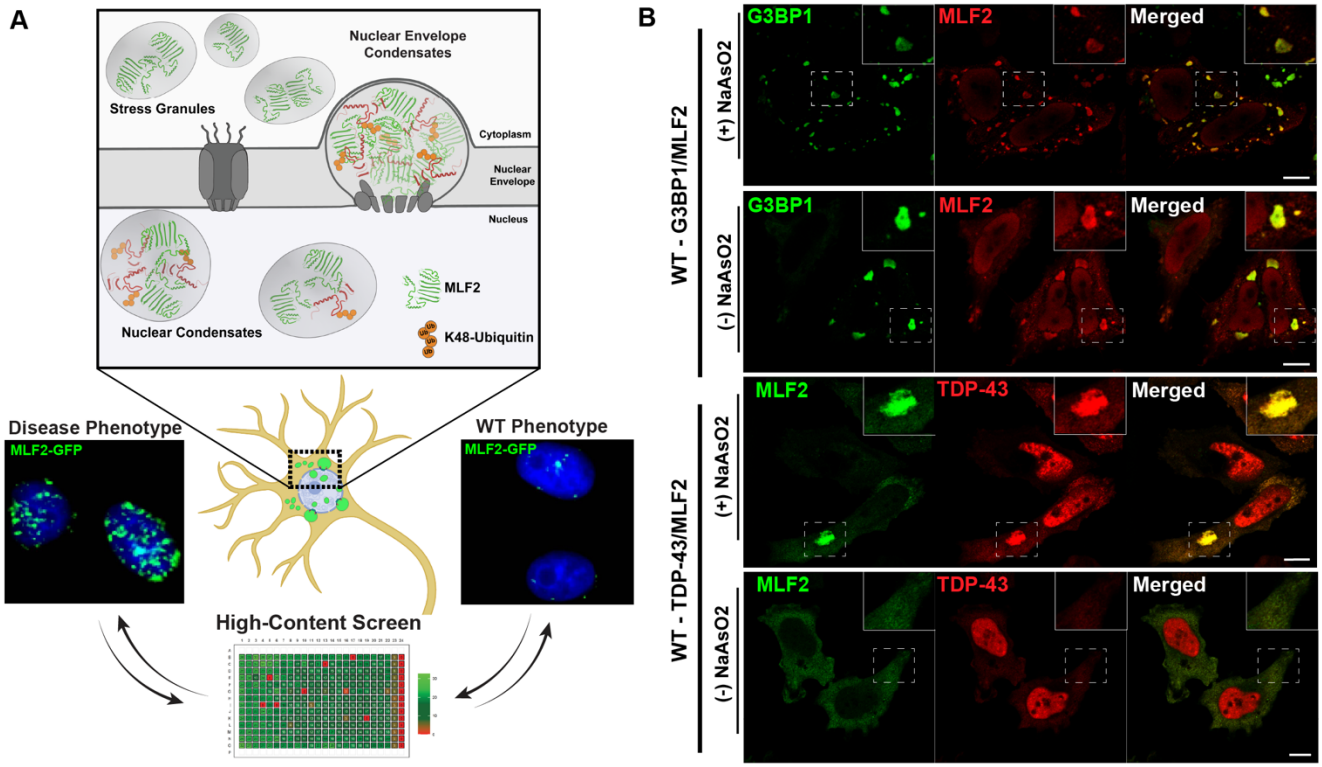

Figure 1, *Poch et al.*

Fig. 1. MLF2 is a biomarker for aberrant condensates across neurological disorders. (**A**) MLF2 detects both nuclear condensates and stress granules. Nuclear envelope condensates that sequester MLF2, FG-nups (not shown for clarity), and K48-Ubiquitin-enriched proteins have been identified in DYT1 dystonia models. The ability of MLF2 to immerse into condensates can be leveraged for high-content imaging platforms to systematically profile additional disease-relevant condensates. (**B**) MLF2 co-localizes with G3BP1 (top panels) and TDP-43 (bottom panels) after stress granule formation in response to oxidative stress (0.5 mM $NaAsO_2$, 30 min).

76

77 **Establishing MLF2-GFP as a Condensate Reporter for High-Content Screening**

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

88 While many robust computational resources are available for monitoring
89 condensates (see *PhaseMetrics* and *Punctatools*)[36,37], *CondenScreen* is specifically
90 designed for large-scale, high-content screening analysis using a customizable
91 *CellProfiler 4* [38,39] pipeline and *R* analysis script. To this end, *CondenScreen* quantifies
92 condensates at both single-cell and population levels and provides a graphical user
93 interface (GUI) allowing users to select plate size and layout, as well as tailoring
94 downstream calculations for a signal “ON” vs signal “OFF” screening campaign. In
95 addition, built-in quality control features automatically assess screen robustness using
96 established statistical metrics, including the Z’-Score, BZ-Score, signal-to-background
97 (SB), normalized effect, and coefficient of variation (CV). In addition, the platform has
98 optional filters to flag potential false positives arising from either spatial plate effects
99 (Figure S1B-S1D) and filter candidate compounds with known false positive liabilities,
100 including pan-assay interference compounds (PAINS) and additional structural alert
101 categories (Figure S1E). Integration of the MLF2-GFP inducible system with our
102 *CondenScreen* analysis pipeline ensures rapid identification of condensate modulators.

103

104 **High-Content Screen Identifies FDA-approved Drugs that Modulate Condensates**

105 To identify chemical matter that can modulate aberrant nuclear envelope condensates,
106 we employed our MLF2-GFP inducible 4TorKO DYT1 cell model. We screened 1760

107 drugs from the Pharmakon library, comprising 1360 US and 400 international drugs, at a
108 10 μM working concentration. To minimize false positives due to differential reporter
109 expression, Dox (0.5 μg/ml) was added for four hours, followed by an additional 2 hours
110 with compound incubation.

111 *CondenScreen* reported an average Z'-score of 0.75 indicating pronounced
112 separation between the positive (0.1% DMSO-treated WT HeLa) and negative (0.1%
113 DMSO-treated 4TorKO HeLa) controls.[40] Additionally, the ratio between the negative
114 and positive controls (SB) was 950, while the SD and CV of DMSO-treated cells were
115 1.9% and 8.3%, respectively, further confirming assay robustness.

116 To control for plate-specific column, row, and edge effects, we assigned each
117 compound a BZ-score based on its ability to decrease the number of nuclear MLF2-
118 GFP foci (Figure 2A, see methods). The distribution of the average number of MLF2-
119 GFP foci per condition was near-normally distributed (Figures 2B, 2C). Using a cutoff
120 BZ-score of three, we identified 20 putative hit compounds that significantly depleted
121 MLF2-GFP condensate foci. Top compounds included plumbagin, sanguinarium
122 chloride, thiram, piroctone olamine, penfluridol, bronopol, thimerosal, triclosan,
123 disulfiram, celastrol, dihydrocelastryl diacetate, and pyrithione zinc (PZ). A full list of the
124 screened drugs can be found in the supplementary materials (Table S1).

125 Top-scoring compounds were resolubilized to construct a 12-point dose response
126 (n = 4 replicates), covering over 5 orders of magnitude to determine the calculated IC_{50}.
127 (Figure 2D, Figure S2A). The normalized percent effect and percent toxicity (calculated
128 as a function of viability, nuclear area, and nuclear morphology) were quantified for each
129 concentration to determine potential therapeutic windows. Four structurally diverse
130 drugs demonstrated a potent dose-dependent relationship, which included celastrol
131 (IC_{50} = 124.8 nM), PZ (IC_{50} = 601.6 nM), plumbagin (IC_{50} = 7.4 μM), and disulfiram (IC_{50}
132 = 10.8 μM) (Figure 2D).
133
134

135

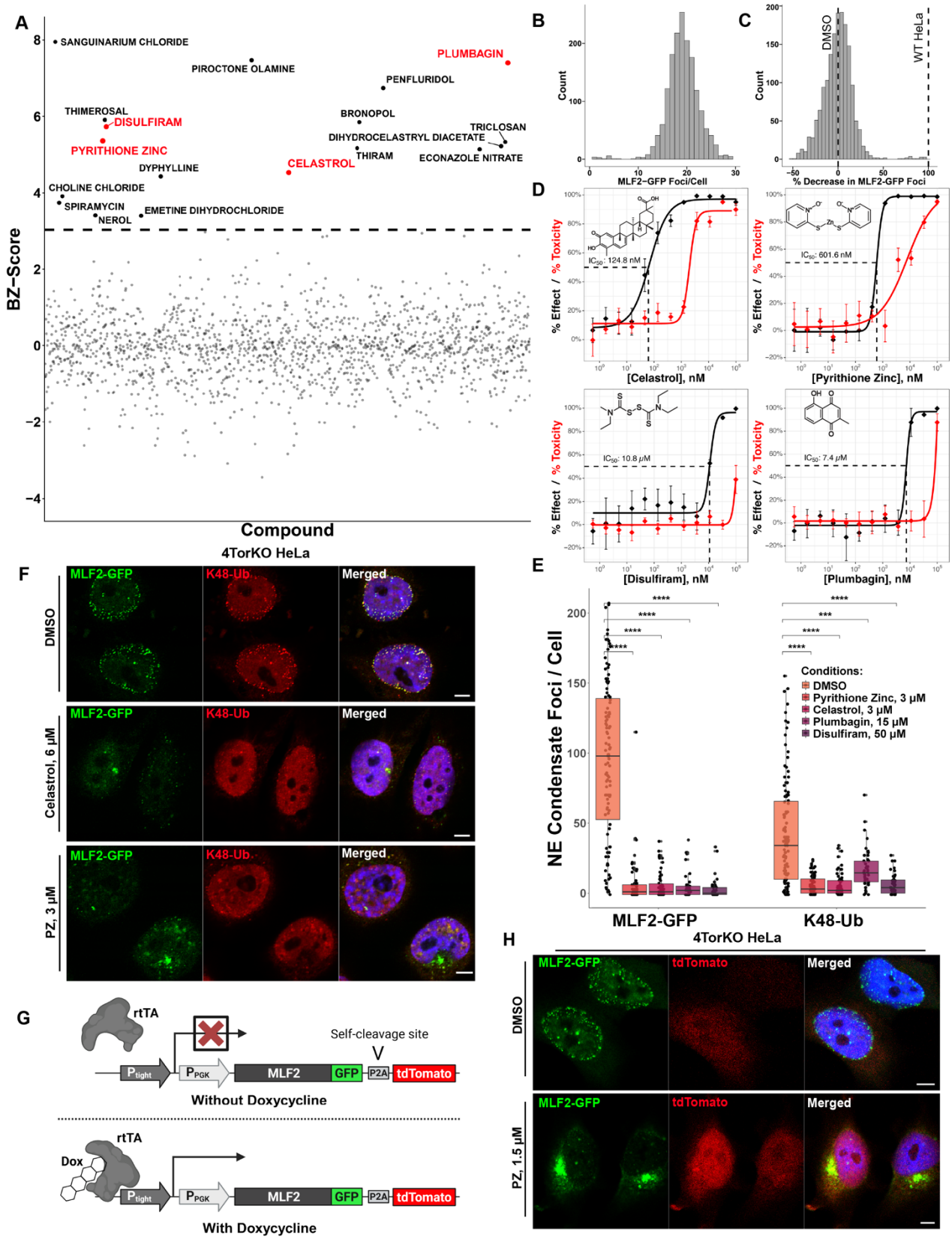

Figure 2, *Poch et al.*

Fig. 2. Chemical screen identifies drugs that modulate nuclear envelope condensates. (A) Drugs from Pharmakon 1760 screen decrease 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 tighter 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 S2B). Consistent with the significant cell-cycle dependency of NE condensate formation previously demonstrated,[20] 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 S2B).

To ensure consistent reduction across different condensate markers, we used K48-linked ubiquitin (K48) as an independent and endogenous biomarker based on our previous observation that 4TorKO cells have reduced protein turnover, resulting in the accumulation of polyubiquitinylated proteins within nuclear envelope condensates.[1,20] 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). Between the optimal concentration profile of PZ and celastrol (Figure 2D), and the undesirable structural elements of plumbagin (quinone, Michael acceptor; Figure S1F), we continued to validate PZ and celastrol by assessing 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 condensates.

161 A significant limitation to phenotype screening is the prevalence of identification
162 of false positives that inhibit biomarker expression by non-specifically disrupting cellular
163 protein transcription or translation. To circumvent this, we constructed a HeLa cell line
164 with an inducible bicistronic MLF2-GFP-P2A-tdTomato reporter (Figure 2G). The P2A
165 site induces ribosomal skipping, allowing the translation of two separate polypeptides,
166 an N-terminal MLF2-GFP and a C-terminal tdTomato, both derived from the same
167 mRNA transcript.[41] We selected MLF2-GFP-P2A-tdTomato 4TorKO cells using
168 Fluorescence-Activated Cell Sorting (FACS) (Figure S2C), before inducing MLF2-GFP
169 expression and treating cells with either PZ (0.75 μM or 1.5 μM, 6 h) or celastrol (1.5 μM
170 or 3 μM, 6 h) . We quantified the number of MLF2-GFP foci and tdTomato expression
171 between either uninduced cells treated with DMSO vehicle control, induced cells treated
172 with DMSO, and induced cells treated with either 100 μg/mL cycloheximide, PZ, or
173 celastrol. Concentrations of both celastrol and PZ at 1.5 μM had a significant decrease
174 in the number of MLF2-GFP foci/cell compared with the DMSO control (Figure S2D).
175 While PZ did not reduce the expression of secondary-tdTomato reporter (Figures 2H,
176 S2E), celastrol slightly, albeit at a non-statistically significant level, decreased secondary
177 reporter expression compared with DMSO-treated cells.
178 To address and quantify the potentially confounding effects of a compound
179 affecting reporter abundance non-specifically, we developed an empirical metric, the
180 Condensate Normalization Score (CNS). The CNS is based on the differential
181 fluorescence between the condensate biomarker and secondary reporter. A high CNS-
182 value indicates a reduction in condensate biomarkers with minimal impact on overall
183 protein levels (see methods). Both celastrol- and PZ-treated cells showed a significantly
184 increased percentage of cells with a median CNS exceeding 5-fold that of the DMSO
185 control (Figure S2F), indicating a mechanism independent of reporter expression.
186 To assess the specificity of the top two compounds in modulating aberrant
187 condensates, we examined their effects on nuclear Promyelocytic Leukemia (PML)
188 bodies, physiological phase-separated compartments involved in protein quality control.
189 As expected, treatment with 5% 1,6-hexanediol (5 minutes), a known disrupter of liquid-
190 liquid phase-separated environments,[42] significantly reduced PML body number (Figure
191 S2H, S2G). In contrast, neither PZ nor celastrol reduced PML body number at a

192 concentration of 1.5 μM that profoundly reduced MLF2 foci in our disease model (Figure
193 S2H, S2G). Interestingly, treatment with PZ resulted in a statistically significant increase
194 to the number of PML body puncta. Overall, this indicates that the activity of PZ and
195 celastrol does not reflect nonspecific disruption of physiological nuclear condensates.
196 However, visualization of condensates in 4TorKO cells using transmission electron
197 microscopy (TEM) did not reveal full dissolution of nuclear envelope condensates after
198 a 3 h treatment of 2 μM PZ (Figure S2I).
199 To identify the structural features responsible for the condensate-modifying
200 activity of PZ, we examined whether this effect originates from its pyrithione moiety (2-
201 Mercaptopyridine N-oxide) or the zinc ion. To this end, we treated 4TorKO cells with
202 either DMSO vehicle control, PZ alone, or PZ with zinc chelator TPEN
203 (N,N,N',N'- *tetrakis*-(2-Pyridylmethyl)ethylenediamine) at a 3 molar equivalent.[43] After
204 zinc chelation with TPEN, the pyrithione compound failed to reduce MLF2-GFP
205 condensate accumulation in 4TorKO cells (Figures S3A, S3B). Furthermore, we tested
206 whether zinc alone is sufficient to deplete nuclear MLF2-GFP condensate foci. Due to
207 its lower membrane permeability, 4TorKO cells were treated with higher concentrations
208 of zinc sulfate monohydrate ($ZnSO_4$). Treatment with 2 mM $ZnSO_4$ alone decreased
209 MLF2-GFP foci (Figure S3C, S3D) while maintaining tdTomato expression (Figure S3E),
210 resulting in a higher relative CNS value (Figure S3F). Thus, PZ modulates nuclear
211 condensates through a zinc-dependent mechanism.
212

213 **Genome-Wide CRISPR Screening Uncovers Condensate Modulators Implicated in**
214 **Microcephaly and Neurodevelopmental Disorders**

215 Having established a robust screening pipeline, we next performed a genome-
216 wide CRISPR/Cas9 KO screen to identify genetic determinants of nuclear condensate
217 accumulation. Knowledge of the corresponding activities may facilitate the identification
218 of future pharmacological targets. We screened an arrayed library of 19,195 human
219 genes using the Horizon Discovery Edit-R synthetic sgRNA collection. To increase KO
220 efficiency, we designed and FACS-sorted a WT HeLa cell line that constitutively
221 expresses Cas9 with Dox-inducible MLF2-GFP. Each well was reverse-transfected in a
222 384-well plate format using a pool of three sgRNAs targeting different regions of the

223 same gene (Figure 3A). Cells were incubated for 72 h, and MLF2-GFP was induced by
224 Dox 16 h before fixation (all within the 72 h window). After running all high-content
225 images through our *CondenScreen* pipeline, the average Z' value was 0.52,
226 demonstrating a dynamic range between the positive (TOR1A/B sgRNAs) and negative
227 (non-targeting (NT) sgRNA) controls. The calculated SB was 4.8, and the SD and CV of
228 the NT control were 3.1 and 9.7, respectively. We determined the assay quality score to
229 be sufficiently robust as indicated by similar scores in the literature.[34,44]

230 We defined a candidate gene as any gene deletion that resulted in a Z-score > 2
231 (Figure 3B). The top 10-ranked candidate genes were: RNF26, WAC, ZNF335,
232 NCAPG2, NCAPH2, ECT2, ASRGL1, SETD1B, ASB6, and FOXS1 (see Table S2 for
233 comprehensive list). Overall, the screen resulted in a near-normal distribution of the
234 number of MLF2-GFP foci per cell (Figures 3C, 3D). Notably, we observed several
235 genes linked to neurodevelopmental disorders, with high enrichment of microcephaly-
236 associated genes (orange labeled dots, Figure 3B), according to human genetics
237 annotated from OMIM[45] and the NIH Genetic Testing Registry (GTR).[46] Both
238 homozygous and compound heterozygous mutations to one of our top hits, zinc finger
239 protein 335 (ZNF335, also identified as MCPH10 or NIF1) have been documented to
240 cause primary microcephaly.[47–54] ZNF335 has been tied to the regulation of
241 neurogenesis through modulation of the REST complex.[48]

242 Additional top hits of interest include, RNF26, an E3-ubiquitin ligase that has
243 been implicated in regulating the endosomal system, serving as a "gatekeeper" for
244 vesicles in the perinuclear ER cloud.[55,56] Furthermore, RNF26 ubiquitinylates
245 SQSTM1[55] and has recently been tied to ER-sheet morphology and maintenance.[56]
246 Other hits include APOL3 (Z-Score = 2.5), a putative regulator of membrane dynamics
247 with detergent-like bactericidal properties;[57,58] Ran (Z-Score = 2.8), a GTP-binding
248 nuclear protein; and EIF4A3/DDX48 (Z-Score 3.8), a DEAD-box ATPase. DEAD-box
249 ATPases have recently been linked to the regulation of phase-separated membraneless
250 organelles,[59] indicating a potential role of EIF4A3 in modulating condensates. These
251 condensate phenotypes observed across gene deletions highlight critical entry points
252 into unexplored biology and promising directions for therapeutic discovery.

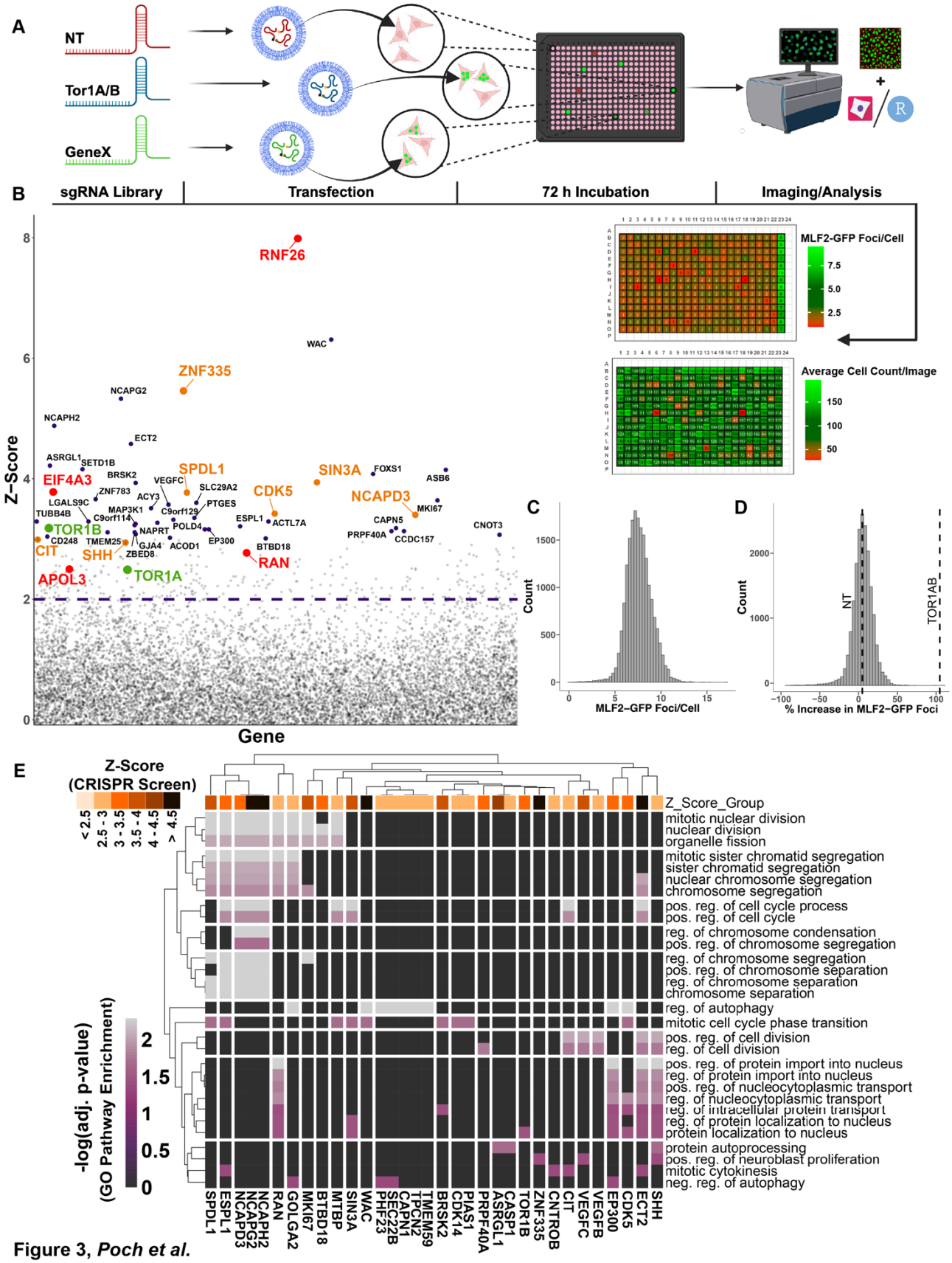

Figure 3, *Poch et al.*

253

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 larger, 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.

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[60] (Figure 3E; interactive webpage available, see methods) and illustrated pathway interconnectedness (Figures S4A-S4C). We observed enrichment (10 out of 30 GO terms) of chromosomal and chromatid condensation, separation, or segregation pathways (Figures 3E, S4A). GO pathways were manually clustered into the following five categories: Autophagy, Cell Cycle, Chromosomal Regulation, Protein Transport, and Other (Figure S4B). Consistent with these enriched pathways, loss of protein homeostasis,[1,61] a cell-cycle dependency of condensates,[20] and disruption to nucleocytoplasmic transport[62,63] have been documented in condensatopathies. Additionally, dysregulation of both cell cycle checkpoints and chromosomal regulation has been linked to primary microcephaly.[64–66] Indeed, seven out of the top 50 gene hits have been linked to microcephaly.[45,46] Beyond the enrichment of individual gene hits linked to microcephaly, encephalopathies, and other neurodevelopmental disorders, additional enriched pathways of interest included organelle fission and positive regulation of neuroblast proliferation.

To validate the three highest candidate genes, we knocked out ZNF335, RNF26, and the WW Domain Containing Adaptor with Coiled-Coil (WAC), and binned them according to the number of nuclear MLF2-GFP foci/cell (5-15, 15-30, or >30 foci/cell) (Figure S4D). While WAC depletion had only a modest effect, RNF26 KO and ZNF335 KO cells displayed pronounced nuclear condensate accumulation as indicated by MLF2-GFP accumulation (Figure S4D), with ZNF335 deficient cells resulting in the greatest number of condensates.

Supervised Machine Learning Links Genetic Perturbations to Distinct Condensate Phenotypes

To resolve phenotypic distinctions and cluster genetic perturbations by their condensate morphology, potentially highlighting divergent biological roles, we developed a deep learning model to capture subtle morphological variations between condensates. Using a residual convolutional neural network (Figure S5A),[67] we trained the model on single-cell patches derived from our strongest validated genetic hits to determine whether condensates differences could be computationally discriminated (Figure 4A). We employed a ResNet18 neural network backbone (ImageNet)[68] to learn a regression mapping across three reference states: RNF26 KO (-1), non-targeting (0), and ZNF335 KO (+1). The model assigned each unlabeled cell a continuous phenotypic similarity score ranging from −1 (RNF26-like) to +1 (ZNF335-like)

After optimization, the model correctly classified 70% of individual cells to their respective anchor conditions (Figure 4B), demonstrating strong performance metrics (area under the curve [AUC] = 0.95, loss = 0.14, true vs. predicted correlation = 0.76; Figures 4C, S5B–C). Notably, fewer than 2% of RNF26 KO cells were misclassified as ZNF335 KO and vice versa (Figure 4E), demonstrating the model's ability to discriminate and report biologically meaningful results. To confirm that predictions were driven by biologically interpretable morphological features, we employed post-hoc saliency mapping based on the integrated gradient (see methods, Figure S5F, S5G).[69] Consistent with this, the RNF26 KO cells resulted in a higher number of smaller, more heterogeneous nuclear condensates when compared to the ZNF335 KO condition (Figure S5G).

Independent of the model's continuous regression output, we extracted raw feature embeddings for each CRISPR hit with a Z-score > 2 and averaged single-cell embeddings per condition to generate centroid vectors. These were projected into two dimensions via Uniform Manifold Approximation and Projection (UMAP), [70] revealing three distinct clusters by KMeans clustering (Figures S5D, S5E). The resulting UMAP visualization (Figure 4F) separated genes into groups phenotypes resembling the ZNF335 KO cells (cluster 1), RNF26 KO cells (cluster 2), or the NT control cells (cluster 3).

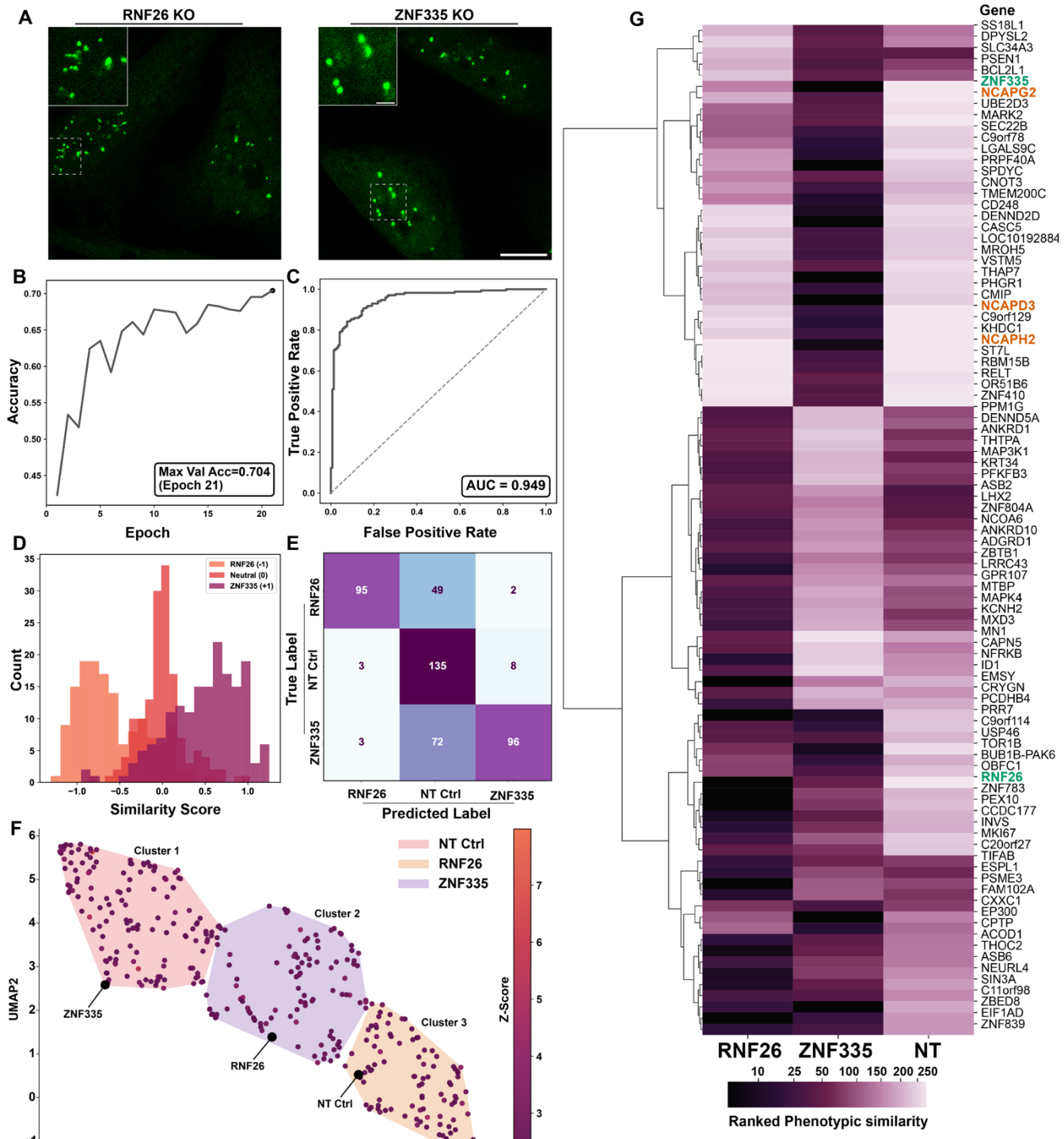

Figure 4, *Poch et al.*

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 model during training. Accuracy reflects the percentage of validation cells correctly classified as RNF26 KO, ZNF335 KO, or NT control. 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 0.95. (**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 (ZNF335-like), cluster 2 (RNF26-like), and cluster 3 (NT-like). (**G**) Hierarchical clustering heatmap showing phenotypic similarity of the top 50 gene knockouts to RNF26 or ZNF335 KO anchors. Darker 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 orange.

Finally, centroid feature vectors were used to compute the similarity metrics for each gene against the RNF26 KO, ZNF335 KO, and NT control conditions. 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 or ZNF335 KO conditions (Figure 4G). Of interest, knocking out any of the non-SMC subunits of the condensin II complex (NCAPG2, NCAPD3, and NCAPH2)[71], and to a lesser extent, the condensin II SMC2 subunit, displayed condensates with high phenotypic similarity to the ZNF335 KO condition (Figure 4G).

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

To validate the deep learning results suggesting that condensates arising from knocking out RNF26 and ZNF335 are phenotypically distinct, we used immunofluorescence staining to determine if these condensates accumulate polyubiquitinylated proteins. Similar to those observed in torsin-deficient cells (see chemical screen), 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 NPC-localized nucleoporins. RNF26 KO cells showed prominent MLF2-GFP co-localization with mAb414 (Figure 5B), analogous to that observed in torsin-deficient cells with stalled NPCs (Figure 5B). In contrast, analysis of confocal Z-stacks revealed that condensates in ZNF335 KO cells rarely shared significant co-localization with mAb414 staining

337 (Figure 5B, 5C), supporting the prediction that RNF26 and ZNF335 KO conditions
338 promote condensates with distinct properties.
339 Consistent with immunofluorescence data, TEM imaging demonstrated that
340 RNF26 KO cells contain condensates between the nuclear envelope (Figure 5D,
341 shaded yellow). Additionally, RNF26 KO cells feature pronounced NE invaginations
342 reminiscent of nucleoplasmic reticulum formation (Figure 5D).[72,73] While we observed
343 several NE invaginations contiguous with the inner nuclear membrane (INM), further
344 electron tomographic studies are needed to determine if all invaginations are contiguous
345 with the NE. We then treated RNF26, ZNF335, or TOR1A/B depleted cells with 5% 1,6-
346 hexanediol for five minutes. Only 1,6-hexanediol, but not DMSO or 2,5-hexanediol,
347 dissolved all three types of condensates (Figure S6). Together, these results indicate
348 that while both RNF26 and ZNF335 have a role in safeguarding against nuclear
349 condensate accumulation, only lack of RNF26 results in a condensate phenotype
350 similar to the one observed upon torsin deletion.

351 **Chemical hit Pyrithione Zinc Modulates Condensates Induced by Genetic**
352 **Knockouts.**
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 NE condensates in RNF26 KO and ZNF335 KO cells. In both KO
358 conditions, treatment with 2 μM PZ significantly reduced accumulation of nuclear MLF2-
359 GFP condensate foci (Figures 5E, S7). Thus, FDA-approved drugs, including PZ, can
360 be repurposed to modulate condensates across phenotypically distinct genetic
361 perturbations.
362
363
364
365

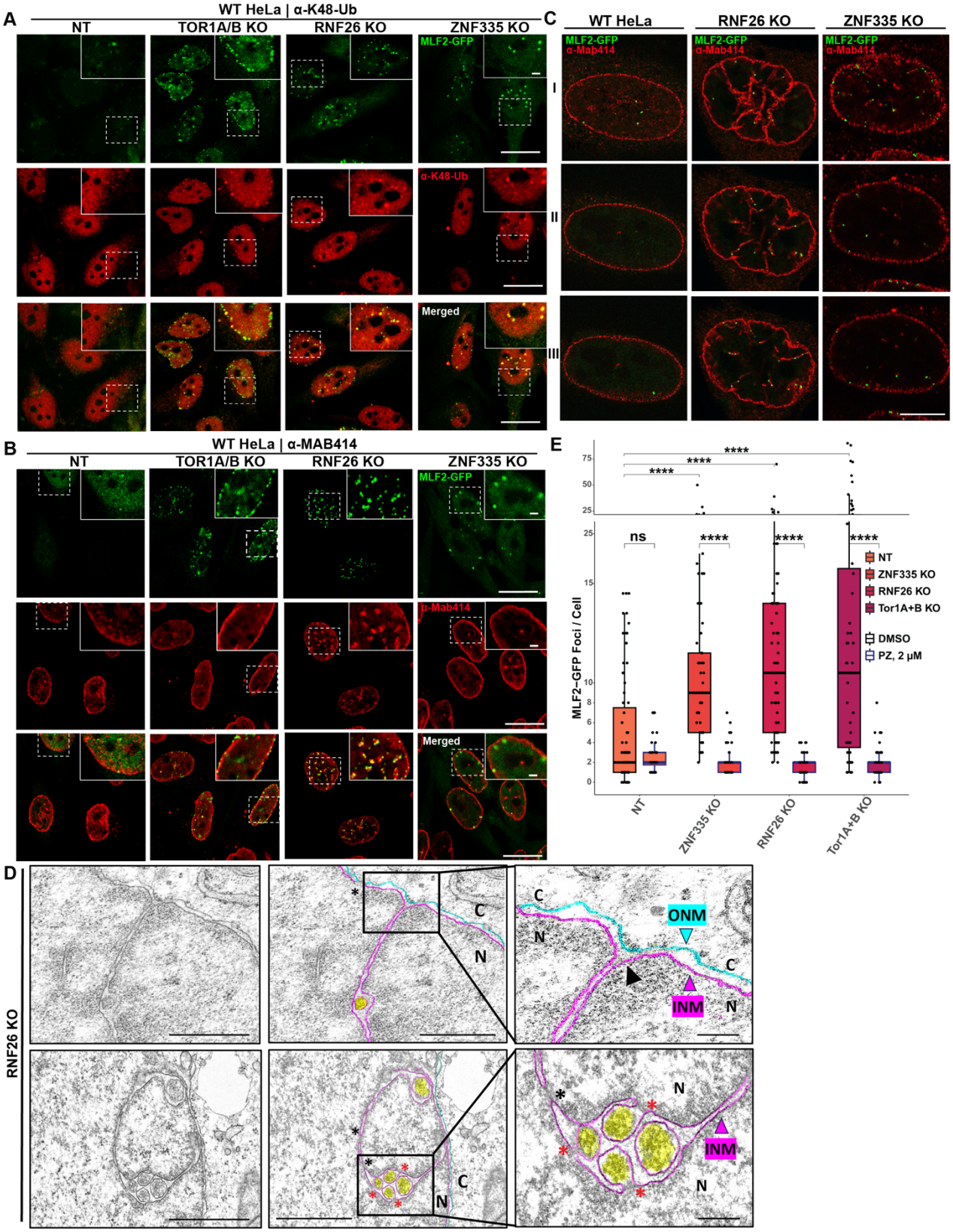

Figure 5, *Poch et al.*

366

Fig. 5. Knockout of RNF26 or ZNF335 induces nuclear condensate accumulation. (**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 (blue outline). 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 scalable high-content platform optimized for identifying chemical and genetic modulators of aberrant condensates that are increasingly implicated in neurological disorders.[12,14,74] We paired a chemical screen in a DYT1 dystonia disease model with a genome-wide CRISPR KO screen in WT cells enabling complementary analysis of condensate modulation. This integrated approach identified FDA-approved compounds that alter the properties of aberrant condensates in a disease model and uncovered genes that are tied to nuclear condensate homeostasis in WT cells. Notably, several genes identified through this study are mutated in patients with microcephaly, dystonia, and related neurodevelopmental disorders, highlighting previously unappreciated commonalities in the cellular pathology underlying these conditions. In addition, our study identifies previously unknown condensate-modifying genes which may inform both patient-based sequencing and mechanistic studies.

Using the known proclivity of MLF2 to immerse into aberrant condensates,[1] we exploit MLF2 derivatives as phenotypic condensate reporters. Remarkably, our reporter allows for detection of DYT1 dystonia-associated condensates under single torsin deletion conditions (Fig 3B) that remain undetectable by conventional markers.[28] In addition, we generalize the utility of MLF2-based reporters for stress granules linked to ALS/FTD (Figure 1B). Notably, MLF2 was recently found as resident in Poly-GA and phosphorylated-TDP-43 inclusions in C9orf72 ALS/FTD patient samples,[75] as well as co-localization of its *Drosophila* homolog (dMLF) with polyglutamine aggregates[76,77], suggesting that our readout can be applied to a variety of neuropathological conditions in the future.

391 Importantly, our screening approach is target-agnostic, utilizing MLF2 solely as a
392 readout for condensate detection at low expression levels to avoid altering condensate
393 dynamics. This allows our platform to be readily adapted to support a range of
394 condensate screening campaigns, even when the affected disease allele is not itself a
395 condensate constituent. This is exemplified by TOR1A/B or ZNF335 KO conditions,
396 which induce condensates via indirect mechanisms, representing a key advantage
397 since condensates are often studied under conditions of disease-allele overexpression,
398 which may perturb physiological condensate equilibria.
399 Inspired by recent condensate-modifying (c-mod) campaigns,[34,35,78,79] we
400 screened drugs that could modulate condensates either directly or indirectly. Several of
401 our identified drug hits caused robust re-localization of both our MLF2-GFP biomarker
402 and endogenous (K48-Ub) nuclear condensate markers (Figures 2E, 2F), without
403 significantly altering reporter expression or non-specifically altering physiological phase-
404 environments (e.g., PML bodies, Figure S2G, S2H). To assess the ability to repurpose
405 our chemical hits for other condensatopathies, we compared our top 20 drug hits with
406 two recent screens that identified modulators of stress granules.[34,35] Half of our drug hits
407 (10 out of 20), including PZ, celastrol, and disulfiram were identified among the top 100
408 hits of these stress granule screens (Table S4).[34,35] The ability of identical chemical
409 matter to modulate disparate disease phenotypes underscores the potential of shared
410 etiological principles across aberrant condensates.
411 To address the incomplete understanding of the dynamics driving condensate
412 formation and the paucity of reliable targets, we repurposed our readout to conduct a
413 genome-wide CRISPR KO screen toward target-directed approaches. Using a deep
414 residual convolutional neural network, further analysis of the resulting condensate
415 phenotypes identified two distinct clusters around either the RNF26 KO or ZNF335 KO
416 condition (Figure 4F). To distinguish these clusters, we performed downstream confocal
417 and electron microscopy on the RNF26 KO and ZNF335 KO conditions. Previous work
418 has implicated RNF26 in ER positioning, endosome retention, and Sec62-mediated
419 recovERphagy.[55,56,80] While condensates in both knockout conditions accumulated K48-
420 linked ubiquitin (Figure 5A), indicating dysregulation of cellular protein quality control,
421 only RNF26-deficient cells showed strong condensate colocalization with nucleoporins

422 (Figure 5B, 5C) and prominent omega-shaped inner nuclear membrane invaginations
423 (Figure 5C, 5D).
424 The identification of RNF26—previously unrelated to Torsins or NE morphology—
425 strongly validates our unbiased genetic approach. Building on our prior discovery that
426 MLF2 is a resident of nuclear envelope blebs,[20] the hallmark phenotype of DYT1
427 dystonia,[1,19,20,28,29] we found here that RNF26 deletion provokes a nearly identical
428 cellular pathology characterized by accumulation of MLF2, nucleoporin-associated
429 nuclear condensates, and omega-shaped inner nuclear membrane invaginations, all of
430 which have previously been observed across dystonia models.[1,19,20,28,29,62] Thus, our
431 study strongly motivates future efforts to explore how RNF26, Torsins, and NE dynamics
432 are functionally connected.
433 Depletion of ZNF335, a C2H2-type zinc finger and component of the trithorax
434 H3K4-methylation complex,[48] resulted in the most pronounced condensate phenotype,
435 with the average number of condensate foci exceeding that found in either the RNF26
436 or TOR1A KO conditions (Figure S4D). Mutations or aberrant splice variants of ZNF335
437 result in severe microcephaly and frequently reduced life expectancy.[48–53] Moreover,
438 ZNF335 has been reported to regulate the REST/NRSF complex with its knockdown
439 inhibiting neuronal differentiation. [48]
440 In addition to ZNF335, pathogenic allelic variants of 15 out of the top 50 gene
441 hits are associated with neurodevelopmental disorders according to human test results
442 from OMIM[45] and the NIH GTR.[46] (Figure 6A). Strikingly, 7/15 of these genes (ZNF335
443 [MCPH10], CIT [MCPH17], NCAPD3 [MCPH22], CDK5, SIN3A, SHH, and SPDL1) are
444 linked to microcephaly, holoprosencephaly, or lissencephaly (OMIM, GTR), along with
445 additional findings implicating NCAPG2, NCAPH2, and EP300 with microcephaly.[48,64,81–
446 84] Given that whole-exome sequencing fails to identify causal disease-causing
447 mutations in more than 70% of microcephaly patients,[85] our curated list of condensate-
448 associated genes provides a valuable resource to prioritize candidates in future patient
449 genetic studies.
450 Several of the microcephaly-associated gene hits identified in our screen are also
451 involved in chromosomal condensation, including several independently identified
452 subunits of the condensin II complex. Microcephaly has previously been labeled as a

453 "condensinopathy",[64] not to be confused with condensatopathies, reflecting prior work
454 that established a relationship between chromosomal condensation defects and
455 reduced cerebral cortex size.[64] Our deep learning model uncovered an unusually similar
456 nuclear MLF2-containing condensate phenotype between cells deficient in ZNF335 and
457 four out of five of the condensin II subunits (NCAPG2, NCAPH2, NCAPD3, and SMC2;
458 Figures 3B, 4G), inspiring future functional studies at the intersection of MLF2,
459 condensin biology, and phase separation.
460 In conclusion, our study lays the foundation for exploring what cellular role
461 biomolecular condensates have in microcephaly or related neurodevelopmental
462 disorders (Figure 6B). We overcome the scarcity of generic phase-separation markers
463 by establishing MLF2 as a versatile condensate biomarker and report several chemical
464 modulators of aberrant condensates. This reinforces similar chemical-modulator
465 screening as a general strategy to identify therapeutic interventions across
466 condensatopathies. Our screen and analysis pipeline are readily adaptable for
467 screening campaigns targeting diverse condensate phenotypes, including stress
468 granules and other disease-relevant condensates. This study uncovers a connection
469 between altered condensate accumulation and neurodevelopmental disorders, inspiring
470 pharmacological strategies to correct aberrant condensates.
471
472 **Limitations of the study**
473 While our study highlights several chemical and genetic modulators of disease-linked
474 condensates, opportunities exist for further refinement. Pyrithione zinc markedly
475 resolved the proteotoxic condensate cargo in torsin-deficient cells, while complete
476 resolution of the underlying condensates was not observed at the tested concentration
477 (Figure S2I). Extended incubation times may result in condensate resolution, but
478 increased cytotoxicity limited this approach. Similarly, while our genetic screen
479 uncovered an enrichment of neurodevelopmentally-linked genes, redundant pathways
480 or those strictly requiring a neuronal context are likely to be underrepresented and we
481 filtered out conditions that resulted in severe cytotoxicity. Large-scale screening of
482 optimized compound libraries and testing genetic redundancy through combinatorial
483 perturbations will likely identify additional condensate modulators. Furthermore,

assessing the condensate burden associated with these neurodevelopmental conditions would benefit from future studies in neuronal or in vivo systems.

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

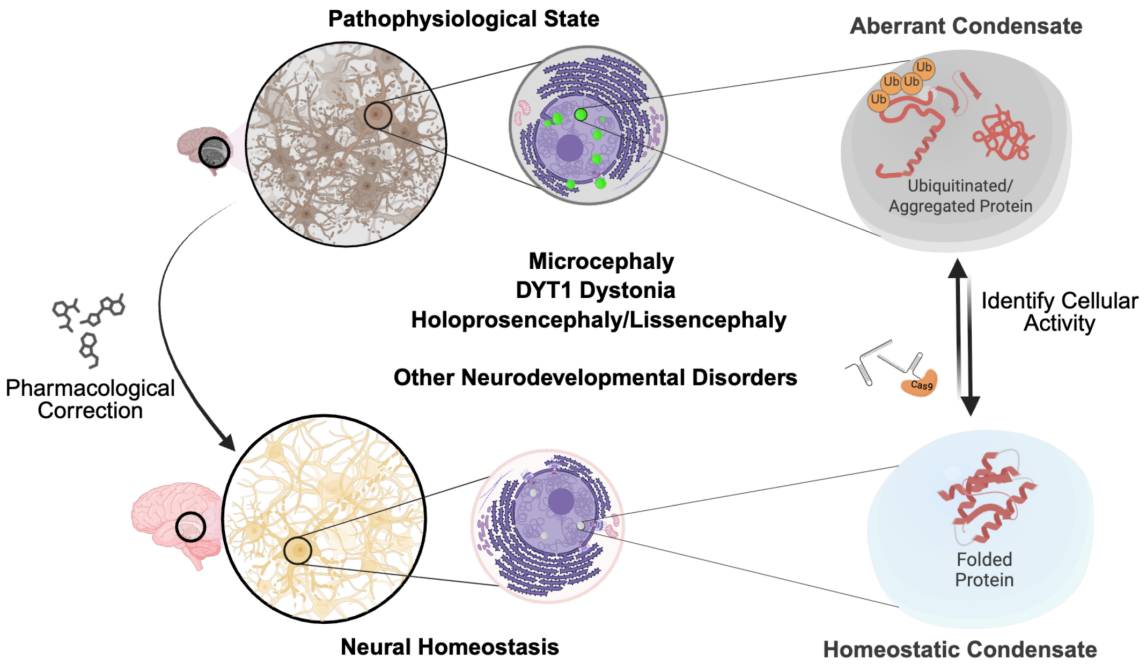

Figure 6, *Poch et al.*

489

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 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.

Supplementary Figures

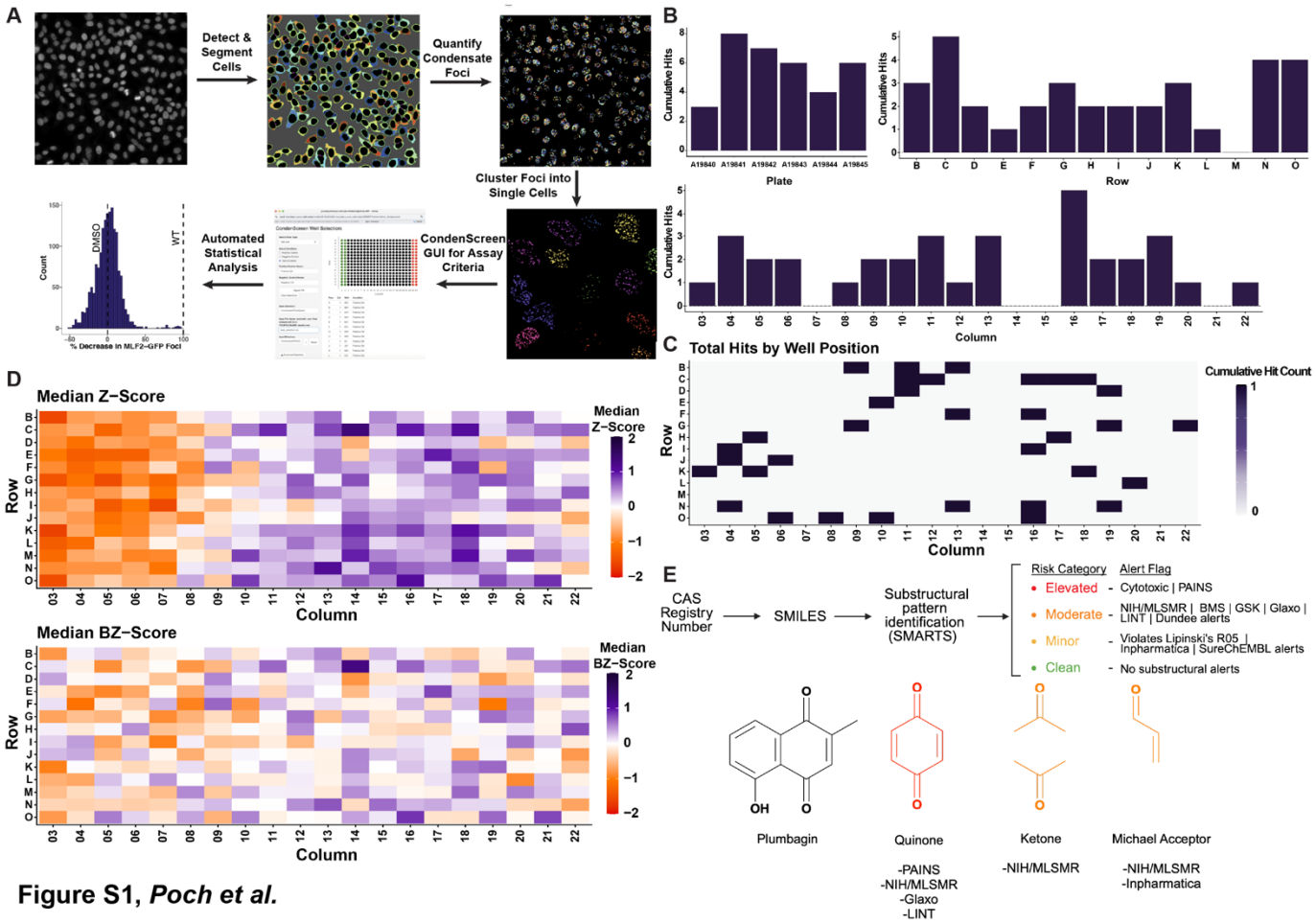

Figure S1, *Poch et al.*

Fig. S1. CondenScreen as a high-content condensate analysis pipeline (**A**) Schematic of the computational *CondenScreen*. Initial cell segmentation and foci detection uses custom *CellProfiler* script and while downstream statistical analysis is completed in *R*. (**B**) Distribution of cumulative hit counts from the Pharmakon 1760 library screen (n = 6 plates) summarized by plate (top left), row (top right), and column (bottom). (**C**) Spatial heatmap of total hits per well position across all plates. Color intensity reflects hit frequency, with darker purple indicating higher cumulative hit counts (a maximum of one hit per well was observed). (**D**) Comparison of hit detection using the Z-Score vs BZ-Score Normalization methods. A heatmap of the median (n = 6) of the Z-Score (top) vs the BZ-Score (bottom) per well. Color indicates score magnitude and direction (orange = negative, purple = positive, white ≈ 0; darker shading reflects larger absolute values). (**E**) Schematic of CondenScreen's built-in chemical structure hit triage workflow. (top) Using either the name or CAS numbers, the SMILES will be returned (PubChem/NCI CACTUS), followed by substructural pattern identification (SMARTS) flags using several industry alert flags, along with the tested cytotoxicity of each compound. (bottom) Example results using plumbagin. Three alerts were flagged for plumbagin, including a quinone, ketone, and Michael acceptor substructural elements. As quinones are listed as a PAINS substructure, plumbagin received a risk category ranking of "Elevated".

493

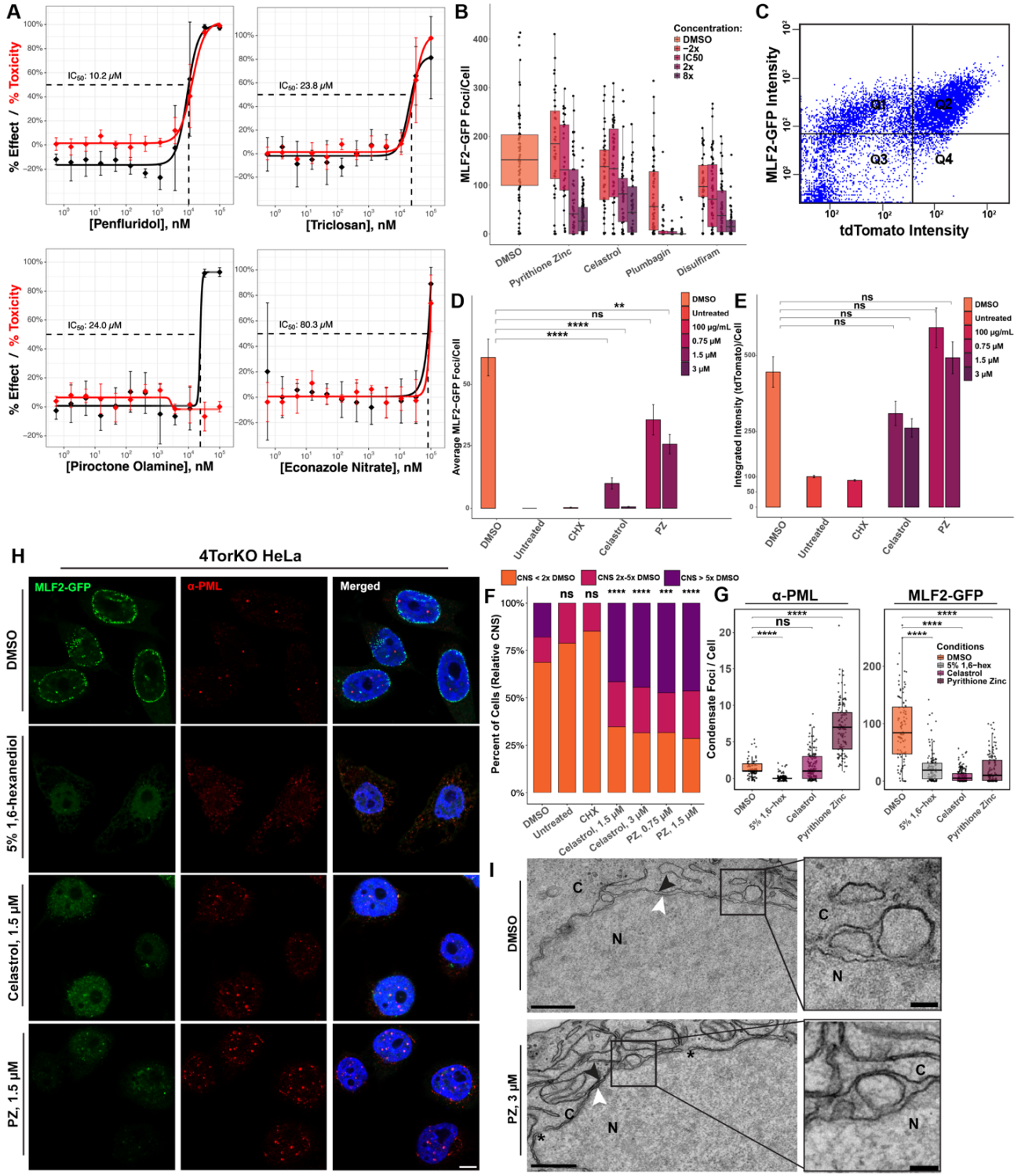

Figure S2, *Poch et al.*

494

Fig. S2. Validation of chemical hits from the Pharmakon 1760 screen. (**A**) Calculated IC_{50} values of additional compounds with dose-dependent relationship. The black and red curves represent % effect and % toxicity, respectively, relative to the DMSO control. (**B**) 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, compared against the DMSO control. Each dot represents the number of MLF2-GFP foci for a single cell, while the horizontal lines indicate the median foci/condition. (**C**) Distribution of FACS results for 4TorKO MLF2-GFP-P2A-tdTomato HeLa cell sorting. Cells within quadrant 2 were sorted. (**D**) Quantification of the number of MLF2-GFP foci/cell (**E**) and tdTomato integrated intensity/cell. Cells (4TorKO) were all (except the untreated condition) induced with Dox and treated with either DMSO, cycloheximide (CHX), celastrol, or PZ. (**F**). Binned CNS value per cell across treatment conditions. Orange indicates <2x the median DMSO CNS value, red indicates 2–5x and purple indicates >5x the median DMSO control. Quantification (**G**) and confocal imaging (**H**) show that PZ does not non-specifically dissolve physiological condensates. Celastrol and PZ reduced MLF2-GFP foci (green) without dissolving PML bodies (red), whereas 5% 1,6-hexanediol dissolved both. All three compounds decreased MLF2-GFP foci relative to DMSO; only 1,6-hexanediol reduced PML foci. PZ increased PML foci number, while celastrol had no significant effect. Scale bars, 5 µm. (**I**) TEM images of 4TorKO HeLa cells treated with either 0.1% DMSO or 2 µM PZ for 3 h. Insets show NE 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. Asterisks indicate Bonferroni-corrected p-values determined by Mann-Whitney U testing (**: $p <= 0.01$; ****: $p <= 0.0001$).

497

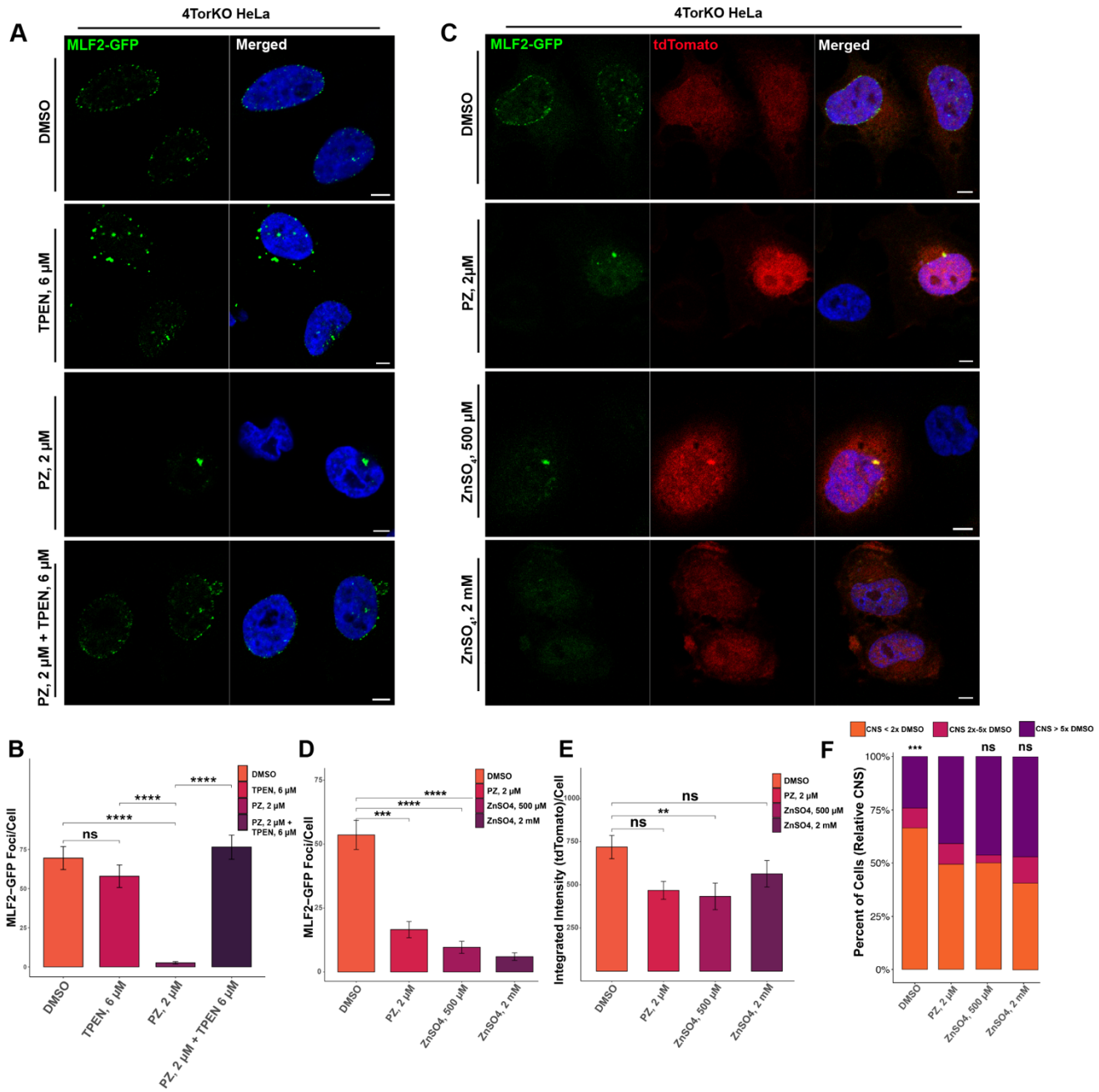

Figure S3, *Poch et al.*

Fig. S3. 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**) Binned CNS value per cell across treatment conditions. Orange indicates <2x the median DMSO CNS value, red indicates 2–5x and purple indicates >5x the median DMSO control. 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).

498

499

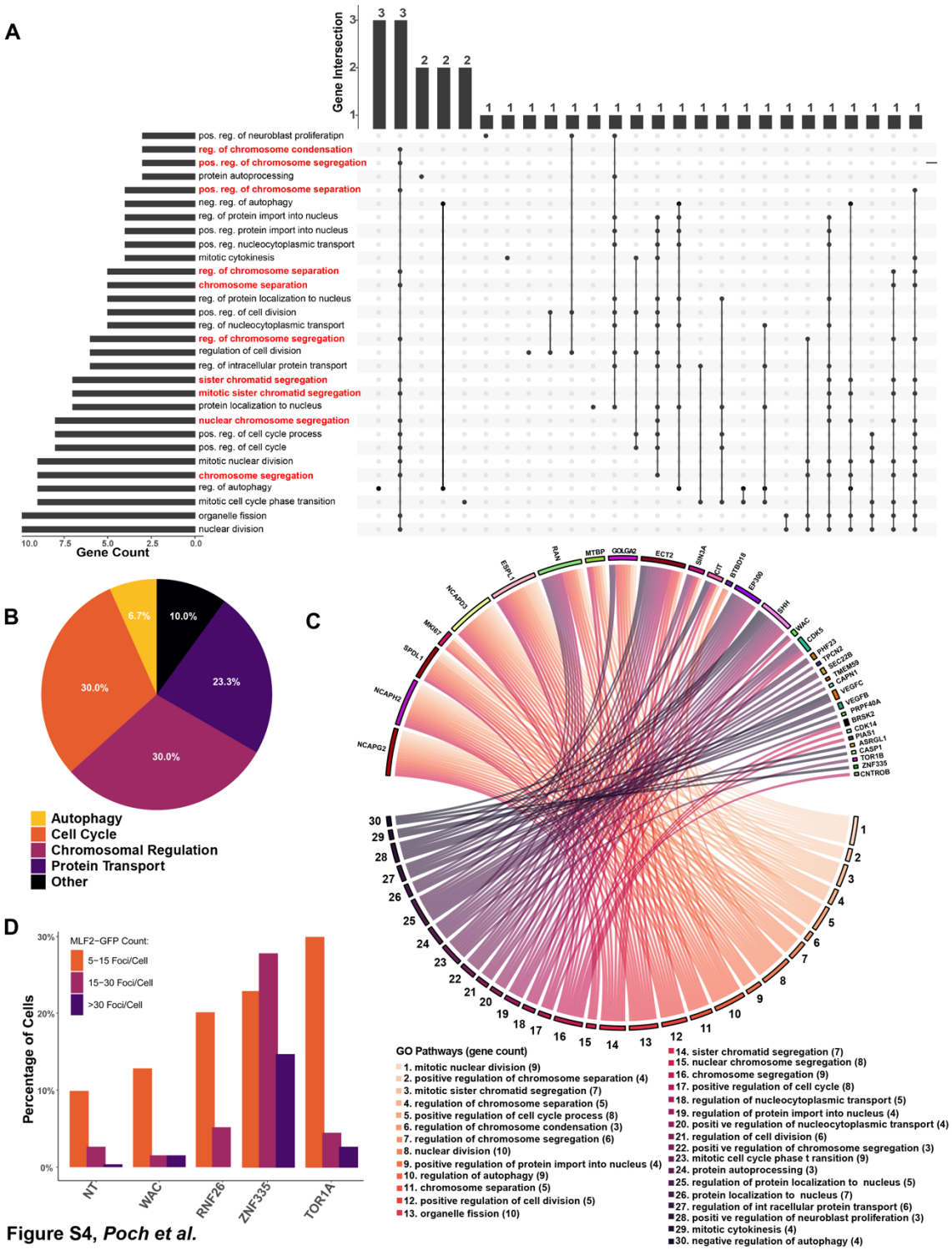

Figure S4, *Poch et al.*

500

501

Fig. S4. Validation and enriched gene pathways of CRISPR KO hits that promote condensate accumulation. (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).

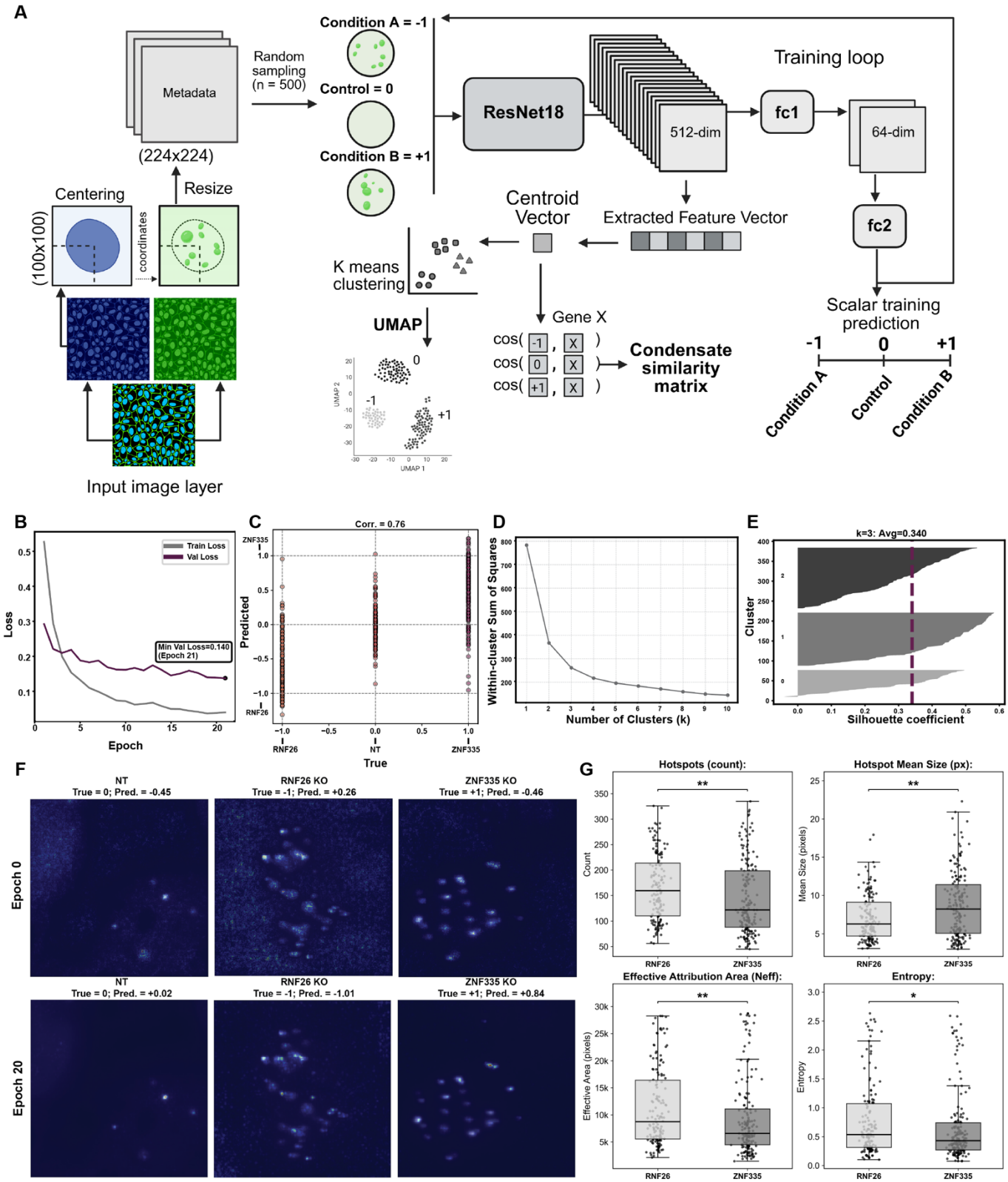

Figure S5, *Poch et al.*

Fig. S5. 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 (gray) and validation set (purple) are reported with a minimum loss (0.14) occurring at 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 0.76. (**D**) Elbow (scree) plot showing the within-cluster sum of squares across cluster numbers (k = 1-10). (**E**) Silhouette plot showing the silhouette coefficient (0.34, vertical dashed line) for three clusters. (**F**) Representative saliency maps prior to training (top, Epoch 0) vs post-training (Epoch 20). From left to right, non-targeting control cell (true label = 0, initial pred. = -0.45, final pred. = +0.02), RNF26 KO cell (true label = -1, initial pred. = +0.26, final pred. = -1.01), and ZNF335 KO cell (true label = +1, initial pred. = -0.46, final pred. = +0.84). (**G**) Boxplots of the hotspot count, mean size of hotspots, effective attribution area, and Shannon's Entropy across the single cells (dots). Asterisks indicate p-values using a Mann-Whitney U test (*: p <= 0.05; **: p <= 0.01).

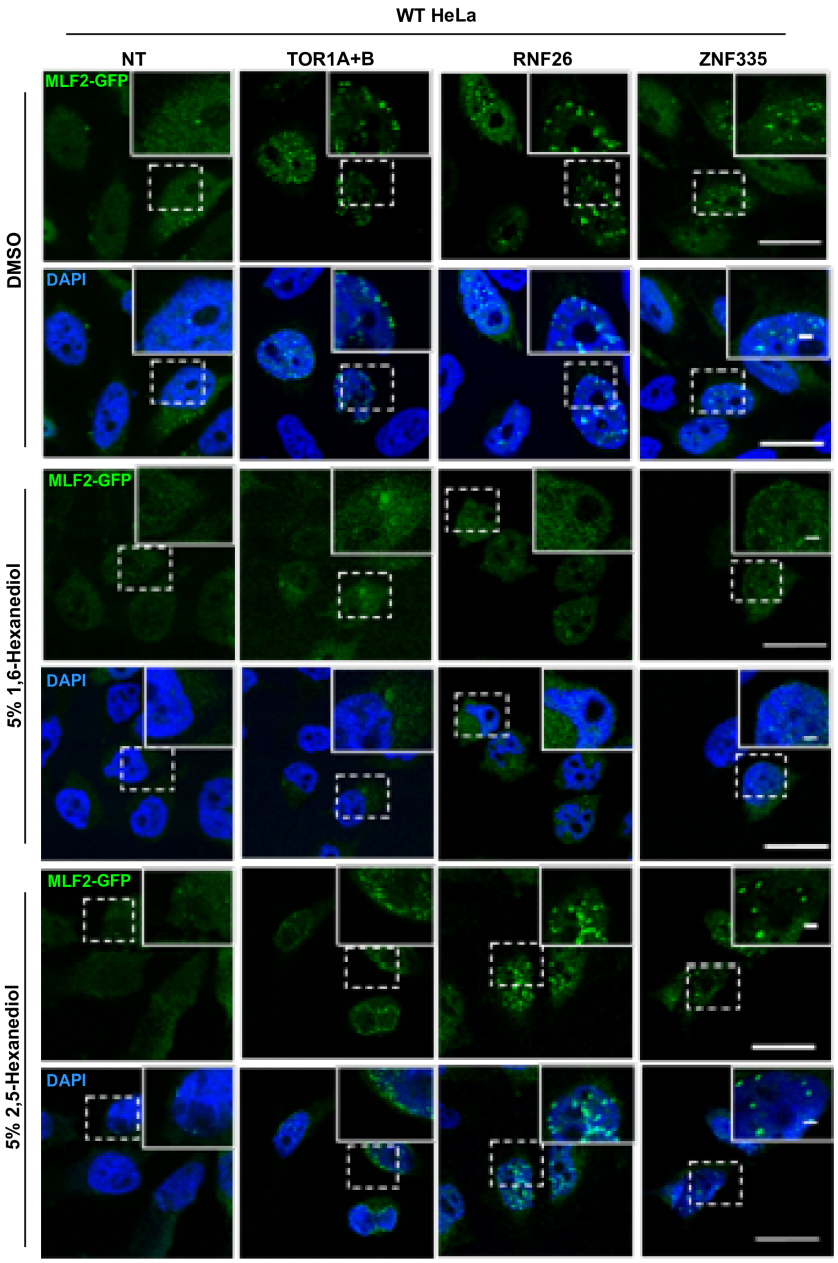

Figure S6, *Poch et al.*

Fig. S6. 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.

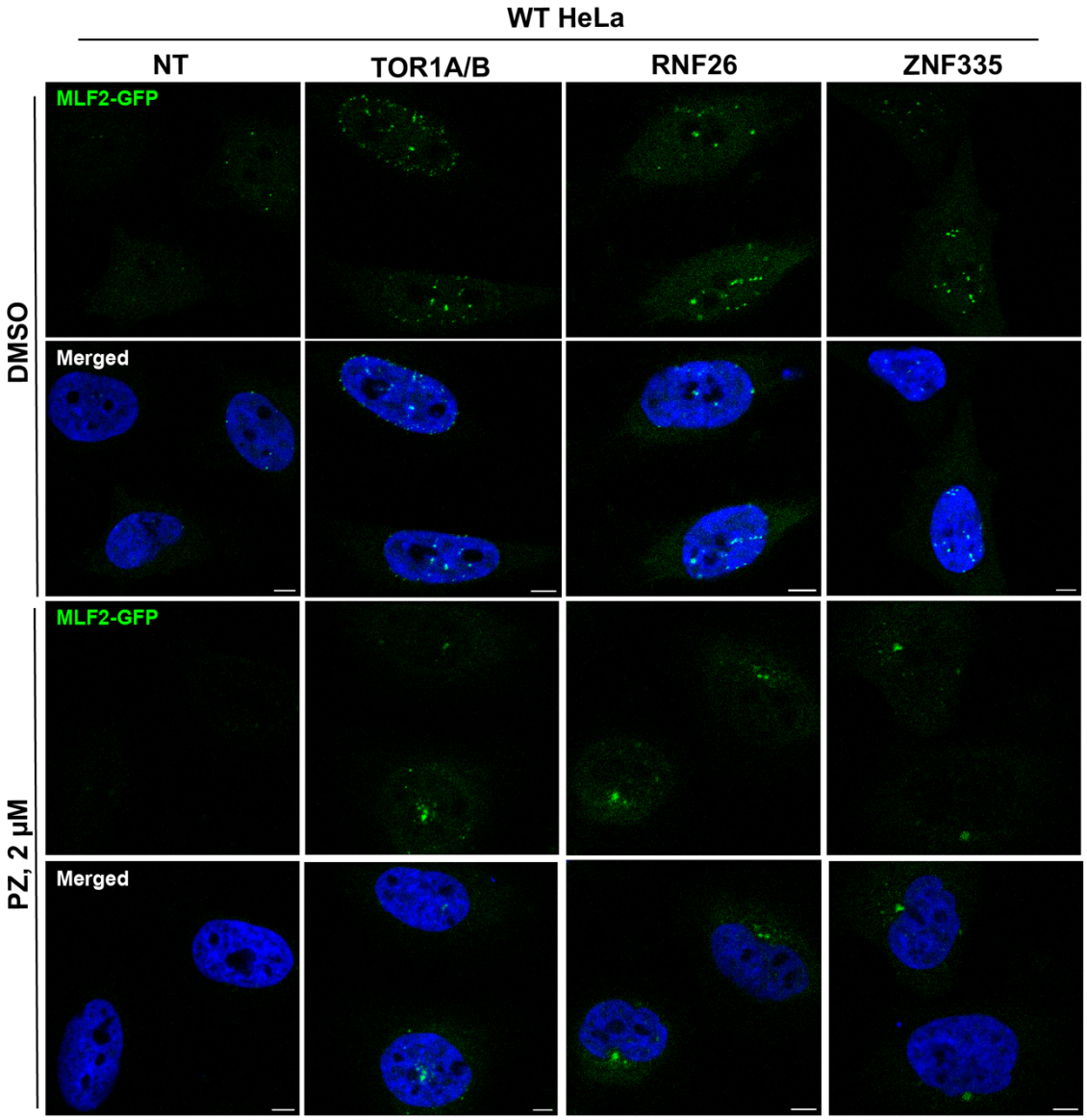

Figure 7, *Poch et al.*

Fig. S7. Pyrithione zinc modulates atypical 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.

523
524

525

526 **Materials and Methods**

527 **Tissue culture**

528 Adherent cultured cell lines (HeLa and HEK293T) were maintained in Dulbecco's
529 modified Eagle's medium (DMEM) supplemented with 10% (vol/vol) heat-inactivated
530 fetal bovine serum (Thermo Fisher Scientific) and 100 units (U) /mL penicillin-
531 streptomycin (Thermo Fisher Scientific). Cells were maintained in either T75 flasks or 10
532 cm dishes at 37℃ (95% humidity; 5% CO_2) and passaged every third day. For splitting,
533 the culture medium was removed, washed once with 1x PBS (Gibco), trypsinized with
534 0.25% trypsin-EDTA (Gibco), and transferred to new dishes (diluted 1:5 or 1:10).

535

536 **Plasmid Constructs**

537 Unless otherwise stated, all plasmid constructs were subcloned into a pcDNA™3.1 (+)
538 Mammalian Expression Vector (Invitrogen, V79020) using standard Gibson Assembly
539 protocols. The cDNA for the MLF2-GFP-P2A-tdTomato construct (MLF2 GenBank
540 accession no. BC000898) was subcloned into the pRetroX-Tight-Pur-GOI (gene of
541 interest) vector (Takara Bio) and co-transfected with a pRetroX-Tet-On advanced
542 inducible expression construct (Takara Bio). Viral constructs were co-transfected with
543 both a VSV-G viral envelope plasmid (Addgene #12259) and an MMLV gag/pol viral
544 packaging plasmid. For generation of HeLa cells expressing Cas9-mCherry, an mCherry
545 tag was inserted into the lentiviral plasmid containing Cas9-Flag (Addgene #52962)
546 immediately after the Flag tag into the BamHI cloning site.

547

548 **Viral Production**

549 Lentivirus production was carried out as previously described.[86,87] Briefly, to generate
550 Cas9-flag lentivirus, low-passage HEK293T cells were plated in a 10 cm dish with a
551 density of 3.5×10^6 cells per dish. Medium was replaced the next day with 6 mL fresh
552 (antibiotic-free) DMEM and transfected with three plasmids encoding: psPAX2
553 (Addgene #12260) (2 µg), viral envelope protein VSV-G (1 µg), and lenti-Cas9-Flag-
554 mCherry (6 µg) using Lipofectamine 2000 (Thermo Fisher Scientific) at a 3:1 ratio of
555 DNA to Lipofectamine. After 72 h, supernatants containing the virus encoding Cas9-

556 mCherry were harvested and filtered through a 0.45 µm Steriflip filter unit
557 (MilliporeSigma), aliquoted into screw-cap tubes, and stored at -80°C or used for
558 transduction immediately. For the Dox-inducible MLF2-GFP-P2A-tdTomato system, two
559 separate retroviruses were produced using three plasmids encoding: MMLV gag/pol (2
560 µg), viral envelope protein VSV-G (1 µg), and either pRetroX-Tight-MLF2-GFP-P2A-
561 tdTomato or pRetroX-Tet-On (6 µg) using Lipofectamine 3000 (Thermo Fisher Scientific)
562 at a 2:1.5 ratio supplemented with Lipofectamine P3000 (2.5:1). After 72 h,
563 supernatants were collected, stored, or used, as described.
564
565 **Generation of stable HeLa cell lines**
566 The detailed protocol regarding the Tor1A/1B/2A/3A KO and integration of a Dox-
567 inducible MLF2-GFP reporter in HeLa cells has been extensively described by our lab
568 (Laudermilch et al. 2016; Rampello et al. 2020).[20,28] The Cas9-flag-mCherry HeLa and
569 Dox-inducible MLF2-GFP-P2A-tdTomato 4TorKO HeLa cell lines were generated by
570 seeding Dox-inducible MLF2-GFP WT HeLa or 4TorKO cells in six-well plates. The next
571 day, media was replaced with DMEM (antibiotic-free) supplemented with 4 µg/mL
572 polybrene (Sigma-Aldrich), and 100 µl of either the Cas9-Flag-mCherry virus (WT HeLa)
573 or both 100 µL of the MLF2-GFP-P2A-tdTomato and pRetro-Tet-On viruses (4TorKO)
574 added dropwise to wells. Medium was replaced 48 hours after transduction (complete
575 DMEM) and grown for another 24 hours. Cells were collected in sorting buffer (ice-cold
576 1x PBS, Gibco) and filtered through a 40 µm nylon mesh filter. Using a BD FACS Aria
577 III, cells were sorted as single cells by selecting those high in either mCherry (WT HeLa)
578 or MLF2-GFP/tdTomato (4TorKO, 16 h after 0.5 µg/mL Dox induction) fluorescence
579 intensity. A stringent gating strategy was applied to enrich for cells positive for the
580 respective fluorophores while excluding those with extremely high signal intensity, to
581 obtain a pure yet physiologically representative population.
582
583 **High-content phenotypic screening strategy**
584 All high-content screening was performed at Yale's Center for Molecular Discovery
585 (YCMD). The Horizon Discovery Edit-R synthetic sgRNA library was used in the whole
586 genome CRISPR KO screen, targeting 19,195 human genes, and comprising of the

587 Drug Targets library (3,683 genes, GA-004650-E2, Lot 20126), the Druggable Genome
588 library (4,643 genes, GA-004670-E2, Lot 20120), and the Human Genome library
589 (10,869 genes, GA-006500-E2, Lot 20130). Each screening plate included 16 replicates
590 of the negative control (NT sgRNA) and 16 replicates of the positive control (combined
591 Tor1A and Tor1B sgRNAs). To ensure robust KO, a pool of three sgRNAs targeting the
592 same gene at different regions was pre-plated into each arrayed well of 384-well glass-
593 bottom plates (Cellvis, #384-1.5H-N). Addition of sgRNA to screening plates was done
594 using the Labcyte Echo 550 Acoustic Dispenser (Labcyte, now Beckman) at a final
595 concentration of 25 nM.

596 Reverse transfection was performed using DharmaFECT1 Transfection Reagent
597 (Horizon/Dharmacon), which was diluted at 1:500 in Opti-MEM (Thermo Fisher
598 Scientific) and mixed with sgRNA. The mixture of sgRNA and transfection reagent was
599 incubated for 30 minutes, after which, 2,000 cells/well (WT HeLa integrated with
600 Cas9/MLF2-GFP) in antibiotic-free DMEM were added using a Thermo MultiDrop Combi
601 (Thermo Fisher Scientific). The plates were then incubated for 72 hours at 37°C. To
602 induce MLF2-GFP reporter expression, 0.5 µg/mL Dox was added to each well 16 h
603 before fixation. Media was removed and cells washed once with PBS and fixed with 4%
604 paraformaldehyde (PFA) for 20 minutes. Cells were washed twice with PBS,
605 permeabilized with 0.1% Triton X-100 for 5 minutes, and washed again with PBS. Cells
606 were stained with Hoechst 33342 dye (1:1000) (Invitrogen, H1399) and HCS CellMask
607 Deep Red Stain (Invitrogen, H32721) (1:5000 of 10 mg/mL, Thermo Fisher Scientific)
608 diluted in PBS for 30 minutes, and washed a final time in PBS. The Molecular Devices
609 ImageXpress 4.0 high-throughput fluorescent microscope (40x magnification) was used
610 to obtain 9-image sets (each consisting of DAPI, MLF2-GFP, and Deep Red CellMask
611 images) per well, resulting in 10,368 TIFF images/plate. Images were uploaded to
612 Yale's McCleary high-performance computing cluster (Yale Center for Research
613 Computing) for analysis.

614 For chemical screening, the Pharmakon 1760 drug library (MicroSource
615 Discovery Systems, Inc.) was used, consisting of 1360 FDA-approved drugs and 400
616 International drugs on six 384-well plates. Each screening plate consisted of 16
617 replicates of the negative control (4TorKO HeLa cells treated with DMSO [0.1% v/v])

618 and 16 replicates of the positive control (WT HeLa cells treated with DMSO [0.1% v/v]).
619 Plates were seeded with 4,000 HeLa cells/well (Multidrop Combi) in 20 µL of DMEM
620 before incubating at 37°C for 24 h. MLF2-GFP expression was induced by adding 0.5
621 µg/mL of Dox into each well 8 h prior to fixation. Compounds were resuspended in
622 DMSO and added to each well 2 h prior to fixation using a Labcyte Echo 550 Acoustic
623 Dispenser (Labcyte, now Beckman), at a final concentration of 10 µM. Cells were
624 washed, fixed, stained, and imaged using the methods outlined above.
625
626 **Image segmentation of phenotypic screen data**
627 We created a custom analysis pipeline to segment condensate foci into single cells
628 using the open-source software *CellProfiler 4.2.6.*[38,39] and downstream statistical
629 analysis using *R Studio* (Version 2023.12.1+402).[88] Individual nuclei were distinguished
630 by shape for objects between 30-100 (+4) pixels using an adaptive (40 pixels) Minimum
631 Cross-Entropy thresholding (0.01-1.0 threshold) in *CellProfiler 4*. The cytoplasm of each
632 nucleus was delineated as a secondary object of each nucleus using a watershed
633 algorithm with global Otsu thresholding (three classes, foreground; 0-1.0 threshold,
634 threshold correction factor = 0.7).[39,89] Each GFP-channel image was enhanced using a
635 white top-hat transform (speckles; maximum 10 pixels), allowing each condensate foci
636 between 1-10 pixels to be detected using adaptive (50 pixels) Otsu thresholding (three
637 classes, foreground; 0.0175-1.0 threshold, threshold correction factor = 1.2). Groups of
638 condensate foci were then distinguished using local intensity and related to their
639 corresponding cytoplasm/nuclei. The same image analysis pipeline was applied across
640 all imaging experiments, with only minor adjustments (e.g., cell size or foci intensity
641 thresholds). Any parameter changes were applied uniformly across all conditions and
642 controls within a given experiment.
643 For parallel analysis, a Bash script was developed allowing the execution of
644 multiple headless CellProfiler instances on a Simple Linux Utility for Resource
645 Management (SLURM)-managed compute cluster. Specifically, after constructing the
646 analysis pipeline in the CellProfiler GUI, the CreateBatchFiles module is used to
647 generate a batch configuration file that defines image metadata and processing
648 instructions. This batch file is then imported to the cluster environment, and all images

649 are registered within a single CellProfiler project in array format, with images from each
650 plate corresponding to a separate SLURM job. This approach enables autonomous
651 analysis of over 1 million high-content images. TIFF images—amounting to multiple
652 terabytes of data—within a matter of hours. Both the CellProfiler pipeline and Bash
653 scripts are open source and available publicly on the SchliekerLab GitHub page as part
654 of the CondenScreen pipeline (https://github.com/SchliekerLab/Condenscreen.git).
655

656 **Statistical analysis of screen data**

657 We developed a condensate screening using a combined *CellProfiler* and *R* pipeline,
658 *CondenScreen*, to statistically analyze all screen data. The *R*-script features an
659 interactive GUI allowing the user to select plate size (6-well to 384-well format), plate
660 layout, and input whether the data was taken from a signal "ON" or a signal "OFF"
661 screen. With limited user input, it then automatically iterates through images from each
662 plate and assigns groups of image sets into their corresponding wells using string
663 parsing. To control for variations in transfection efficiencies that could inadvertently
664 impact downstream scoring, the program has an optional filter to select only the top
665 25% of cells showing the highest (or lowest) signal using the modest assumption that at
666 least 1 out of 4 cells, on average, will have received the sgRNAs of interest. Additionally,
667 for both screens, the image with the highest score was filtered out, as there were
668 several conditions where one image-set had screening aberrations skewing the
669 condition toward a false positive result. For each plate, aggregate statistics of the
670 respective positive and negative controls are compared, including signal-to-background
671 (SB), standard deviation (SD) of the controls, and the coefficient of variation (CV) of the
672 controls, among others. To determine separation between positive and negative
673 controls, a Z' assay quality score was calculated: $Z' = 1 - (3(\sigma_{PC} + \sigma_{NC})/|\mu_{PC} -$
674 $\mu_{NC}|)$,[40] where σ_{PC} and σ_{NC} is the SD and μ_{PC} and μ_{NC} is the condensate mean of the
675 positive and negative controls, respectively.

676 The percent effect of each condition was quantified by either $E\% = \left(\frac{\mu - \overline{x}}{\mu}\right) \cdot 100$
677 (for signal "OFF" selection) or $E\% = \left(\frac{\overline{x} - \mu}{\mu}\right) \cdot 100$ (for signal "ON" selection) where $\overline{x}$ is
678 the average number of condensate foci/cell per tested condition, and μ is the average

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.[90,91]

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, which was used during validation experiments. Specifically, we added several modules to *CellProfiler* to measure the size, shape, intensity, and texture of each nucleus.[38] 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.

709 Two optional *CondenScreen* quality control features allow for flagging potential
710 false-positives or problematic hits arising from screening large libraries. In addition to
711 detecting spatial intra-plate bias, the program flags screen actives arising from global
712 systematic inter-plate spatial biases. These include shifts in continuous significance
713 score metrics (Z-Scores and BZ-Scores) and enrichment of binary hits across rows,
714 columns, individual well positions, and plates with a statistically enriched hit rate.
715 Continuous metric shifts are assessed using a Wilcoxon rank-sum (Mann–Whitney U)
716 test comparing each group to all other wells, while hit enrichment is assessed using
717 Fisher's exact test. P-values are Benjamini–Hochberg adjusted across groups and
718 flagged at an adjusted p-value < 0.05.
719 *CondenScreen* enables flagging potentially problematic chemical structures
720 using compound SMILES (Simplified Molecular Input Line Entry System), either
721 provided manually (recommended) or can attempt to be retrieved automatically from
722 PubChem[92] using the CAS number, with CACTUS[93] (NCI Chemical Identifier Resolver,
723 https://cactus.nci.nih.gov/chemical/structure) used as a fallback. Smiles are processed
724 using RDKit (via Reticulate in R), and compounds are evaluated against SMARTS
725 (SMILES Arbitrary Target Specification) substructures appearing in alert libraries
726 compiled from RDKit and ChEMBL resources. We utilized a pre-compiled list of these
727 structural alert filters from a csv document allowing SMARTS identification within the
728 RDKit environment, compiled by Pat Walters (see GitHub for more details:
729 https://github.com/PatWalters/rd_filters.git). We defined compounds as having an
730 "Elevated" risk if they include SMARTS overlapping with the Pan-Assay Interference
731 Compounds (PAINS) alerts,[94,95] NIH Molecular Libraries Small Molecule Repository
732 (MLSMR) alerts, or have a significant cytotoxicity (≥25% viability loss); "Moderate" risk if
733 overlapping with the Bristol-Myers Squibb (BMS), GlaxoSmithKline (GSK/Glaxo), Pfizer
734 LINT alerts, or the University of Dundee filter sets, or tdTomato reporter suppression
735 (≥25% decrease); and "Minor" risk if violating the Lipinski rule-of-five[96] or appearing in
736 the Inpharmatica or SureChEMBL medicinal chemistry alert libraries. Compounds
737 lacking these features were classified as "Clean." These categories are intended to
738 support chemical triage rather than exclude compounds.

739 *CondenScreen* provides additional downstream inter-plate and intra-plate
740 visualizations using open-source libraries, including *ggplot2*[97] and *ggplate*
741 (10.32614/CRAN.package.ggplate). The output includes a ranked list of hits, plots of the
742 screening distributions (raw condensate foci/cell and %Effect), a Z/BZ-score plot, and
743 individual plate heatmaps for each 384-well plate. Finally, top hits in both screens were
744 inspected manually to ensure quality control and rule out any additional potential
745 screening aberrations. Notably, manual inspection was rarely a reliable measure of hit
746 robustness. In several cases, hits that were initially overlooked during first-pass
747 screening (false negatives) were later validated and correctly identified by the program.
748 The *CondenScreen* platform is publicly available on the SchliekerLab GitHub page:
749 https://github.com/SchliekerLab/Condenscreen.git
750

751 **Development of phenotypic machine learning model to distinguish condensate**
752 **phenotypes**

753 We developed a supervised machine learning model to differentiate between subtle
754 phenotypic differences in nuclear condensates. The model was built around the
755 ResNet18 architecture[67] with pretrained initial starting weights borrowed from
756 ImageNet1k IMAGENET1K_V1[68] using torchvision[98] on Python 3.11.8 (conda-forge).
757 Using images directly from the CRISPR KO screen, 4,523 single cells were extracted,
758 cropped (100x100) and resized (224x224) before centering (DAPI-channel) each cell.
759 Single cells were labeled in accordance with their file names, with NT Ctrl cells = 0,
760 RNF26 KO cells = -1, and ZNF335 KO cells = +1. To ensure even sampling numbers
761 across conditions, cells from the overrepresented NT Ctrl conditioned were
762 downsampled to match KO conditions, with only the MLF2-GFP channel selected for
763 analysis. Data were split 80:20 into train and validation sets (sklearn.model_selection:
764 train_test_split)[99]. Tensors were packaged (torch.utils.data) before employing the
765 following parameters during model training (PyTorch[100]): AdamW (lr = 1e-4, weight
766 decay = 5e-2), dropout = 0.5, hidden dimension = 64, SmoothL1Loss,
767 CosineAnnealingLR, 100 max epochs, and early stopping (patience = 25). Model
768 performance was evaluated using several modules from sklearn.metrics and
769 scipy.stats,[99,101] including: mean_absolute_error(); r2_score(); pearsonr();

770 accuracy_score(); confusion_matrix(); roc_curve(); roc_auc_score(); and
771 ConfusionMatrixDisplay(). The integrated gradient[69] for the validation data was used for
772 post-hoc saliency mapping in order to ensure the model was learning biological features
773 compared to technical noise. Briefly, for each cell, an attribution map was calculated
774 using integrated gradients, which averages model gradients across 24 interpolation
775 steps between the input image and a zero baseline. This attribution map was min-
776 maxed normalized and converted to a single-cell 2D heatmap. For each cell, the spatial
777 distribution of saliency attribution (Shannon entropy;[102] heat_entropy), effective number
778 of pixels contributing to the saliency distribution (heat_effective_npix), and hotspot
779 number and mean size (heat_hotspot_components), where hotspots were defined as
780 connected components within the top 2% of saliency intensities. The per-cell
781 distributions for ZNF335 vs RNF26 cells were statistically assessed using a two-sided
782 Wilcoxon rank-sum (Mann–Whitney U) test.
783 After optimization, all CRISPR KO gene hits with a Z-Score > 2 were cropped,
784 resized, and centered before randomly downsampled to 500 cells per condition
785 (accounting for differences in cell viability). Cells were then analyzed by the optimized
786 model and their latent representation embeddings (512-dimensional latent embeddings
787 per cell, e.g., condensate morphology, size, intensity, etc.) were extracted. Relative
788 centroid vector scores were computed (mean embedding across cells), and KMeans
789 clustering with k = 3 was applied to UMAP (n_neighbors = 15, min_dist = 0.1)
790 coordinates to assign genes to discrete clusters. The optimal cluster number was
791 selected using the elbow method (KMeans, k = 1–10) and silhouette analysis
792 (sklearn.metrics: silhouette_samples | silhouette_score), yielding three distinct clusters.
793 The difference in the cosine similarity for each gene's centroid vector embedding was
794 compared with the three anchor centroids (RNF26/NT Ctrl/ZNF335) and exported to an
795 .xlsx file (see Table S3). A hierarchical heatmap was constructed using the top 50 genes
796 with highest similarity to the RNF26 and ZNF335 KO conditions, via pandas,
797 seaborn.clustermap, and matplotlib.pyplot. For a full list of modules imported and model
798 training parameters, please refer to our source script:
799 https://github.com/SchliekerLab/Condenscreen.git
800 **Solubilization of hit compounds**

801 All putative hit compounds were resuspended up to 100 mM in DMSO. The following
802 compounds were solubilized: pyrithione zinc (PZ) (CAS 13463-41-7), celastrol (CAS
803 34157-83-0), nerol (CAS 106-25-2), spiramycin (CAS 8025-81-8), choline chloride (CAS
804 67-48-1), disulfiram (CAS 97-77-8), bronopol (CAS 52-51-7), penfluridol (CAS 26864-
805 56-2), plumbagin (CAS 481-42-5), dyphylline (CAS 479-18-5), econazole nitrate (CAS
806 24169-02-6), ethacrynic acid (CAS 58-54-8), piroctone olamine (CAS 68890-66-4),
807 dihydrocelastryl (CAS 80-97-7), fenipentol (CAS 583-03-9), triclosan (CAS 3380-34-5),
808 eprodisate disodium (CAS 36589-58-9), L-phenylalanine (CAS 63-91-2), nabumetone
809 (CAS 42924-53-8), edrophonium chloride (CAS 116-38-1), josamycin (CAS 16846-24-
810 5), mevastatin (CAS 73573-88-3), clofibric acid (CAS 882-09-7), docosahexaenoic acid
811 (CAS 6217-54-5).
812
813 **Dose-dependent validation of small molecules**
814 Based on the change in MLF2-GFP foci, viability, and structural diversity, the 22 top-
815 scoring putative hit compounds were resuspended to 100 mM in DMSO. Using
816 analogous methods described in the "High-content screening strategy" section, a 12-
817 point dose response was constructed with concentrations ranging from < 100 nM to 100
818 μM of each compound, with a final working DMSO concentration of 0.1%. All
819 compounds were tested using four replicates, and 12-positive (WT HeLa) and 12-
820 negative (DMSO) controls were included on each screening plate to ensure assay
821 robustness. MLF2-GFP was induced with 0.5 μg/mL of Dox 8 h before fixation, and
822 compounds were added 3 h prior to fixation (5 h after initial Dox addition). Cells were
823 washed, fixed, stained, and imaged as described above. All dose-response images
824 were analyzed using an analogous *CellProfiler* + *R*-script pipeline.
825 The effect of each concentration was interpolated (*R*, *drc* & *ggplot*) using a four-
826 parameter log-logistic function for each compound: $f(x) = \min + (\max - \min)/(1 +$
827 $e^{h(\log(x) - \log(IC_{50}))})$, where x is the concentration, min represents the lower asymptote,
828 max represents the higher asymptote, IC_{50} is the concentration at which response is
829 50% of the way between min and max, and h is the hill coefficient.[97,103] The percent
830 toxicity was calculated by fitting the additive inverse of the nuclear health function
831 values to the same model.

832

833 **Pathway enrichment analysis**

834 Gene Ontology (GO) term enrichment analysis was used for the top 100 CRISPR gene
835 hits.[60] The gene/chemical hits were separately analyzed in *R* to determine enriched
836 biological process pathways using the enrichGO function (*clusterProfiler*).[104] Enriched
837 pathways were compared against an annotation database (org.Hs.eg.db). The top 30
838 pathways were then plotted using the pheatmap function and additionally plotted to an
839 interactive .html graphic using heatmaply.[105] Pathways were grouped by hierarchical
840 clustering using Euclidean distance with complete linkage. To explore the interactive
841 heatmap, please visit the Schlieker Lab GitHub page and download the corresponding
842 .html files under the "Interactive Figures" folder.

843

844 **Immunofluorescence and confocal microscopy:**

845 HeLa cells were grown on 13 mm round coverslips and fixed with 4% (v/v)
846 paraformaldehyde (Thermo Fisher Scientific), washed with 1x PBS (Gibco),
847 permeabilized with 0.1% (v/v) Triton X-100 (Sigma-Aldrich) for 5 minutes, and blocked
848 with a PBS mixture containing 2% (w/v) BSA and 2% (w/v) glycine. Primary antibodies
849 were diluted in blocking buffer and added dropwise to coverslips for a one-hour
850 incubation. After washing 3x with blocking buffer, fluorescent secondary antibodies
851 diluted in blocking buffer were added for an additional hour, followed by a 1x PBS wash.
852 Coverslips were then stained using a 1:2000 dilution of Hoechst 33342 (Life
853 Technologies). Cells were then washed 3x with PBS before mounting onto slides
854 (Fluoromount-G, Southern Biotech). Primary antibodies were diluted 1:500: anti-HA
855 produced in rat (11867423001, Roche), anti-K48 produced in rabbit (ZRB2150, Millipore
856 Sigma), and anti-Mab414 produced in mouse. Secondary antibodies were conjugated to
857 Alexa Fluor 568 (Life Technologies).

858 Epifluorescence images were captured using a Zeiss AXIO Observer Z1 inverted
859 fluorescence microscope using a ×63/1.40 oil objective. All confocal images were
860 acquired at Yale's Imaging Facility on Science Hill using a Zeiss LSM 880 laser
861 scanning confocal microscope with a C Plan-Apochromat ×63/1.40 oil objective. Images

862 were analyzed with *CondenScreen* and compiled for publication using Zeiss v3.4 and
863 Image J (Fiji) software.
864
865 **Validation of hit compounds**
866 Based on the dose-response results, we focused validation efforts on four compounds:
867 PZ (IC_{50} = 601.6 nM), celastrol (IC_{50} = 124.8 nM), disulfiram (IC_{50} = 10.8 μM), and
868 plumbagin (IC_{50} = 7.41 μM). To determine the optimal concentrations using a tighter set
869 of concentrations, MLF2-GFP expression was induced in 4TorKO HeLa cells (0.5 μg/mL
870 Dox, 8 h) and treated with either DMSO or one of four different
871 concentrations/compounds: 2x below the IC_{50}, IC_{50}, 2x above the IC_{50}, or 8x above the
872 IC_{50}. Compounds were added 3 h before fixation. For each condition, images (n > 50
873 cells/condition) were captured on a Zeiss AXIO Observer Z1 inverted fluorescence
874 microscope and analyzed using our *CellProfiler* pipeline. *R Studio* was used to cluster
875 MLF2-GFP foci into single cells grouped by treatment condition and construct jittered
876 boxplots. To determine the effect of compounds on K48-ubiquitinylated proteins
877 accumulation within condensates, we incubated cells with compounds for 3h (PZ, 3 μM;
878 celastrol, 3 μM; plumbagin, 15 μM; and disulfiram, 50 μM), followed by
879 immunofluorescent labeling with α-K48 primary antibody (1:500). Images were captured
880 and analyzed as described above. High-resolution representative images were acquired
881 using confocal microscopy (Zeiss LSM 880).
882 To determine differential protein expression, we constructed a 4TorKO HeLa cell
883 line with a Dox inducible MLF2-GFP-P2A-tdTomato marker, where tdTomato serves as
884 a diffuse secondary reporter for global protein synthesis. Cells were induced with 0.5
885 μg/mL Dox and treated with their respective compound for 6 h before fixation. Imaging
886 was carried out on the Zeiss AXIO and the Zeiss LSM 880 for quantification and high-
887 resolution, respectively. The integrated intensity of both GFP and tdTomato signal was
888 quantified (n > 50) with *CellProfiler* and *R*. To ascertain the relative influence of global
889 protein expression inhibition on the reduced number of MLF2-GFP foci, we derived the
890 following empirical equation to determine the condensate normalization score ($CNS =$
891 $(1 - (\overline{F}_x/\overline{F}_{DMSO})) \cdot (\overline{II}_x/\overline{II}_{DMSO})^2$, where $\overline{F}$ and $\overline{II}$ are the average number of NE
892 condensate foci and the average integrated intensity of the secondary marker,

respectively, for either the tested condition (x) or the DMSO control. The scoring system is designed to prioritize compounds that robustly reduce condensate foci without affecting tdTomato reporter expression and penalizes those that have little effect on foci or substantially suppress reporter expression. It penalizes modulators that have either little effect on foci or those that substantially suppress reporter expression. The weighting is particularly sensitive to changes in tdTomato intensity due to the squared exponent, thereby strongly penalizing compounds which impair protein expression. A low CNS-score suggests non-specific or off-target activities, such as decreased transcription, translation, or increased biomarker turnover.

Guide RNA-mediated KO experiments:

Each gene KO consisted of a pool of 3 distinct, synthetic sgRNAs, each targeting a different portion of the same gene (Horizon Discovery). Gene knock-out was conducted using either DharmaFECT or DharmaFECT Duo Transfection Reagent (Horizon discovery) using Tor1A KO + Tor1B KO as a positive control and NT sgRNAs as a negative control (Horizon discovery). The sgRNAs were transfected into stably-integrated Cas9 expressing HeLa cells and incubated for 72 h with a final sgRNA concentration of 25 nM. The cells were treated with 0.5 µg/mL Dox to induce the MLF2-GFP expression 6 h before fixation.

The pooled sgRNA sequences used were: ZNF335 sgRNAs (SG-016173, Horizon Discovery), GATGATGGTGGCGCCGCCCA, GTGCCAGGGTGATCTGCGGT, ACCAAATGGACCAGGTGCCA; RNF26 sgRNAs (SG-007060, Horizon Discovery), TATCATGGAGGCAGTGTACC, GCATGACATCCTCTCGAAGA, CAGTGTACCTGGTAGTGAAT; TOR1A sgRNAs (SG-011023, Horizon Discovery) TATGAACATGGCTTTCTGGT, TATGAAGATGGACCTCGCAC, AGTGCAATGTGGCCACAAAC; and TOR1B sgRNAs (SG-014203-01, Horizon Discovery), AGTGCAGAGTCGATACAAAC, CAAGAGCCGTCTTAGTTATA, TTGGCGGGCCGGAAGAAAGA.

Sodium arsenite and 1,6-hexanediol treatment

To determine the co-localization of MLF2 into stress granules, Lipofectamine 2000 was used to co-transfect MLF2 with either G3BP1-GFP (Addgene #135997) + MLF2-HA, TDP-43-tdTomato-HA (Addgene #28205) + MLF2-GFP-Flag (Addgene #230986), or

924 pcDNA3.1 empty vector control. Next day (16 h post-transfection), HeLa cells were
925 treated with $NaAsO_2$ (0.5 mM) and incubated for 30 min. Cells were fixed for 20 min with
926 PFA, stained, and imaged as outlined. For hexanediol treatment, cells were incubated in
927 complete DMEM medium containing 5% (v/v) 1,6-hexanediol (CAS 629-11-8) or 5%
928 (v/v) 2,5-hexanediol (CAS 2935-44-6) for 5 min before fixation and immunofluorescence
929 staining as described.
930
931 **Electron Microscopy:**
932 Electron microscopy was performed at the Center for Cellular and Molecular Imaging,
933 Yale School of Medicine, using our previously described workflow (Laudermilch et al.,
934 2016).[28] Briefly, cells were fixed for 1 h in 2.5% glutaraldehyde in 0.1 M sodium
935 cacodylate buffer, pH 7.4. Cells were rinsed and scraped in 1% gelatin and centrifuged
936 in a 2% agar solution. Chilled cell blocks were processed with osmium and
937 thiocarbohydrazide-osmium ligand as described (West et al., 2010), and samples were
938 embedded in Durcupan ACM resin (Electron Microscopy Science). Polymerization was
939 performed by incubating samples at 60°C overnight. These blocks were cut into 60 nm
940 sections with a Leica UltraCut UC7 and stained with 2% uranyl acetate and lead citrate
941 on Formvar/carbon-coated grids. Samples were visualized with an FEI Tecnai Biotwin
942 TEM at 80 kV, and pictures were taken with Morada CCD and iTEM (Olympus)
943 software.
944
945 **Statistical Analysis:**
946 Genetic and chemical hits were ranked based on their Z-Score or BZ-Score,
947 respectively, as described in detail within the *Statistical analysis of screen data* section.
948 GO pathways were ranked using a p-value cutoff of 0.05, after controlling for false
949 discovery rate using the Benjamini-Hochberg (BH) procedure. The significance of all
950 chemical and genetic validation experiments was assessed using a Bonferroni multiple
951 comparisons corrected Mann–Whitney U test with > 50 cells per condition individually
952 analyzed. Unless otherwise stated, asterisks indicate Bonferroni corrected p-values
953 using a Mann-Whitney U test where *: $p <= 0.05$; **: $p <= 0.01$; ***: $p <= 0.001$; and ****:
954 $p <= 0.0001$.

Acknowledgements

Funding:

This work was funded by the DOD PR200788 (C.S.) 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 (D.P.), the European Molecular Biology Organization (EMBO) ALTF 910-2022 (E.F.F.K), and the Dutch Research Council (NWO) 019.222EN.007 (E.F.E.K.) 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 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
Data Analysis: DP
Supervision: CS
Writing—original draft: DP
Writing—review & editing: DP, CS, EFEK, CM, SM, ND, YVS

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.git .

988 At the time of publication, all raw screening images will be uploaded to the EMBL-EBI
989 BioImage Archive: https://www.ebi.ac.uk/bioimage-archive/.
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018

1019

1020 **References**

1021 1. Prophet, S.M., Rampello, A.J., Niescier, R.F., Gentile, J.E., Mallik, S., Koleske, A.J.,
1022 and Schlieker, C. (2022). Atypical nuclear envelope condensates linked to
1023 neurological disorders reveal nucleoporin-directed chaperone activities. Nat. Cell
1024 Biol. *24*, 1630–1641. https://doi.org/10.1038/s41556-022-01001-y.

1025 2. Mackenzie, I.R., Nicholson, A.M., Sarkar, M., Messing, J., Purice, M.D., Pottier, C.,
1026 Annu, K., Baker, M., Perkerson, R.B., Kurti, A., et al. (2017). TIA1 Mutations in
1027 Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase
1028 Separation and Alter Stress Granule Dynamics. Neuron *95*, 808-816.e9.
1029 https://doi.org/10.1016/j.neuron.2017.07.025.

1030 3. Alberti, S., and Dormann, D. (2019). Liquid-Liquid Phase Separation in Disease.
1031 Annu. Rev. Genet. *53*, 171–194. https://doi.org/10.1146/annurev-genet-112618-
1032 043527.

1033 4. Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler,
1034 M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P. (2005). Stress
1035 granules and processing bodies are dynamically linked sites of mRNP remodeling.
1036 J. Cell Biol. *169*, 871–884. https://doi.org/10.1083/jcb.200502088.

1037 5. Portz, B., Lee, B.L., and Shorter, J. (2021). FUS and TDP-43 Phases in Health and
1038 Disease. Trends Biochem. Sci. *46*, 550–563.
1039 https://doi.org/10.1016/j.tibs.2020.12.005.

1040 6. Alberti, S., Gladfelter, A., and Mittag, T. (2019). Considerations and Challenges in
1041 Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell *176*,
1042 419–434. https://doi.org/10.1016/j.cell.2018.12.035.

1043 7. Mittag, T., and Pappu, R.V. (2022). A conceptual framework for understanding
1044 phase separation and addressing open questions and challenges. Mol. Cell *82*,
1045 2201–2214. https://doi.org/10.1016/j.molcel.2022.05.018.

1046 8. Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C.,
1047 Gharakhani, J., Jülicher, F., and Hyman, A.A. (2009). Germline P granules are
1048 liquid droplets that localize by controlled dissolution/condensation. Science *324*,
1049 1729–1732. https://doi.org/10.1126/science.1172046.

1050 9. Schmidt, H.B., and Görlich, D. (2016). Transport Selectivity of Nuclear Pores, Phase
1051 Separation, and Membraneless Organelles. Trends Biochem. Sci. *41*, 46–61.
1052 https://doi.org/10.1016/j.tibs.2015.11.001.

1053 10. Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L., Llaguno, M.,
1054 Hollingsworth, J.V., King, D.S., Banani, S.F., et al. (2012). Phase transitions in the
1055 assembly of multivalent signalling proteins. Nature *483*, 336–340.
1056 https://doi.org/10.1038/nature10879.

1057 11. Wolozin, B., and Ivanov, P. (2019). Stress granules and neurodegeneration. Nat.
1058 Rev. Neurosci. *20*, 649–666. https://doi.org/10.1038/s41583-019-0222-5.

1059 12. Alberti, S., and Hyman, A.A. (2021). Biomolecular condensates at the nexus of
1060 cellular stress, protein aggregation disease and ageing. Nat. Rev. Mol. Cell Biol.
1061 *22*, 196–213. https://doi.org/10.1038/s41580-020-00326-6.

1062 13. Yan, X., Kuster, D., Mohanty, P., Nijssen, J., Pombo-García, K., Morato, J.G.,
1063 Rizuan, A., Franzmann, T.M., Sergeeva, A., Ly, A.M., et al. (2025). Intra-
1064 condensate demixing of TDP-43 inside stress granules generates pathological
1065 aggregates. Cell *188*. https://doi.org/10.1016/j.cell.2025.04.039.

1066 14. Vendruscolo, M., and Fuxreiter, M. (2022). Protein condensation diseases:
1067 therapeutic opportunities. Nat. Commun. *13*, 5550. https://doi.org/10.1038/s41467-
1068 022-32940-7.

1069 15. Ng, S.C., and Görlich, D. (2022). A simple thermodynamic description of phase
1070 separation of Nup98 FG domains. Nat. Commun. *13*, 6172.
1071 https://doi.org/10.1038/s41467-022-33697-9.

1072 16. Kozai, T., Fernandez-Martinez, J., van Eeuwen, T., Gallardo, P., Kapinos, L.E.,
1073 Mazur, A., Zhang, W., Tempkin, J., Panatala, R., Delgado-Izquierdo, M., et al.
1074 (2023). Dynamic molecular mechanism of the nuclear pore complex permeability
1075 barrier. bioRxiv, 2023.03.31.535055. https://doi.org/10.1101/2023.03.31.535055.

1076 17. Cowburn, D., and Rout, M. (2023). Improving the hole picture: towards a consensus
1077 on the mechanism of nuclear transport. Biochem. Soc. Trans. *51*, 871–886.
1078 https://doi.org/10.1042/BST20220494.

1079 18. Frey, S., Richter, R.P., and Görlich, D. (2006). FG-rich repeats of nuclear pore
1080 proteins form a three-dimensional meshwork with hydrogel-like properties. Science
1081 *314*, 815–817. https://doi.org/10.1126/science.1132516.

1082 19. Kim, S., Phan, S., Tran, H.T., Shaw, T.R., Shahmoradian, S.H., Ellisman, M.H.,
1083 Veatch, S.L., Barmada, S.J., Pappas, S.S., and Dauer, W.T. (2024). TorsinA is
1084 essential for neuronal nuclear pore complex localization and maturation. Nat. Cell
1085 Biol. *26*, 1482–1495. https://doi.org/10.1038/s41556-024-01480-1.

1086 20. Rampello, A.J., Laudermilch, E., Vishnoi, N., Prophet, S.M., Shao, L., Zhao, C.,
1087 Lusk, C.P., and Schlieker, C. (2020). Torsin ATPase deficiency leads to defects in
1088 nuclear pore biogenesis and sequestration of MLF2. J. Cell Biol. *219*.
1089 https://doi.org/10.1083/jcb.201910185.

1090 21. Fan, Y., Si, Z., Wang, L., and Zhang, L. (2023). DYT-TOR1A dystonia: an update on
1091 pathogenesis and treatment. Front. Neurosci. *17*, 1216929.
1092 https://doi.org/10.3389/fnins.2023.1216929.

1093 22. Balint, B., Mencacci, N.E., Valente, E.M., Pisani, A., Rothwell, J., Jankovic, J.,
1094 Vidailhet, M., and Bhatia, K.P. (2018). Dystonia. Nat. Rev. Dis. Primer *4*, 1–23.
1095 https://doi.org/10.1038/s41572-018-0023-6.

1096 23. Ozelius, L.J., Hewett, J.W., Page, C.E., Bressman, S.B., Kramer, P.L., Shalish, C.,
1097 de Leon, D., Brin, M.F., Raymond, D., Corey, D.P., et al. (1997). The early-onset
1098 torsion dystonia gene (DYT1) encodes an ATP-binding protein. Nat. Genet. *17*, 40–
1099 48. https://doi.org/10.1038/ng0997-40.

1100 24. Liang, C.-C., Tanabe, L.M., Jou, S., Chi, F., and Dauer, W.T. (2014). TorsinA
1101 hypofunction causes abnormal twisting movements and sensorimotor circuit
1102 neurodegeneration. J. Clin. Invest. *124*, 3080–3092.
1103 https://doi.org/10.1172/JCI72830.

1104 25. Zhao, C., Brown, R.S.H., Chase, A.R., Eisele, M.R., and Schlieker, C. (2013).
1105 Regulation of Torsin ATPases by LAP1 and LULL1. Proc. Natl. Acad. Sci. U. S. A.
1106 *110*, E1545–E1554. https://doi.org/10.1073/pnas.1300676110.

1107 26. Brown, R.S.H., Zhao, C., Chase, A.R., Wang, J., and Schlieker, C. (2014). The
1108 mechanism of Torsin ATPase activation. Proc. Natl. Acad. Sci. U. S. A. *111*,
1109 E4822–E4831. https://doi.org/10.1073/pnas.1415271111.

1110 27. Demircioglu, F.E., Sosa, B.A., Ingram, J., Ploegh, H.L., and Schwartz, T.U. (2016).
1111 Structures of TorsinA and its disease-mutant complexed with an activator reveal
1112 the molecular basis for primary dystonia. eLife *5*, e17983.
1113 https://doi.org/10.7554/eLife.17983.

1114 28. Laudermilch, E., Tsai, P.-L., Graham, M., Turner, E., Zhao, C., and Schlieker, C.
1115 (2016). Dissecting Torsin/cofactor function at the nuclear envelope: a genetic
1116 study. Mol. Biol. Cell *27*, 3964–3971. https://doi.org/10.1091/mbc.E16-07-0511.

1117 29. Goodchild, R.E., Kim, C.E., and Dauer, W.T. (2005). Loss of the Dystonia-
1118 Associated Protein TorsinA Selectively Disrupts the Neuronal Nuclear Envelope.
1119 Neuron *48*, 923–932. https://doi.org/10.1016/j.neuron.2005.11.010.

1120 30. Kuiper, E.F.E., Gallardo, P., Bergsma, T., Mari, M., Kolbe Musskopf, M., Kuipers, J.,
1121 Giepmans, B.N.G., Steen, A., Kampinga, H.H., Veenhoff, L.M., et al. (2022). The
1122 chaperone DNAJB6 surveils FG-nucleoporins and is required for interphase
1123 nuclear pore complex biogenesis. Nat. Cell Biol. *24*, 1584–1594.
1124 https://doi.org/10.1038/s41556-022-01010-x.

1125 31. Martin, E.W., Holehouse, A.S., Peran, I., Farag, M., Incicco, J.J., Bremer, A., Grace,
1126 C.R., Soranno, A., Pappu, R.V., and Mittag, T. (2020). Valence and patterning of
1127 aromatic residues determine the phase behavior of prion-like domains. Science
1128 *367*, 694–699. https://doi.org/10.1126/science.aaw8653.

1129 32. Colombrita, C., Zennaro, E., Fallini, C., Weber, M., Sommacal, A., Buratti, E., Silani,
1130 V., and Ratti, A. (2009). TDP-43 is recruited to stress granules in conditions of

1131 oxidative insult. J. Neurochem. *111*, 1051–1061. https://doi.org/10.1111/j.1471-
1132 4159.2009.06383.x.

1133 33. Palangi, F., Samuel, S.M., Thompson, I.R., Triggle, C.R., and Emara, M.M. (2017).
1134 Effects of oxidative and thermal stresses on stress granule formation in human
1135 induced pluripotent stem cells. PLOS ONE *12*, e0182059.
1136 https://doi.org/10.1371/journal.pone.0182059.

1137 34. Fang, M.Y., Markmiller, S., Vu, A.Q., Javaherian, A., Dowdle, W.E., Jolivet, P.,
1138 Bushway, P.J., Castello, N.A., Baral, A., Chan, M.Y., et al. (2019). Small-Molecule
1139 Modulation of TDP-43 Recruitment to Stress Granules Prevents Persistent TDP-43
1140 Accumulation in ALS/FTD. Neuron *103*, 802-819.e11.
1141 https://doi.org/10.1016/j.neuron.2019.05.048.

1142 35. Uechi, H., Sridharan, S., Nijssen, J., Bilstein, J., Iglesias-Artola, J.M., Kishigami, S.,
1143 Casablancas-Antras, V., Poser, I., Martinez, E.J., Boczek, E., et al. (2025). Small-
1144 molecule dissolution of stress granules by redox modulation benefits ALS models.
1145 Nat. Chem. Biol. *21*, 1–12. https://doi.org/10.1038/s41589-025-01893-5.

1146 36. Bergsma, T., Steen, A., Kamenz, J.L., Otto, T.A., Gallardo, P., and Veenhoff, L.M.
1147 (2024). Imaging-based quantitative assessment of biomolecular condensates
1148 in vitro and in cells. J. Biol. Chem. *301*, 108130.
1149 https://doi.org/10.1016/j.jbc.2024.108130.

1150 37. Baggett, D.W., Medyukhina, A., Tripathi, S., Shirnekhi, H.K., Wu, H., Pounds, S.B.,
1151 Khairy, K., and Kriwacki, R. (2022). An Image Analysis Pipeline for Quantifying the
1152 Features of Fluorescently-Labeled Biomolecular Condensates in Cells. Front.
1153 Bioinforma. *2*. https://doi.org/10.3389/fbinf.2022.897238.

1154 38. Stirling, D.R., Swain-Bowden, M.J., Lucas, A.M., Carpenter, A.E., Cimini, B.A., and
1155 Goodman, A. (2021). CellProfiler 4: improvements in speed, utility and usability.
1156 BMC Bioinformatics *22*, 433. https://doi.org/10.1186/s12859-021-04344-9.

1157 39. Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O.,
1158 Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., et al. (2006). CellProfiler:
1159 image analysis software for identifying and quantifying cell phenotypes. Genome
1160 Biol. *7*, R100. https://doi.org/10.1186/gb-2006-7-10-r100.

1161 40. Zhang, J.-H., Chung, T.D.Y., and Oldenburg, K.R. (1999). A Simple Statistical
1162 Parameter for Use in Evaluation and Validation of High Throughput Screening
1163 Assays. J. Biomol. Screen. *4*, 67–73.
1164 https://doi.org/10.1177/108705719900400206.

1165 41. Liu, Z., Chen, O., Wall, J.B.J., Zheng, M., Zhou, Y., Wang, L., Ruth Vaseghi, H.,
1166 Qian, L., and Liu, J. (2017). Systematic comparison of 2A peptides for cloning
1167 multi-genes in a polycistronic vector. Sci. Rep. *7*, 2193.
1168 https://doi.org/10.1038/s41598-017-02460-2.

1169 42. Kroschwald, S., Maharana, S., and Simon, A. (2017). Hexanediol: a chemical probe
1170 to investigate the material properties of membrane-less compartments. Matters.
1171 https://doi.org/10.19185/matters.201702000010.

1172 43. Wang, W.-M., Liu, Z., Liu, A.-J., Wang, Y.-X., Wang, H.-G., An, D., Heng, B., Xie, L.-
1173 H., Duan, J.-L., and Liu, Y.-Q. (2015). The Zinc Ion Chelating Agent TPEN
1174 Attenuates Neuronal Death/apoptosis Caused by Hypoxia/ischemia Via Mediating
1175 the Pathophysiological Cascade Including Excitotoxicity, Oxidative Stress, and
1176 Inflammation. CNS Neurosci. Ther. *21*, 708–717.
1177 https://doi.org/10.1111/cns.12428.

1178 44. Martin, H.L., Adams, M., Higgins, J., Bond, J., Morrison, E.E., Bell, S.M., Warriner,
1179 S., Nelson, A., and Tomlinson, D.C. (2014). High-Content, High-Throughput
1180 Screening for the Identification of Cytotoxic Compounds Based on Cell Morphology
1181 and Cell Proliferation Markers. PLOS ONE *9*, e88338.
1182 https://doi.org/10.1371/journal.pone.0088338.

1183 45. Amberger, J.S., Bocchini, C.A., Schiettecatte, F., Scott, A.F., and Hamosh, A.
1184 (2015). OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online
1185 catalog of human genes and genetic disorders. Nucleic Acids Res. *43*, D789–
1186 D798. https://doi.org/10.1093/nar/gku1205.

1187 46. Rubinstein, W.S., Maglott, D.R., Lee, J.M., Kattman, B.L., Malheiro, A.J., Ovetsky,
1188 M., Hem, V., Gorelenkov, V., Song, G., Wallin, C., et al. (2013). The NIH genetic
1189 testing registry: a new, centralized database of genetic tests to enable access to
1190 comprehensive information and improve transparency. Nucleic Acids Res. *41*,
1191 D925–D935. https://doi.org/10.1093/nar/gks1173.

1192 47. Patel, D.D., Gripp, K.W., Wadman, E., Mishra, I., and Kandula, V. (2024). Novel
1193 biallelic ZNF335 variant causing primary microcephaly: A case report and
1194 radiological review. Am. J. Med. Genet. A. *194*, e63593.
1195 https://doi.org/10.1002/ajmg.a.63593.

1196 48. Yang, Y.J., Baltus, A.E., Mathew, R.S., Murphy, E.A., Evrony, G.D., Gonzalez, D.M.,
1197 Wang, E.P., Marshall-Walker, C.A., Barry, B.J., Murn, J., et al. (2012).
1198 Microcephaly Gene Links Trithorax and REST/NRSF to Control Neural Stem Cell
1199 Proliferation and Differentiation. Cell *151*, 1097–1112.
1200 https://doi.org/10.1016/j.cell.2012.10.043.

1201 49. Li, D., Quan, Z., Ni, J., Li, H., and Qing, H. (2023). The many faces of the zinc finger
1202 protein 335 in brain development and immune system. Biomed. Pharmacother.
1203 *165*, 115257. https://doi.org/10.1016/j.biopha.2023.115257.

1204 50. Stouffs, K., Stergachis, A.B., Vanderhasselt, T., Dica, A., Janssens, S., Vandervore,
1205 L., Gheldof, A., Bodamer, O., Keymolen, K., Seneca, S., et al. (2018). Expanding
1206 the clinical spectrum of biallelic ZNF335 variants. Clin. Genet. *94*, 246–251.
1207 https://doi.org/10.1111/cge.13260.

1208 51. Sato, R., Takanashi, J., Tsuyusaki, Y., Kato, M., Saitsu, H., Matsumoto, N., and
1209 Takahashi, T. (2016). Association Between Invisible Basal Ganglia and ZNF335
1210 Mutations: A Case Report. Pediatrics *138*, e20160897.
1211 https://doi.org/10.1542/peds.2016-0897.

1212 52. Caglayan, A.O., Yaghouti, K., Kockaya, T., Kemer, D., Cankaya, T., Ameziane, N.,
1213 Cogulu, O., Coker, M., and Yalcinkaya, C. (2021). Biallelic ZNF335 mutations
1214 cause basal ganglia abnormality with progressive cerebral/cerebellar atrophy. J.
1215 Neurogenet. *35*, 23–28. https://doi.org/10.1080/01677063.2020.1833006.

1216 53. Nishida, T., Mihara, T., Horiki, T., and Ka, K. (2016). Anaesthesia and orphan
1217 disease: primary autosomal recessive microcephaly-10 caused by a mutation in the
1218 ZNF335 gene. Eur. J. Anaesthesiol. *33*, 543–545.
1219 https://doi.org/10.1097/EJA.0000000000000358.

1220 54. Theusch, E., Ting, F.Y., Qin, Y., Stevens, K., Naidoo, D., King, S.M., Yang, N.V.,
1221 Orr, J., Han, B.Y., Cyster, J.G., et al. (2024). Participant-derived cell line
1222 transcriptomic analyses and mouse studies reveal a role for ZNF335 in plasma
1223 cholesterol statin response. Genome Med. *16*, 93. https://doi.org/10.1186/s13073-
1224 024-01366-9.

1225 55. Cremer, T., Jongsma, M.L.M., Trulsson, F., Vertegaal, A.C.O., Neefjes, J., and
1226 Berlin, I. (2021). The ER-embedded UBE2J1/RNF26 ubiquitylation complex exerts
1227 spatiotemporal control over the endolysosomal pathway. Cell Rep. *34*.
1228 https://doi.org/10.1016/j.celrep.2020.108659.

1229 56. Cremer, T., Voortman, L.M., Bos, E., Jongsma, M.L., ter Haar, L.R., Akkermans,
1230 J.J., Talavera Ormeño, C.M., Wijdeven, R.H., de Vries, J., Kim, R.Q., et al. (2023).
1231 RNF26 binds perinuclear vimentin filaments to integrate ER and endolysosomal
1232 responses to proteotoxic stress. EMBO J. *42*, e111252.
1233 https://doi.org/10.15252/embj.2022111252.

1234 57. Lecordier, L., Heo, P., Graversen, J.H., Hennig, D., Skytthe, M.K., Cornet d'Elzius,
1235 A., Pincet, F., Pérez-Morga, D., and Pays, E. (2023). Apolipoproteins L1 and L3
1236 control mitochondrial membrane dynamics. Cell Rep. *42*, 113528.
1237 https://doi.org/10.1016/j.celrep.2023.113528.

1238 58. Gaudet, R.G., Zhu, S., Halder, A., Kim, B.-H., Bradfield, C.J., Huang, S., Xu, D.,
1239 Mamiñska, A., Nguyen, T.N., Lazarou, M., et al. (2021). A human apolipoprotein L
1240 with detergent-like activity kills intracellular pathogens. Science *373*.
1241 https://doi.org/10.1126/science.abf8113.

1242 59. Hondele, M., Sachdev, R., Heinrich, S., Wang, J., Vallotton, P., Fontoura, B.M.A.,
1243 and Weis, K. (2019). DEAD-box ATPases are global regulators of phase-separated
1244 organelles. Nature *573*, 144–148. https://doi.org/10.1038/s41586-019-1502-y.

1245 60. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis,
1246 A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene Ontology: tool for
1247 the unification of biology. Nat. Genet. *25*, 25–29. https://doi.org/10.1038/75556.

1248 61. Shroff, K., Caffall, Z.F., and Calakos, N. (2021). DYT-TOR1A subcellular proteomics
1249 reveals selective vulnerability of the nuclear proteome to cell stress. Neurobiol. Dis.
1250 *158*, 105464. https://doi.org/10.1016/j.nbd.2021.105464.

1251 62. Ding, B., Tang, Y., Ma, S., Akter, M., Liu, M.-L., Zang, T., and Zhang, C.-L. (2021).
1252 Disease Modeling with Human Neurons Reveals LMNB1 Dysregulation Underlying
1253 DYT1 Dystonia. J. Neurosci. *41*, 2024–2038.
1254 https://doi.org/10.1523/JNEUROSCI.2507-20.2020.

1255 63. Pappas, S.S., Liang, C.-C., Kim, S., Rivera, C.O., and Dauer, W.T. (2018). TorsinA
1256 dysfunction causes persistent neuronal nuclear pore defects. Hum. Mol. Genet. *27*,
1257 407–420. https://doi.org/10.1093/hmg/ddx405.

1258 64. Martin, C.-A., Murray, J.E., Carroll, P., Leitch, A., Mackenzie, K.J., Halachev, M.,
1259 Fetit, A.E., Keith, C., Bicknell, L.S., Fluteau, A., et al. (2016). Mutations in genes
1260 encoding condensin complex proteins cause microcephaly through decatenation
1261 failure at mitosis. Genes Dev. *30*, 2158–2172.
1262 https://doi.org/10.1101/gad.286351.116.

1263 65. Trimborn, M., Bell, S.M., Felix, C., Rashid, Y., Jafri, H., Griffiths, P.D., Neumann,
1264 L.M., Krebs, A., Reis, A., Sperling, K., et al. (2004). Mutations in Microcephalin
1265 Cause Aberrant Regulation of Chromosome Condensation. Am. J. Hum. Genet.
1266 *75*, 261–266. https://doi.org/10.1086/422855.

1267 66. Jean, F., Stuart, A., and Tarailo-Graovac, M. (2020). Dissecting the Genetic and
1268 Etiological Causes of Primary Microcephaly. Front. Neurol. *11*.
1269 https://doi.org/10.3389/fneur.2020.570830.

1270 67. He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep Residual Learning for Image
1271 Recognition. arXiv. https://doi.org/1512.03385.

1272 68. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. ImageNet: A large-
1273 scale hierarchical image database. In 2009 IEEE Conference on Computer Vision
1274 and Pattern Recognition, pp. 248–255.
1275 https://doi.org/10.1109/CVPR.2009.5206848.

1276 69. Sundararajan, M., Taly, A., and Yan, Q. (2017). Axiomatic Attribution for Deep
1277 Networks. arXiv.

1278 70. McInnes, L., Healy, J., and Melville, J. (2020). UMAP: Uniform Manifold
1279 Approximation and Projection for Dimension Reduction. Preprint at arXiv,
1280 https://doi.org/10.48550/arXiv.1802.03426
1281 https://doi.org/10.48550/arXiv.1802.03426.

1282 71. Hirano, T. (2005). Condensins: Organizing and Segregating the Genome. Curr. Biol.
1283 *15*, R265–R275. https://doi.org/10.1016/j.cub.2005.03.037.

1284 72. Drozdz, M.M., and Vaux, D.J. (2016). Shared mechanisms in physiological and
1285 pathological nucleoplasmic reticulum formation. Nucleus *8*, 34–45.
1286 https://doi.org/10.1080/19491034.2016.1252893.

1287 73. Drozdz, M.M., Jiang, H., Pytowski, L., Grovenor, C., and Vaux, D.J. (2017).
1288 Formation of a nucleoplasmic reticulum requires de novo assembly of nascent
1289 phospholipids and shows preferential incorporation of nascent lamins. Sci. Rep. *7*,
1290 7454. https://doi.org/10.1038/s41598-017-07614-w.

1291 74. Mathieu, C., Pappu, R.V., and Taylor, J.P. (2020). Beyond aggregation: Pathological
1292 phase transitions in neurodegenerative disease. Science *370*, 56–60.
1293 https://doi.org/10.1126/science.abb8032.

1294 75. Schludi, M.H., Becker, L., Garrett, L., Gendron, T.F., Zhou, Q., Schreiber, F.,
1295 Popper, B., Dimou, L., Strom, T.M., Winkelmann, J., et al. (2017). Spinal poly-GA
1296 inclusions in a C9orf72 mouse model trigger motor deficits and inflammation
1297 without neuron loss. Acta Neuropathol. (Berl.) *134*, 241–254.
1298 https://doi.org/10.1007/s00401-017-1711-0.

1299 76. Kazemi-Esfarjani, P., and Benzer, S. (2002). Suppression of polyglutamine toxicity
1300 by a Drosophila homolog of myeloid leukemia factor 1. Hum. Mol. Genet. *11*,
1301 2657–2672. https://doi.org/10.1093/hmg/11.21.2657.

1302 77. Kim, W.-Y., Fayazi, Z., Bao, X., Higgins, D., and Kazemi-Esfarjani, P. (2005).
1303 Evidence for sequestration of polyglutamine inclusions by Drosophila myeloid
1304 leukemia factor. Mol. Cell. Neurosci. *29*, 536–544.
1305 https://doi.org/10.1016/j.mcn.2005.04.005.

1306 78. Mitrea, D.M., Mittasch, M., Gomes, B.F., Klein, I.A., and Murcko, M.A. (2022).
1307 Modulating biomolecular condensates: a novel approach to drug discovery. Nat.
1308 Rev. Drug Discov. *21*, 841–862. https://doi.org/10.1038/s41573-022-00505-4.

1309 79. Patel, A., Mitrea, D., Namasivayam, V., Murcko, M.A., Wagner, M., and Klein, I.A.
1310 (2022). Principles and functions of condensate modifying drugs. Front. Mol. Biosci.
1311 *9*. https://doi.org/10.3389/fmolb.2022.1007744.

1312 80. Jongsma, M.L.M., Berlin, I., Wijdeven, R.H.M., Janssen, L., Janssen, G.M.C.,
1313 Garstka, M.A., Janssen, H., Mensink, M., van Veelen, P.A., Spaapen, R.M., et al.
1314 (2016). An ER-Associated Pathway Defines Endosomal Architecture for Controlled
1315 Cargo Transport. Cell *166*, 152–166. https://doi.org/10.1016/j.cell.2016.05.078.

1316 81. Shaheen, R., Maddirevula, S., Ewida, N., Alsahli, S., Abdel-Salam, G.M.H., Zaki,
1317 M.S., Tala, S.A., Alhashem, A., Softah, A., Al-Owain, M., et al. (2019). Genomic
1318 and phenotypic delineation of congenital microcephaly. Genet. Med. *21*, 545–552.
1319 https://doi.org/10.1038/s41436-018-0140-3.

1320 82. Pavinato, L., Carestiato, S., Trajkova, S., Sorasio, L., Mantovani, G., De Sanctis, L.,
1321 Kerkhof, J., McConkey, H., Rzasa, J., Todd, E., et al. (2025). Skipping of Exon 20
1322 in EP300: A Novel Variant Linked to Rubinstein–Taybi Syndrome With Atypical and
1323 Severe Clinical Manifestations. Clin. Genet. *107*, 354–358.
1324 https://doi.org/10.1111/cge.14654.

1325 83. Roessler, E., Belloni, E., Gaudenz, K., Jay, P., Berta, P., Scherer, S.W., Tsui, L.C.,
1326 and Muenke, M. (1996). Mutations in the human Sonic Hedgehog gene cause
1327 holoprosencephaly. Nat. Genet. *14*, 357–360. https://doi.org/10.1038/ng1196-357.

1328 84. Loo, C.K.C., Pearen, M.A., and Ramm, G.A. (2021). The Role of Sonic Hedgehog in
1329 Human Holoprosencephaly and Short-Rib Polydactyly Syndromes. Int. J. Mol. Sci.
1330 *22*, 9854. https://doi.org/10.3390/ijms22189854.

1331 85. Rump, P., Jazayeri, O., van Dijk-Bos, K.K., Johansson, L.F., van Essen, A.J.,
1332 Verheij, J.B.G.M., Veenstra-Knol, H.E., Redeker, E.J.W., Mannens, M.M.A.M.,
1333 Swertz, M.A., et al. (2016). Whole-exome sequencing is a powerful approach for
1334 establishing the etiological diagnosis in patients with intellectual disability and
1335 microcephaly. BMC Med. Genomics *9*, 7. https://doi.org/10.1186/s12920-016-0167-
1336 8.

1337 86. Tsai, P.-L., Cameron, C.J.F., Forni, M.F., Wasko, R.R., Naughton, B.S., Horsley, V.,
1338 Gerstein, M.B., and Schlieker, C. (2022). Dynamic quality control machinery that
1339 operates across compartmental borders mediates the degradation of mammalian
1340 nuclear membrane proteins. Cell Rep. *41*, 111675.
1341 https://doi.org/10.1016/j.celrep.2022.111675.

1342 87. Tsai, P.-L., Zhao, C., and Schlieker, C. (2019). Methodologies to monitor protein
1343 turnover at the inner nuclear membrane. Methods Enzymol. *619*, 47–69.
1344 https://doi.org/10.1016/bs.mie.2018.12.033.

1345 88. Ihaka, R., and Gentleman, R. (1996). R: A Language for Data Analysis and
1346 Graphics. J. Comput. Graph. Stat. *5*, 299–314.
1347 https://doi.org/10.1080/10618600.1996.10474713.

1348 89. Vincent, L., and Soille, P. (1991). Watersheds in digital spaces: an efficient algorithm
1349 based on immersion simulations. IEEE Trans. Pattern Anal. Mach. Intell. *13*, 583–
1350 598. https://doi.org/10.1109/34.87344.

1351 90. Brideau, C., Gunter, B., Pikounis, B., and Liaw, A. (2003). Improved Statistical
1352 Methods for Hit Selection in High-Throughput Screening. SLAS Discov. *8*, 634–
1353 647. https://doi.org/10.1177/1087057103258285.

1354 91. Shun, T.Y., Lazo, J.S., Sharlow, E.R., and Johnston, P.A. (2011). Identifying Actives
1355 from HTS Data Sets: Practical Approaches for the Selection of an Appropriate HTS
1356 Data-Processing Method and Quality Control Review. J. Biomol. Screen. *16*, 1–14.
1357 https://doi.org/10.1177/1087057110389039.

1358 92. Wang, Y., Xiao, J., Suzek, T.O., Zhang, J., Wang, J., and Bryant, S.H. PubChem: a
1359 public information system for analyzing bioactivities of small molecules.

1360 93. NCI/CADD Chemical Identifier Resolver
1361 https://cactus.nci.nih.gov/chemical/structure.

1362 94. Baell, J.B., and Holloway, G.A. (2010). New Substructure Filters for Removal of Pan
1363 Assay Interference Compounds (PAINS) from Screening Libraries and for Their
1364 Exclusion in Bioassays. J. Med. Chem. *53*, 2719–2740.
1365 https://doi.org/10.1021/jm901137j.

1366 95. Baell, J.B., and Nissink, J.W.M. (2018). Seven Year Itch: Pan-Assay Interference
1367 Compounds (PAINS) in 2017-Utility and Limitations. ACS Chem. Biol. *13*, 36–44.
1368 https://doi.org/10.1021/acschembio.7b00903.

1369 96. Lipinski, C.A. (2004). Lead- and drug-like compounds: the rule-of-five revolution.
1370 Drug Discov. Today Technol. *1*, 337–341.
1371 https://doi.org/10.1016/j.ddtec.2004.11.007.

1372 97. Wickham, H. (2011). ggplot2. WIREs Comput. Stat. *3*, 180–185.
1373 https://doi.org/10.1002/wics.147.

1374 98. Marcel, S., and Rodriguez, Y. (2010). Torchvision the machine-vision package of
1375 torch. In Proceedings of the 18th ACM international conference on Multimedia
1376 (ACM), pp. 1485–1488. https://doi.org/10.1145/1873951.1874254.

1377 99. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O.,
1378 Blondel, M., Müller, A., Nothman, J., Louppe, G., et al. (2018). Scikit-learn:
1379 Machine Learning in Python. Preprint at arXiv,
1380 https://doi.org/10.48550/arXiv.1201.0490 https://doi.org/10.48550/arXiv.1201.0490.

1381 100. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T.,
1382 Lin, Z., Gimelshein, N., Antiga, L., et al. (2019). PyTorch: An Imperative Style,
1383 High-Performance Deep Learning Library. Preprint at arXiv,
1384 https://doi.org/10.48550/arXiv.1912.01703
1385 https://doi.org/10.48550/arXiv.1912.01703.

1386 101. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau,
1387 D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0--
1388 Fundamental Algorithms for Scientific Computing in Python. Nat. Methods *17*, 261–
1389 272. https://doi.org/10.1038/s41592-019-0686-2.

1390 102. Shannon, C.E. (1948). A Mathematical Theory of Communication. Bell Syst. Tech.
1391 J. *27*, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

1392 103. Ritz, C., Baty, F., Streibig, J.C., and Gerhard, D. (2015). Dose-Response Analysis
1393 Using R. PLOS ONE *10*, e0146021. https://doi.org/10.1371/journal.pone.0146021.

1394 104. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W.,
1395 Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for
1396 interpreting omics data. The Innovation *2*.
1397 https://doi.org/10.1016/j.xinn.2021.100141.

1398 105. Galili, T., O'Callaghan, A., Sidi, J., and Sievert, C. (2018). heatmaply: an R
1399 package for creating interactive cluster heatmaps for online publishing.
1400 Bioinformatics *34*, 1600–1602. https://doi.org/10.1093/bioinformatics/btx657.

1401