Nature Communications

https://doi.org/10.1038/s41467-026-69781-7

Article in Press

# NEDAMSS syndrome-related truncating and missense mutations are associated with aberrant liquid-liquid phase separation of IRF2BPL

Received: 13 June 2024

Accepted: 6 February 2026

Cite this article as: Dell'Oca, M., Boggio Bozzo, S., Vaglietti, S. *et al.* NEDAMSS syndrome-related truncating and missense mutations are associated with aberrant liquid-liquid phase separation of IRF2BPL. *Nat Commun* (2026). https://doi.org/10.1038/s41467-026-69781-7

**Marco Dell'Oca, Stefania Boggio Bozzo, Serena Vaglietti, Chiara Marchetti, Chiara Di Luca, Pietro Munarin, Michael Nicoli, Rossella Indellicato, Davide Ravanelli, Giorgia Falanga, Angelo Iannielli, Mirko Luoni, Alessia Loffreda, Valeria Berno, Paolo Bianchini, Sarah Sertic, Antonio Conforti, Shima Rashidiani, Eleonora Aimaretti, Massimo Collino, Ilaria Cecere, Angelo Gallo, Federica Santoro, Diego Brancaccio, Samuele Rosso, Giovanna Di Nardo, Raffaele Pertusio, Federico Cesano, Francisco J. Monje Quiroga, Mirella Ghirardi, Silvestro Roatta, Luca Colnaghi & Ferdinando Fiumara**

# NEDAMSS syndrome-related truncating and missense mutations are associated with aberrant liquid-liquid phase separation of IRF2BPL

Marco Dell'Oca[1], Stefania Boggio Bozzo[1], Serena Vaglietti[1], Chiara Marchetti[1], Chiara Di Luca[2,3], Pietro Munarin[2,3], Michael Nicoli[2,3], Rossella Indellicato[2,3], Davide Ravanelli[2,3], Giorgia Falanga[2,3], Angelo Iannielli[3,4], Mirko Luoni[3,4], Alessia Loffreda[3], Valeria Berno[3], Paolo Bianchini[5], Sarah Sertic[6], Antonio Conforti[1], Shima Rashidiani[1], Eleonora Aimaretti[1], Massimo Collino[1], Ilaria Cecere[7], Angelo Gallo[7], Federica Santoro[8], Diego Brancaccio[8], Samuele Rosso[9], Giovanna Di Nardo[9], Raffaele Pertusio[1], Federico Cesano[7], Francisco J. Monje[10], Mirella Ghirardi[1], Silvestro Roatta[1], Luca Colnaghi[2,3]*, and Ferdinando Fiumara[1]*

[1]'*Rita Levi Montalcini*' Department of Neuroscience, University of Turin, 10125 Turin, Italy
[2]School of Medicine, *Vita-Salute San Raffaele* University, 20132 Milan, Italy
[3]IRCCS *San Raffaele* Scientific Institute, 20132 Milan, Italy
[4]National Research Council (CNR), Institute of Neuroscience, 20129, Milan, Italy
[5] Nanoscopy and NIC@IIT, CHT, Istituto Italiano di Tecnologia, 16152 Genoa, Italy
[6] Department of Biosciences, University of Milan, 20133 Milan, Italy
[7]Department of Chemistry, University of Turin, 10125 Turin, Italy
[8]Department of Pharmacy, University of Naples *Federico II*, 30131 Naples, Italy
[9]Department of Life Sciences and Systems Biology, University of Turin, 10123, Turin, Italy
[10]Department of Neurophysiology and Neuropharmacology,
Medical University of Vienna, 1090 Vienna, Austria

* *Correspondence*: ferdinando.fiumara@unito.it;
colnaghi.luca@hsr.it

**ABSTRACT**

Since 2018, truncating and missense mutations in the *IRF2BPL* gene have been associated with the neurodevelopmental NEDAMSS syndrome and other *IRF2BPL*-related disorders. These mutations mainly affect the gene region encoding the central portion of the IRF2BPL protein, whose physiopathological roles are poorly understood. Here we show that this region contains three low-complexity regions ($LCR_{1-3}$) and a higher-complexity domain (HCD) similar to DNA-binding domains. The polyA/polyQ-rich $LCR_1$, together with an upstream zinc finger domain, drives IRF2BPL liquid-liquid phase separation (LLPS) in both neuronal and non-neuronal cells. Disease-related truncating and missense mutations disrupt the physiological IRF2BPL LLPS in cells, leading to the formation of aberrant cytoplasmic condensates sequestering wild-type IRF2BPL from the nuclear compartment, activating molecular disease mechanisms, i.e., WNT1 upregulation, and altering neuronal cell electrophysiological properties. These findings identify aberrant LLPS as a key biophysical mechanism underlying the nuclear depletion of wild-type IRF2BPL, a fundamental pathogenic alteration underlying NEDAMSS, triggering downstream cellular and molecular disease mechanisms.

## INTRODUCTION

A recently identified neurodevelopmental syndrome with regression, abnormal movements, loss of speech, and seizures (NEDAMSS) has been linked to mutations in the intronless interferon regulatory factor 2 binding protein-like (*IRF2BPL*) gene[1,2]. Mutations in *IRF2BPL* can also lead to other diseases with heterogeneous clinical presentations, such as progressive myoclonus epilepsy, that have been categorized as *IRF2BPL*-related disorders[3,4,5].

The *IRF2BPL* gene encodes a protein whose structure and function are poorly understood, that was listed until recently in the Pharos Tdark database of understudied therapeutic targets[6]. Based on sequence homology and partial structural characterization, two phylogenetically conserved functional domains have been identified in the IRF2BPL protein, i.e., an N-terminal zinc finger (ZnF) domain involved in transcriptional regulation, and a C-terminal RING domain acting as a ubiquitin ligase[1,7,8,9]. Evidence indicates that IRF2BPL is physiologically involved in transcriptional regulation in primates[10], a process that may be altered in disease. Indeed, NEDAMSS-related IRF2BPL downregulation, or loss, enhances *WNT1* transcription in the nervous system affecting neuronal function and maintenance[8].

IRF2BPL also contains a largely uncharacterized central region (>600 residues) targeted by most of the known disease-related nonsense and missense mutations. This strongly suggests that it may have significant, yet unidentified, roles in IRF2BPL function. This central region, currently considered a 'variable' low-complexity region (LCR)[1,10,11] contains long polyA and polyQ repeats of undetermined function, three PEST sequences which may promote degradation, and a nuclear localization signal (NLS).

PolyA and polyQ repeats can regulate physiological protein function and interactions by forming coiled coils (CCs) which mediate protein-protein interactions and oligo-/polymerization [12,13,14,15,16]. The pathological expansion of these repeats may lead to neurodegenerative and neurodevelopmental disorders, such as Huntington disease[17].

Recent evidence indicates that both polyA and polyQ repeats, as well as CC structures in general, can mediate liquid-liquid phase separation (LLPS)[16,18,19,20,21], a biophysical process that drives

the reversible formation of membraneless, liquid-like protein condensates, thus regulating protein localization and function[22]. PolyA and polyQ CCs undergo LLPS and oligo-/polymerization in a length-dependent manner[16,20]. LLPS is also involved in the pathogenesis of neurological disorders and may facilitate the organization of proteins into fibrillary structures[16,23,24]. LLPS can also be promoted by other types of structured domains (e.g., ZnF)[25]and disordered LCRs[22], which are also present in IRF2BPL. However, the ability of IRF2BPL to undergo LLPS, and the possible role of its domains/regions in this process, have not yet been explored.

NEDAMSS and *IRF2BPL*-related disorders primarily result from IRF2BPL nonsense and frameshift mutations, leading to the production of variably truncated forms of the protein[1,2]. Some disease-related missense mutations, which appear to be associated with milder phenotypes, have also been identified[1,2,26,27,28]. Most disease-related mutations affect the central part of the IRF2BPL protein, and cause disease through gain-of-function and dominant-negative mechanisms, some of which have begun to be elucidated[8,27]. In particular, the upregulation of WNT1 has been recently identified as a key molecular mechanism of NEDAMSS[8].

To better define the physiology of IRF2BPL and potentially identify additional pathogenetic mechanisms of NEDAMSS and *IRF2BPL*-related disorders, it would be essential to gain a deeper understanding of the functional and dysfunctional roles of the large central portion of the protein, and its potential interplay with the other two flanking domains.

To address these issues, we focused here on the central region of IRF2BPL in search of possible pathophysiological roles. We found that its most proximal polyA/polyQ portion, together with the flanking ZnF which has a triggering role, physiologically drives LLPS. Disease-related mutations, which primarily target this region of the protein, alter IRF2BPL LLPS, leading to the formation of aberrant condensates that sequester wild-type IRF2BPL, depleting it from neuronal cell nuclei and axons. Aberrant LLPS, common to both truncating and missense IRF2BPL mutants, activates downstream disease pathways with WNT1 upregulation and electrophysiological alterations in neuronal cells. These findings identify a role for LLPS in IRF2BPL physiology and establish aberrant phase separation as a biophysical mechanism underlying NEDAMSS and *IRF2BPL*-related disorders.

## RESULTS

### The IRF2BPL central region contains three LCRs and a higher-complexity domain with distinct compositional features and evolutionary dynamics

To better define the possible structural and functional roles of the IRF2BPL central region, we first undertook a comprehensive analysis of its compositional features and predicted structure (**Fig. 1a-d**; **Supplementary Figs. 1 and 2; Supplementary Data 1-3**).

We identified numerous low-complexity segments within the central region of IRF2BPL using the SEG algorithm[29] (available on PlaToLoCo at https://platoloco.aei.polsl.pl), which we grouped into three main LCRs according to their compositional features (**Fig. 1a, Supplementary Fig. 1a; Supplementary Fig. 2a,b; Supplementary Data 2**). The first one ($LCR_1$) is enriched in A/Q, the second one ($LCR_2$) in P/G/S, and the third ($LCR_3$) in both subsets of residues (A/Q and P/G/S). These enrichments in small (A/G), polar (Q/S), and cyclic (P) residues contrast sharply with the overall depletion of hydrophobic and charged amino acids across the three LCRs (**Supplementary Fig. 2a,b**). By contrast, a fourth portion of the central region (highlighted in green in **Fig. 1b,c**) displays a proteome-like composition, similar to that of the ZnF and RING domains, which are also enriched in sulphurated residues (cysteines; **Supplementary Fig. 2a,b**). Together, these findings indicate that IRF2BPL is organized into alternating high- and low-complexity domains with distinct compositional signatures (**Fig. 1c; Supplementary Fig. 2a,b**). This conclusion is further supported by significant differences across these regions in two sequence complexity metrics, i.e., simplicity (SIM) and repetitiveness (REP)[30], calculated in a sliding 20-residue window along the IRF2BPL primary sequence (**Supplementary Fig. 2c**).

From the structural point of view, the AlphaFold IRF2BPL atomic-level structural model (**Fig. 1b**; **Supplementary Data 1**) indicates that the $LCR_1$ polyA and polyQ repeats may form α-helical structures with CC propensity ($CC_1$), consistent with a Paircoil2 prediction (**Fig. 1a, Supplementary Fig. 1b**) and previous experimental studies[12,13], whereas $LCR_2$ and $LCR_3$ are mostly disordered, except

for two short α-helices in $LCR_3$, consistent with their enrichment in structure-breaking P/G residues[12]. The fourth region is predicted to be a structured higher-complexity domain (HCD) of unknown function, with a CC-prone α-helix ($CC_2$) followed by a complex α/β-sandwich (**Fig. 1a,b; Supplementary Fig., 1b**). Other α-helical and β-strand structural elements in IRF2BPL are predicted within the terminal ZnF and RING domains, $LCR_1$ (polyA/polyQ CCs), and $LCR_3$ (**Fig. 1a**, 'α/β' blocks).

These findings overall indicate that the central region is formed by four heterogeneous subregions that may differentially contribute to IRF2BPL function and disease-related dysfunction.

The central portion of IRF2BPL has been generally viewed as a 'variable region', as opposed to the conserved ZnF and RING domains. Based on these findings, we systematically analyzed IRF2BPL evolution across >450 vertebrate species (**Supplementary Fig. 3a; Supplementary Data 4**) to test whether its four subregions, in addition to having distinct compositional and structural features, also display differential evolutionary dynamics (**Fig. 1a,d; Supplementary Fig. 3b-e**). We found that, while the LCRs are quite variable, the HCD is highly conserved, like the ZnF and RING domains. Indeed, the mean entropy at each position of a primary sequence alignment of the IRF2BPL orthologs, a measure of evolutionary variation rate, is significantly higher within the three LCRs and lower within the other domains (ZnF, HCD, RING; **Fig. 1a** (Ent) and **Fig. 1d**). A one-way ANOVA ($F_{(5,1115)} = 64.031$, $p<0.001$) indicated significant differences in the mean entropy (measured across all residues of each domain) between each one of the LCRs and the other three domains ($p<0.001$, Newman-Keuls (NK) post hoc test in all pairwise comparisons).

Furthermore, we found that the overall amino acid composition varied differentially across LCR and non-LCR domains, as shown by the mean coefficient of variation (CV) of amino acid occurrence across nine major clades of different stem age along the vertebrate lineage (**Supplementary Fig. 3a,b; Supplementary Data 2**). Indeed, the mean CV for the 20 amino acids was higher for the LCRs in comparison with the other domains (**Supplementary Fig. 3b**). The LCR compositional dynamics across clades are primarily related to marked changes in the occurrence of a few amino acids, such as A, G, P, Q, and S (**Supplementary Data 2**). Some of these changes significantly correlate with

clade stem ages, with a gradual increase or decrease in amino acid occurrence going from older (Chondrichthyes) to younger clades (Euarchontoglires; **Supplementary Fig. 3c**). Remarkably, these graded changes were mainly related to the length variation of amino acid repeats (AARs) that are found not only within $LCR_1$ (long polyA and polyQ tracts) but also within $LCR_{2/3}$, which bear multiple short polyP and polyG stretches (**Supplementary Fig. 3d,e; Supplementary Data 3**). These observations of regular clade stem age-related trends of LCR/AAR parameters, instead of random clade-specific oscillations, suggest the possibility that selective pressure may have favored gradual compositional change in IRF2BPL along the vertebrate lineage for functional reasons[30,31]. These trends may be related to the emergence of longer polyA/polyQ CCs in $LCR_1$, which promote protein oligo-/polymerization, and of downstream P/G-rich LCRs with opposite effects[12,13] (see below and Discussion).

**Truncating and non-truncating disease-related mutations differentially affect LCRs and other IRF2BPL domains**

Given the considerable differences in their composition and evolutionary dynamics, we investigated whether the LCR and non-LCR portions of IRF2BPL also differ in their predicted sensitivity to mutation and observed frequency of disease-related mutation types, i.e., truncating (TR, i.e. nonsense and frameshift) *vs* non-truncating (nTR; i.e., missense and in-frame indels; **Supplementary Data 5**).

First, we assessed the potential impact of nTR missense mutations across IRF2BPL using MetaDome[32,33] and AlphaMissense[34], which respectively predict functional tolerance to mutation and pathogenicity of amino acid substitutions (**Fig. 1a; Supplementary Fig. 4a,b**). MetaDome identified most IRF2BPL residues as relatively mutation-intolerant, with low tolerance predicted for the ZnF and RING domains, as expected, but also for large portions of the HCD and LCRs (**Fig. 1a**). Consistently, AlphaMissense (**Fig. 1a**, 'αM'; **Supplementary Fig. 4a**) predicted high pathogenicity for mutations in the ZnF, HCD, and RING domains, as well as in multiple LCR regions.

Second, we evaluated the impact of known disease-associated TR and nTR mutations using MutPred2[35] (**Supplementary Data 6**). This analysis predicted that TR mutations disrupt post-

translational modification sites (e.g., SUMOylation, amidation), whereas nTR mutations are more likely to alter local structure (e.g., 'gain of helix' and 'altered coiled coil' for P372R in the HCD) or impair functional sites (e.g., 'altered metal binding' for F30L in the ZnF).

Third, as NEDAMSS and *IRF2BPL*-related disorders can be caused by both TR and nTR mutations, as also found in other neurodevelopmental disorders[36], we analyzed their respective distribution along the LCR and non-LCR regions of IRF2BPL. Remarkably, TR and nTR mutations affect these regions in a significantly different manner (**Fig. 1e**). Indeed, 75% of the truncating mutations target the LCRs whereas 77.4% of the missense mutations impact the non-LCR domains ($p<0.01$, Fisher's exact test). Notably, 50% of disease-related truncating mutation sites fall within $LCR_1$, which represents only 12% of the protein length (**Fig. 1a**). Similarly, 48% of the missense mutations target the HCD, which accounts for 23% of the protein length. These findings are consistent with the notion that structured domains may be sensitive even to single amino acid substitutions caused by missense mutations, while LCRs may be more tolerant in this respect[37,38]. However, evidence exists that missense mutations, such as G195V in IRF2BPL, can also be pathogenic within LCRs[39].

It is also noteworthy that $LCR_1$ is predicted to contain codon repeat-stimulated frameshifting sites (CRFS) (**Supplementary Fig. 4c**) and that several stop codons introduced by $LCR_1$- and $LCR_2$-targeting nonsense mutations may be prone to translational readthrough promoted by small molecules (**Supplementary Fig. 4d**).

Together with the previous findings, these observations indicate that the IRF2BPL central region is formed by four subregions with distinct compositional, evolutionary, and mutational features, whose physiopathological roles are yet to be defined. Furthermore, they identify $LCR_1$ and the HCD as mutational hotspots for disease-related TR and nTR mutations, respectively.

**IRF2BPL undergoes condensation in neural progenitor cell (NPC)-derived human neurons and in other neuronal and non-neuronal cells**

In search of possible functional roles of the IRF2BPL central region, we considered that LCRs and CC domains are known to promote LLPS[16]. Interestingly, ZnF domains can also mediate LLPS[25]. Based on these premises, we hypothesized that the LCRs, CCs, and/or the ZnF domain may mediate IRF2BPL LLPS.

Consistent with this hypothesis, the FuzDrop algorithm[40] identified four major segments predicted to undergo LLPS which overlap the three LCRs, including $CC_1$. Smaller regions with LLPS propensity were predicted at the N-terminal and C-terminal ends of the ZnF domain and within the HCD (**Fig.** 1A).

These findings strongly suggest that IRF2BPL undergoes LLPS, possibly driven by the LCRs and potentially by flanking domains such as ZnF and HCD.

To experimentally test this hypothesis, we first explored whether the protein may form the typical LLPS-driven condensates in the cellular context (**Fig. 1f, Fig. 2**). We observed that both endogenous and exogenous IRF2BPL form discrete, rounded condensates in neuronal and non-neuronal cells, a hallmark of LLPS[16,20], that were detected by confocal microscopy and further characterized using super-resolution imaging.

**Figs. 1f and 2a** show the subcellular distribution of endogenous IRF2BPL in human neural progenitor cell (NPC)-derived neurons, as detected after immunocytochemistry (ICC) using STED super-resolution imaging and conventional confocal microscopy. The protein forms numerous nuclear and extranuclear condensate-like rounded foci of variable size which were also present in the axon and neurites.

The same condensation pattern was observed in differentiated neuronal cell lines (NG108-15)[41,42,43]. **Fig. 2b** and **Supplementary Fig. 5a** show Airyscan super-resolution images of the cell-wide distribution of either endogenous (detected using immunocytochemistry, ICC) or exogenous GFP-tagged IRF2BPL condensates in differentiated neuronal cells. Condensates were detected in the nucleus as well as in axons, neurites, and growth cones.

Non-neuronal human cells (HEK293; **Fig. 2c**; **Supplementary Fig. 5b**) also displayed a similar nuclear and extranuclear pattern of condensation. The IRF2BPL-GFP condensates in these cells can

coalesce with each other into large assemblies in a time-dependent manner, as is typical of LLPS-driven protein complexes[44], especially in the cytoplasm where they have a higher degree of mobility and can coalesce rapidly once in contact (**Fig. 2c,d**; **Supplementary Fig. 5b,c**; **Supplementary Movie 1**).

Taken together, these findings show that IRF2BPL forms intracellular assemblies with typical morphological features and coalescence behavior of LLPS-driven condensates.

**IRF2BPL condensation relies on LLPS**

The morphology of IRF2BPL condensates and their coalescence behavior strongly suggest that LLPS mediates their assembly. To corroborate this conclusion, we used established criteria[22] to assess their LLPS-based nature in cells through ultrastructural, fluorescence dynamics, and biochemical approaches.

First, through correlative light and electron microscopy (CLEM), we found that, at the ultrastructural level, the condensates formed by GFP-tagged IRF2BPL in cells are anisotropic membraneless assemblies, as is typical of LLPS-driven condensates (**Fig. 2e**).

Second, we found that IRF2BPL condensates are largely disrupted after a 15' incubation of cells with 1,6-Hexanediol (1,6-Hex), a compound known to destabilize the weak molecular interactions underlying LLPS, but not with the control compound 2,5-Hexanediol (2,5-Hex), which is known not to disrupt LLPS-driven condensates[45] (**Fig. 2f-j**; **Supplementary Fig. 5d-f**). A one-way ANOVA indicated overall significant differences among the groups ($F_{(2,36)} = 10.54$, $p<0.001$), with a significant reduction of condensation for the 1,6-Hex treatment (condensate area: 0.20 ± 0.05, n = 13 microscopy fields, for the 1,6-Hex group vs. 1.00 ± 0.11, n = 13, for the control group, values normalized to the control group mean, $p<0.01$, NK post hoc test), but not for the 2,5-Hex group (1.15 ± 0.24, n = 13, $p = 0.51$ NK test vs. the control group), in comparison with the control group (**Fig. 2f,g**).

The dissolution of condensates (**Fig. 2h,i**; **Supplementary Fig. 5e**) was already substantial ($F_{(2,20)} = 8.67$, $p<0.01$, one way-ANOVA) even after 5' of 1,6-Hex application (condensate area: 0.32 ±

0.05, n = 9, vs. 1 ± 0.21, n = 9, in the control group, p<0.01, NK post hoc test), and was maintained at 15' (0.13 ± 0.02, n = 6, p<0.01, NK post hoc test vs. control).

The protein from the dissolved condensates diffused to the cytoplasm (**Fig. 2h,j**), with significant differences ($F_{(2,20)}$ = 63.92, p<0.001, one way-ANOVA) already at 5' (Log N/C fluorescence intensity ratio: -0.10 ± 0.01, n = 9, vs. 0.29 ± 0.03, n = 9, in the control group, p<0.001, NK post hoc test), which were maintained at 15' (-0.08 ± 0.01, n = 5, p<0.001, NK post hoc test vs. control).

Similar phenomena were observed for the exogenous GFP-tagged IRF2BPL protein, which was also able to recondense spontaneously after 1,6-Hex washout (**Supplementary Fig**. 5f). Indeed, while the condensate area was significantly reduced after 15' of exposure to 1,6-Hex (one-way ANOVA: $F_{(2, 51)}$ = 159.82, p <0.001; condensate area 0.39 ± 0.02, n = 18, vs. 1.00 ± 0.03, n = 18, in the control group, p<0.001 NK post hoc test), condensates reformed within 1 h after 1,6-Hex washout, covering a similar area as found under control conditions, even with a modest rebound effect (1.14 ± 0.08, n = 18, p<0.001, NK post hoc test vs. control). Such rapid reversibility of protein condensation and decondensation is another hallmark feature of LLPS-driven processes.

Third, we found that IRF2BPL condensates undergo rapid fluorescence recovery after photobleaching (FRAP) (**Fig. 2k**; **Supplementary Fig. 5g,h**), indicative of a highly dynamic exchange of the protein between condensates and the surrounding cellular environment, which is characteristic of phase-separated protein assemblies.

Together with their morphological features and coalescence propensity, these ultrastructural, biochemical, and dynamic properties of the IRF2BPL condensates indicate that they are held together by LLPS.

**The N-terminal region (NTR) encompassing the ZnF and $LCR_1$ drives IRF2BPL LLPS**

Based on the previous findings, we sought to determine whether a specific region of IRF2BPL is mainly responsible for its LLPS. As the three LCRs were the most prominent candidates based on the FuzDrop

prediction (**Fig. 1a**, 'LLPS propensity', pDP score), we used a molecular dissection approach to determine which one of them may be required for condensation (**Fig. 3**).

We initially tested whether LLPS may rely on the polyA/polyQ-bearing $LCR_1$ rather than on the other two P/G/S-rich LCRs. Thus, we expressed in cells either the N-terminal region (NTR; a.a. 1-173) of the protein, containing $LCR_1$ with the upstream ZnF, or the remaining part of it (ΔNTR), containing $LCR_2$/$LCR_3$ with the HCD and RING domains, as GFP fusions (**Fig. 3a**). We found that the NTR alone largely recapitulated the condensation behavior of wild-type (WT) IRF2BPL, albeit with some differences (see below). The NTR fragment underwent significant condensation compared with the GFP control ($F_{(3, 99)} = 45.681$, $p < 0.001$, one-way ANOVA; 0.67 ± 0.05, n = 38 microscopy fields vs. 0.01 ± 0.08, n = 17, values normalized to WT IRF2BPL; $p < 0.001$, NK post hoc test), although the total area occupied by condensates was reduced relative to full-length WT IRF2BPL (1.00 ± 0.06, n = 29; $p < 0.01$, NK post hoc test). In contrast, the ΔNTR fragment did not form condensates (0.01 ± 0.08, n = 19; $p = 0.98$, NK post hoc test vs. GFP control), displaying a diffuse GFP-like distribution confined to the nucleus, consistent with the presence of the NLS in it. These results indicate that the N-terminal region of IRF2BPL, which contains $LCR_1$ but not $LCR_2$ or $LCR_3$, serves as the primary physiological driver of LLPS.



**The ZnF and $LCR_1$ together shape the LLPS behavior of the IRF2BPL NTR**

Besides $LCR_1$ (a.a. 76-173), the NTR contains the ZnF with short flanking peptides (ZnF domain, a.a. 1-75) which also displayed LLPS propensity in the FuzDrop analyses (**Fig. 1a**). Thus, we proceeded further with the molecular dissection approach to determine the relative contribution to LLPS of the two constituent parts of the NTR, which were alternatively expressed in cells as GFP fusions (**Fig. 3b**).

While the ZnF fragment formed condensates in cells, although in a limited number, the $LCR_1$ exhibited a completely diffuse distribution. Indeed, cultures expressing the ZnF alone displayed significantly more condensates than the negative control cultures (GFP; $F_{(3, 37)} = 856.24$, $p<0.001$, one-way ANOVA; condensate area: 0.11 ± 0.01, n = 13 microscopy fields, vs. 0.005 ± 0.002, n = 5,

respectively, values normalized to the NTR group mean, $p<0.001$, NK post hoc test), while this was not the case for cultures expressing $LCR_1$ alone ($0.02 \pm 0.01$, $n = 13$, $p = 0.66$, NK post hoc test). The condensation ability of the ZnF alone was, however, considerably reduced in comparison with that of the whole NTR ($1.00 \pm 0.06$, $n = 10$, $p<0.001$, NK post hoc test). This indicates that the ZnF and $LCR_1$ act synergistically to drive NTR-mediated LLPS, with the ZnF possessing an intrinsic ability to undergo phase separation and playing a triggering role in the LLPS of the NTR region and of the whole IRF2BPL protein.

These conclusions were confirmed when we expressed in cells IRF2BPL mutants lacking either the ZnF or $LCR_1$ (**Fig. 3c**). The ZnF deletion (ΔZnF) was sufficient to essentially abolish IRF2BPL condensation, consistent with a key role of this domain in driving LLPS (one-way ANOVA: $F_{(2,34)}=126.41$, $p<0.00001$; condensate area: $0.06 \pm 0.02$, $n = 14$ microscopy fields, vs. $1.00 \pm 0.02$, $n = 10$, values normalized to the WT IRF2BPL group, $p<0.001$, NK post hoc test).

Conversely, the internal $LCR_1$ deletion ($\Delta LCR_1$) only reduced, in comparison with WT IRF2BPL ($0.81 \pm 0.02$, $n = 13$, $p<0.01$, NK post hoc test), but did not abolish condensation ($p<0.001$ vs. ΔZnF, NK post hoc test), thus suggesting that other portions of the protein may functionally replace $LCR_1$ in promoting LLPS together with the ZnF.

Taken together, these findings indicate that the ZnF is essential for the LLPS of IRF2BPL and its NTR. However, since the ZnF alone exhibits limited condensation capacity, the presence of $LCR_1$ appears necessary to enable substantial condensation of the NTR region.

**The IRF2BPL NTR can rapidly undergo photo-triggered LLPS using the optoDroplet system**

Another hallmark of LLPS-prone peptides is their ability to undergo phase separation with relatively rapid kinetics. To assess whether the LLPS-driving region of IRF2BPL exhibits this dynamic behavior, we employed the optoDroplet system[46], which enables rapid, temporally controlled, and reversible induction of condensation of LLPS-prone peptides over tens of seconds (**Fig. 3d,g**).

To perform these experiments, the NTR was expressed as a fusion protein with mCherry (mCh) and an *A. thaliana* cryptochrome 2 protein (Cry2) fragment (**Fig. 3d**). The NTR-mCh-Cry2 fusion protein displayed, as expected, some spontaneous condensation after 24 h of expression. Moreover, photoactivation (488 nm) rapidly triggered massive condensation (**Fig. 3e,f**; $F_{(3,77)} = 20.804$, $p<0.001$, two-way ANOVA, construct x treatment interaction; condensation area: 2.28 ± 0.44 %, n = 12, vs. 6.22 ± 0.47 %, n = 12, for non-photoactivated and photoactivated cultures, respectively, $p<0.001$ NK post hoc test). Neither spontaneous nor photoactivated condensation were observed, as expected, for the control construct mCh-Cry2. Thus, both spontaneous and triggered condensation of the NTR were significantly increased in comparison with the mCh-Cry2 control (spontaneous condensation area: 2.28 ± 0.44 %, n = 12 microscopy fields, for the NTR vs. 0.08 ± 0.04 %, n = 7, for the control; triggered condensation 6.22 ± 0.47 %, n = 12, for the NTR, and 0.08 ± 0.07 %, n = 6, for the control; $p<0.001$, NK post hoc test for both comparisons).

We used the same approach to further characterize the relative ability of the ZnF and the $LCR_1$ to undergo LLPS (**Fig. 3e,f**). Similar to what observed in the previous experiments, the ZnF displayed some spontaneous condensation (condensate area: 1.09 ± 0.09 %, n = 12, *vs*, 0.08 ± 0.04 %, n = 7, for the control; $p<0.05$, NK post hoc test), which was moderately enhanced by photoactivation (~25% of that of the NTR construct). Conversely, $LCR_1$ did not display any significant spontaneous condensation (0.07 ± 0.03 %, n = 12, *vs*, 0.08 ± 0.04 %, n = 7, for the control; $p = 0.99$, NK post hoc test) nor a significant increase in condensation induced by photoactivation ($p = 0.73$, NK post hoc test).

These findings show that the NTR can rapidly undergo triggered LLPS and further confirm that both the ZnF and $LCR_1$ are required to determine its overall condensation behavior.

**The ZnF domain of IRF2BPL undergoes LLPS also in vitro**

To further characterize the ability of IRF2BPL and its ZnF domain to undergo LLPS, as found in the cellular context, we sought to study their intrinsic phase separation behavior also in a purified form in vitro, independently of other cellular cofactors. We produced WT IRF2BPL in *E. coli* but the protein

was not soluble, consistent with the known in vitro aggregation propensity of LLPS-prone proteins bearing long polyQ and polyA tracts[47,48,49] (see Methods and **Supplementary Fig. 6**).

Thus, we chemically synthesized the ZnF domain (a.a. 1-75), the key LLPS-promoting region of the protein, conjugated to the fluorophore FITC for visualization during in vitro LLPS assays[50]. This peptide, which has a relatively low condensation propensity in cells, was also considerably more soluble than the WT protein and allowed us to perform in vitro analyses of condensation. We found that the fluorescently-labelled peptide can form condensates also in vitro, even in a physiological phosphate saline buffer (pH 7.4) devoid of crowding agents (such as PEG-8000), as shown by simultaneous differential interference contrast (DIC) and fluorescence imaging (**Fig. 3g**).

These findings indicate that the ZnF, i.e., the key LLPS-promoting region of IRF2BPL, can autonomously undergo condensation in vitro, at physiological pH and ionic strength, mimicking its behavior in the cellular context.



**Disease-related truncated IRF2BPL mutants form aberrant condensates with altered size, shape, and subcellular localization**

Most of the known disease-related mutations lead to the production of truncated forms of the protein[1,8]. A plot of the local frequency of the truncation sites along IRF2BPL (**Fig. 4a**) indicates that they are mostly clustered within or near $LCR_1$ and its junction with $LCR_2$ (cluster 1; a.a. 116-221). These mutations lead to the production of N-terminal IRF2BPL protein fragments bearing the ZnF and variable portions of $LCR_1$ overlapping, entirely or partially, the LLPS-driving region that we identified (i.e., a.a. 1-173). Interestingly, disease-causing truncating mutations also target sites clustered at the junction of $LCR_2$ with the HCD (cluster 2) and of $LCR_3$ with the RING domain (cluster 3).

Thus, we tested whether the different disease-related truncating mutations may alter the IRF2BPL LLPS behavior. Towards this aim, we generated constructs for the cellular expression of representative mutant forms of the protein (**Fig. 4a**, upper panel). Given the quantitative prominence of cluster 1 mutations, we included in this analysis three of its disease-related mutant forms of the protein,

i.e., Y173X[8], which bears an intact $LCR_1$ with its entire polyA and polyQ repeats, Q123X[26], with a partial polyQ truncation, and Q116X (reported in the ClinVar database (variation ID: 1801711) as likely pathogenic) with an even shorter polyQ tract. We also generated two mutants reproducing cluster 2 truncations between $LCR_2$ and HCD (R331X) and cluster 3 truncations between $LCR_3$ and the RING domain (A708X).

When we expressed GFP-tagged cluster 1 mutant fragments for 24 h in cells (**Fig. 4b-d**; **Supplementary Fig. 7a,b**), we found that all of them still formed condensates. However, these condensates displayed length-dependent alterations in their subcellular localization, morphology, ultrastructure, and biochemical properties, which were more pronounced with the progressive shortening of $LCR_1$.

As observed previously, WT IRF2BPL forms both nuclear and cytoplasmic condensates (**Fig. 4b**). The former are numerous and small, whereas the latter have more variable dimensions and tend to coalesce into larger ones with time. To characterize quantitatively the condensation patterns of WT and mutant IRF2BPL, we measured five parameters which all varied significantly between WT and mutant IRF2BPL, i.e. the total cell area occupied by condensates (**Fig. 4c**, first panel from the left; $F_{(3, 58)} = 15.892$, $p<0.001$, one-way ANOVA), the number of condensates per cell unit area (**Fig. 4c**, second panel; $F_{(3, 58)} = 19.674$, $p<0.001$, one-way ANOVA), the mean size of individual condensates (**Fig. 4c**, third panel; $F_{(3, 45523)} = 266.29$, $p<0.001$, one-way ANOVA), as well as the proportion of condensates in the nucleus (**Fig. 4c**, fourth panel; $F_{(3, 20)} = 367.84$, $p<0.001$, one-way ANOVA) and the proportion of condensate area in the nucleus (**Supplementary Fig. 7b**; $F_{(3, 20)} = 957.35$, $p<0.001$, one-way ANOVA).

The Y173X mutant had a pattern of condensation relatively closer to WT IRF2BPL, although with a strong reduction in the number of small nuclear condensates and an increase in that of larger cytoplasmic ones. The shorter truncation mutants (Q123X, Q116X) prevalently formed a few, mostly cytoplasmic large condensates, which were seldom found in the nucleus (**Fig. 4b,c; Supplementary Fig. 7b**). Smaller intranuclear condensates formed by these mutants were only transiently observed in the very first phases of their expression. The progressive coalescence of these mutant condensates in

the cytoplasm led to a further reduction in condensate number and an increase in condensate size after 24 h of expression (see below and **Supplementary Movies 3-4**).

The progressive shortening of $LCR_1$ in the mutants, from Y173X to Q116X, resulted in the gradual loss of the small nuclear condensates, which were replaced by fewer and larger cytoplasmic ones. Quantitatively, this corresponded to a decrease in both the total area occupied by condensates and their number, with an increase in their average size compared with WT IRF2BPL ($p<0.01$ in all instances, NK post hoc test; **Fig. 4c**). Notably, the longer mutant (Y173X) differed from the two shorter ones in all these respects ($p<0.05$ in all instances, NK post hoc test), indicating a length-dependency for the condensation phenotypes of cluster 1 mutations. Moreover, with the shortening of the expressed N-terminal fragment, we also noticed that condensates, which are typically rounded for WT IRF2BPL, became progressively more eccentric and elongated (**Fig. 4b,d**). This latter shape distortion may be related to the apparent reduction in the average size of the condensates formed by Q116X in comparison with those of Q123X. The condensates of both mutants were anyway larger than those of the longer mutant Y173X.

The mutant forms of IRF2BPL also displayed an altered subcellular distribution in comparison with the WT protein. Indeed, all three mutants showed a dramatic reduction in the proportion of their condensate area occupying the nuclear compartment (**Supplementary Fig. 7b**), with a decrease in the number of intranuclear condensates (**Fig. 4c**; $p<0.001$ vs. WT IRF2BPL in all instances, NK post hoc test).

These findings indicate that the truncation mutants undergo qualitatively and quantitatively aberrant LLPS, with altered condensate dynamics and distribution in comparison with WT IRF2BPL. These changes appear to be related to the length of the N-terminal fragments.

The differential LLPS behavior of Y173X and Q116X, together with the lower condensation tendency of the ZnF (a.a. 1-75; **Fig. 3e**), suggest the hypothesis that the 76-115 region, which overlaps with $CC_1$ and contains the first polyA and part of the polyQ, may represent the strongest LLPS-promoting $LCR_1$ subregion. Together with our previous observations on the relative stability of polyA versus polyQ CCs[13,51], these findings suggest that the progressive loss of the region downstream $CC_1$,

and even of its less stable polyQ part, may stabilize CC formation by the remaining polyA tract, which may in turn enhance LLPS[20] (see Discussion). Moreover, other downstream parts of the protein may have important regulatory roles on the physiological LLPS of IRF2BPL, which are lost when they are elided by truncating mutations.

The two more downstream truncations R331X and A708X also led to substantial aberrations in condensation patterns (**Supplementary Fig. 7c-f**). Indeed, while the total cell area occupied by condensates of the two mutants was not different from that occupied by WT IRF2BPL condensates after 24 h of expression ($F_{(2,66)} = 0.14$, $p = 0.86$, one-way ANOVA; **Supplementary Fig. 7e**, first graph from the left), the number, size, and subcellular distribution (proportion of condensate area and number in the nucleus) of the condensates were significantly altered ($F_{(2,66)} = 17.41$, $p<0.001$, $F_{(2,64002)} = 236.47$, $p<0.001$, $F_{(2,15)} = 44.61$, $p<0.001$, $F_{(2,15)} = 28.88$, $p<0.001$, respectively; **Supplementary Fig. 7d,e**). Indeed, as found for cluster 1 mutants, both R331X and A708X formed fewer (**Supplementary Fig. 7e**, second graph) and larger (**Supplementary Fig. 7e**, third graph) condensates in comparison with the WT protein ($p<0.001$ in all instances, NK post hoc test). These condensates were also mislocalized, i.e., mostly cytoplasmic for R331X and quite evenly split between the cytoplasm and the nucleus for A708X ($p<0.02$ vs. WT IRF2BPL in all instances, NK post hoc test, for both localization-related parameters; **Supplementary Fig. 7d and 7e**, fourth graph). For these two mutants, we did not observe the marked changes in condensate eccentricity as found for the cluster 1 mutants (**Supplementary Fig. 7f**).

Together with the previous findings, these observations indicate that truncations along the protein up to the end of $LCR_3$ can significantly alter the LLPS behavior of IRF2BPL in both quantitative and qualitative terms in the cellular context. For technical reasons (see above), we could not study the impact of the same mutations on the ZnF-dependent LLPS of IRF2BPL in vitro.

**The aberrant condensates formed by disease-related IRF2BPL truncated mutants display accelerated coalescence kinetics, reduced sensitivity to 1,6-Hex, and blunted FRAP in the cellular context**

Building on these findings, we investigated whether condensates formed by truncated IRF2BPL mutants may exhibit, in addition to aberrant morphology and localization, an altered dynamic behavior. Thus, we studied their coalescence kinetics, sensitivity to 1,6-Hex-mediated disruption, and molecular exchange with the surrounding cellular environment as assessed by FRAP.

First, we compared the coalescence dynamics of condensates formed by two disease-related mutants (Q123X and Y173X) with those of WT IRF2BPL condensates. Towards this aim, we performed live-cell fluorescence imaging experiments, starting 24 h after transfection over a time course of 12 hours, in cell expressing GFP-tagged IRF2BPL-Q123X, IRF2BPL-Y173X, and WT IRF2BPL (**Fig. 5a,b and Supplementary Movies 2-4**). These experiments revealed that the condensates formed by the Q123X and Y173X mutants have a significantly stronger tendency to coalesce than those formed by WT IRF2BPL. Indeed, at the beginning of the time course experiment, the cells expressing the two mutants already had a significantly lower number of condensates than cells expressing WT IRF2BPL (**Fig. 5b**, left two panels; $F_{(2,380)} = 29.30$, $p<0.001$, two-way ANOVA, with significant differences between the mean number of condensates for WT IRF2BPL vs. the two mutants, both at 24 and 36 h, $p<0.01$ in all instances, NK post hoc test). Furthermore, the mutant condensates coalesced more rapidly than the wild-type ones during the subsequent 12 hours (**Fig. 5b**, right two panels). At 36 h after transfection, only a few (1-4) large cytoplasmic condensates were usually found in cells expressing IRF2BPL-Q123X and -Y173X, whereas WT IRF2BPL still displayed numerous nuclear small condensates. The reduction in condensate number was significantly stronger for the two mutants in comparison with the WT protein (one-way ANOVA: $F_{(2,190)}=4.7065$, $p<0.05$, NK test, in both instances).

Second, we analyzed the differential effect of 1,6-Hex on the condensation of WT and mutant IRF2BPL (Y173X, Q123X, and Q116X; **Fig. 5c,d**). These experiments revealed that sensitivity to 1,6-Hex of Y173X mutant condensates is significantly reduced in comparison to that of WT condensates in terms of total cellular area occupied by condensates. Indeed, a two-way ANOVA indicated a differential effect of 1,6-Hex across constructs for total condensate area, condensate number per cell unit area, and individual condensate size ($F_{(3,34)} = 20.47$, $F_{(3,34)} = 24.17$, $F_{(3,34)} = 10.55$, respectively, $p<0.001$ in all

instances, one-way ANOVA for normalized condensate area in 1,6-Hex-treated cultures). This area was not significantly reduced for the two shortest mutants (Q123X and Q116X). However, 1,6-Hex was still able to induce fragmentation of their large condensates into more numerous, smaller ones ($p<0.001$ for both mutants vs. WT, NK post hoc test), indicating that even the aberrant condensates formed by these mutants retain some sensitivity to the action of the diol.

Third, we compared the FRAP of WT and mutant (Y173X, Q123X) IRF2BPL condensates (**Fig. 5e,f**). We found that the condensates formed by both mutants display substantially reduced FRAP in comparison with those formed by WT IRF2BPL. Indeed, while the mobile fraction ($M_f$) of WT IRF2BPL condensates was 83%, that of the Y173X and Q123X mutants was reduced to 52% and 45%, respectively. These results indicate that the mutant protomers are exchanged at a lower rate between the condensates and the surrounding cellular environment in comparison with WT ones.

The concurrent alterations in condensate coalescence, sensitivity to 1,6-Hex, and FRAP indicate that the dynamics of IRF2BPL phase separation are significantly altered by disease-related mutations, with the formation of aberrant condensates that are significantly more stable than those formed by the WT protein.



**The aberrant condensates formed by disease-related IRF2BPL truncated mutants display altered ultrastructure and a length-dependent reduction in degradation rate**

Based on the previous findings, we tested whether the altered dynamics of condensates may be related to ultrastructural changes in the molecular organization of the mutant IRF2BPL protomers.

Towards this aim, we performed CLEM of cells expressing either WT IRF2BPL or its Y173X and Q123X mutants in search of possible differences in the internal ultrastructural organization of their condensates (**Fig**. 6; **Supplementary Fig. 8**). Inverse fast Fourier transform (iFFT) analysis of CLEM images revealed that whereas the rounded FL condensates are isotropic and amorphous, the elongated Y173X and Q123X condensates exhibit a regular internal organization of evenly spaced fibrillary profiles, generally aligned with the major condensate axis. The mean spacing ($\bar{s}$) between fibrils, as

calculated using FFT analyses, was ~16.5 nm for both mutants, which was conserved in different areas of one same condensate and across condensates (**Supplementary Fig. 8a,b**). The mean spacing was not significantly different across mutants (16.45 ± 0.19 nm, n = 9 fields, vs. 16.82 ± 0.33 nm, n = 6 fields, for Q123X and Y173X, respectively; p = 0.31, t-test; **Supplementary Fig. 8c**).

These findings suggested that the packing of mutant proteins into larger, more stable condensates, with a more ordered internal organization, may reduce their molecular turnover, thus increasing their half-life and persistence in cells. To test this hypothesis, we compared the relative degradation rate of GFP-tagged WT and mutant forms of IRF2BPL, and of GFP alone, through a cycloheximide (CHX) chase assay (**Fig. 7a**). We found that the degradation rate was lower for all IRF2BPL proteins in comparison with GFP alone (**Fig. 7b**; $F_{(3,8)} = 9.25$, $p<0.01$, one-way ANOVA, n = 3 per group; $p<0.05$ in all comparisons, NK test) and for the truncated mutants in comparison with the WT protein (**Fig. 7c**; $F_{(2,6)} = 7.14$, $p<0.03$, n = 3 per group, values normalized to the WT group; $p<0.05$ in both cases, NK test). The lower degradation rate of the truncated mutants may be the combined result of the absence of the three PEST sequences and the reduced accessibility of their protomers within large condensates.

These findings show that the truncated IRF2BPL mutants display a length-dependent alteration in stability and persistence in cells in comparison with the WT protein.



**The aberrant cytoplasmic condensates of disease-related truncated mutants recruit WT IRF2BPL, subtracting it from nuclear and axonal compartments, in a length-dependent manner**

PolyA/polyQ CC domains, such as those in $LCR_1$, can mediate physiological protein homo-/hetero-typic interactions[12,13,52]. In certain diseases, polyA/polyQ proteins can entrap in their condensates and aggregates other cellular proteins bearing similar repeats, causing their loss of function[12,13,14,15,52]. Sinha Ray et al. (2022)[27] reported how truncated IRF2BPL mutants sequester the normal WT IRF2BPL protein in their cytoplasmic condensates. Based on this evidence and our previous findings of length-

dependent condensate alterations, we tested whether the recruitment of WT IRF2BPL into the aberrant condensates of its mutant forms may also be dependent on the length of their variably truncated $LCR_1$.

We found that all three truncated mutants, when exogenously expressed in NPC-derived human neurons, as well as in neuronal and non-neuronal cell lines, were able to sequester WT IRF2PBL (**Fig. 7d-h**). Indeed, the WT and mutant proteins appeared to co-condensate mostly in the cytoplasm. WT IRF2BPL was detected through ICC using an antibody against an epitope that is not present in the truncated mutant proteins. Such sequestration visibly subtracted WT IRF2BPL from the nuclear and axonal compartments of neuronal cells (**Fig. 7d,g,h**).

To quantify this phenomenon, we measured the mean number of condensates in nuclei of transfected cells ($n_t$) expressing a given IRF2BPL mutant and compared it with that found in control nuclei ($n_c$) of neighboring, non-transfected cells in the same culture (**Fig. 7d**). All mutants displayed a significantly reduced number of IRF2BPL condensates in the nuclei of transfected cells ($n_t$) in comparison with the nuclei of neighboring non-transfected control cells ($n_c$; **Fig. 7e**; $F_{(3,3335)} = 15{,}68$, $p<0.001$, two-way ANOVA, construct x transfection interaction; $p<0.001$ for all mutants, but not for GFP, NK test). Such recruitment was not observed in control cells expressing GFP alone (**Supplementary Fig. 9a**). Moreover, this reduction was more intense for the shortest mutants (**Fig. 7f**; $p<0.02$ for Q116X *vs* Y173X, NK post hoc test).

These results show that the length-dependent alteration of the molecular properties of the aberrant condensates varies their ability to recruit WT IRF2BPL. This can cause a length-dependent gain-of-function of cluster 1 truncated mutants contributing to the molecular pathogenesis of *IRF2BPL*-related disorders by sequestering the WT copies of the protein produced from the normal allele. Ultimately, these observations identify aberrant LLPS as the biophysical basis of a key mechanism in the cellular pathogenesis of NEDAMSS, i.e., the sequestration of WT IRF2BPL outside of the nucleus within the aberrant cytoplasmic condensates of its mutant forms.

**Mutant forms of IRF2BPL undergoing aberrant LLPS in cells and sequestering the WT protein trigger molecular disease mechanisms**

The sequestration of WT IRF2BPL out of the nucleus by aberrant condensates may in turn trigger downstream cellular and molecular mechanisms, ultimately leading to neural damage and disease. To identify some of these mechanisms, we performed genetic and electrophysiological experiments in the cellular context (**Fig. 8**). These experiments aimed at testing whether the expression of NEDAMSS-related IRF2BPL mutants undergoing aberrant LLPS may trigger, by sequestering WT IRF2BPL, cellular and molecular mechanisms of the disease.

To this end, we first used CRISPR/Cas9-mediated gene editing of the IRF2BPL locus in human HEK293 cells to generate cell lines carrying a monoallelic truncating mutation from cluster 1 (Q127X; **Fig. 8a,b**), thereby mimicking the genetic makeup observed at the IRF2BPL locus in NEDAMSS, or a knockout (KO) of the *IRF2BPL* gene (**Fig. 8d,e**). Isogenic cell lines, for both Q127X and KO HEK293 cells, were used as controls. The Q127X mutation was chosen for the presence of a favorable CRISPR/Cas9 editing site at codon 127 of the IRF2BPL ORF.

Like other cluster 1 mutations (e.g., Q123X), Q127X caused marked depletion of WT IRF2BPL from the nucleus. Indeed, immunostaining with available antibodies targeting a C-terminal epitope of IRF2BPL, thus not recognizing its N-terminal fragments, showed that the full-length WT protein produced by the normal allele largely mislocalized to extranuclear condensates (**Fig. 8b**). This sequestration was sufficient to induce a significant upregulation of WNT1 in these cells, as shown by qPCR experiments (**Fig. 8c**; $p<0.05$, t-test), a recently identified key molecular mechanism of the disease[8]. Remarkably, a similar degree of WNT1 upregulation was found in KO cells (**Fig. 8f**; $p<0.05$, t-test).

These results show how the cytoplasmic sequestration of functional WT IRF2BPL, mediated by the aberrant LLPS of its mutant forms, can trigger dominant-negative disease mechanisms (WNT1 upregulation) in a manner that is quantitatively comparable to the complete knockout of the WT protein.

**Mutant forms of IRF2BPL undergoing aberrant LLPS produce electrophysiological alterations in neuronal cell resting potential and excitability**

Based on the previous findings, we also tested whether the extranuclear sequestration of WT IRF2BPL into aberrant condensates, and the consequent dysregulation of disease-related molecular pathways such as WNT1, ultimately leads to functional impairments in neuronal cells. As IRF2BPL has been implicated in neuronal function and maintenance[8], we searched for possible electrophysiological phenotypes in differentiated neuronal cells (NG108-15) in which we alternatively expressed WT and mutant forms of IRF2BPL (**Fig. 8g,k**). First, we found that the expression for 48 h of IRF2BPL mutants (i.e., IRF2BPL-Q173X-GFP and -Q123X-GFP; **Fig. 8g**) significantly reduced the cell resting potential in comparison to control cells expressing either WT IRF2BPL or GFP ($F_{(3,86)} = 4.54$, $p<0.01$ one-way ANOVA; $p<0.05$ NK post hoc test; **Fig.** 8h). Input resistance was not different across groups (**Supplementary Fig. 9b,c**). Second, we found that the same two mutants significantly impaired spiking ability, as determined by measuring spike peak amplitude (both from resting potential, as in **Fig. 8i** ($F_{(3,86)} = 3.75$, $p<0.02$, one-way ANOVA), and from the steady-state (sp) potential in the plateau phase following depolarization, as in **Fig. 8j**; $F_{(3,85)} = 4.53$, $p<0.01$, one-way ANOVA), in comparison to either WT IRF2BPL or GFP. Indeed, the proportion of cells with a spike peak amplitude ≥70 mV was significantly lower for both the Y173X and Q123X mutant experimental groups ($p<0.02$ in both instances, Fisher's exact (FE) test; **Fig. 8k**). Threshold currents did not vary across experimental groups (**Supplementary Fig. 9d**).

Taken together, these findings indicate that even the expression of mutant forms of IRF2BPL undergoing aberrant LLPS can significantly alter neuronal excitability and induce neuronal damage, as signaled in particular by the reduction of the resting potential[53], in agreement with evidence of other phenotypes indicating neuronal damage in NEDAMSS[1,8].

Together with the results of the gene-editing experiments, these findings indicate that the aberrant LLPS of IRF2BPL NEDAMSS-related mutants can trigger cellular and molecular mechanisms of the disease.

**Disease-related missense mutation sites target the HCD in the IRF2BPL central region with predicted structural homology to methyl-DNA-binding domains**

The previous findings prompted us to test whether the non-truncating disease-related missense mutations may also lead to aberrant LLPS, as found for the truncating ones (**Figs. 9** and **10**). Our initial bioinformatics screening (**Fig. 1a**) had shown how, differently from truncating mutation sites that cluster within $LCR_1$, most disease-related missense mutation sites cluster in gene region coding for the middle portion of the IRF2BPL central region (cluster 4; see below and **Figs. 1a** and **10a,b**). Notably, this region appeared to overlap with the HCD, which was identified in the AF structural model, consisting of a long CC-prone helix ($CC_2$) followed by an α/β-sandwich (**Fig. 9a**, upper left panel) comprised between the mostly disordered $LCR_2$ and $LCR_3$. Thus, before studying the impact on LLPS of missense mutations targeting this region, we decided to characterize it in further detail, also because of its possible relevance to IRF2BPL physiological function.

Remarkably, this region of the protein did not display primary sequence similarity to other proteins in a BLAST search, except for IRF2BP1, an IRF2BPL paralog containing a similar domain (**Supplementary Figs. 9e,f**, **10**, and **11a**). However, a DALI structural similarity search in the PDB, revealed instead striking resemblance of the HCD α/β-sandwich to that of known DNA-binding domains. Using this α/β-sandwich region as the query in a more refined DALI search, the first hit was the methyl-DNA binding protein MPND bound to DNA (PDB: 7ydw). In matching the two structures, we found that a DNA-facing MPND region, formed by a β-hairpin, an α-helix, and an interposed loop, is also present within the core of the IRF2BPL HCD, possibly representing a DNA-binding site. Furthermore, when we used this core region (**Fig. 9a** upper middle panel) as the query, we found numerous hits represented by methyl-DNA-binding proteins (**Supplementary Data 7**). Other hits included RNA-binding proteins.

A structural matching of the α/β-sandwich core to human methyl-DNA-binding domains (MBDs) belonging to MBD-2, -4, -5, and MeCP2, whose structures are available in the PDB, revealed a striking similarity (**Fig.** 9A, upper right panel). Furthermore, while the primary sequence similarity

between the IRF2BPL HCD core region and the other human MBDs is low, many key residues forming the MBD hydrophobic core or involved in DNA binding are conserved or similar (**Fig. 9b**). **Fig. 9a** (lower left and middle panels) highlights how the HCD core overlaps very well with the resolved structure of a MeCP2 MBD fragment bound to DNA.

In the sequence alignment of the HCD core and the MBDs (**Fig. 9b**), several residues that bind DNA in human MBDs are present, or are conservatively replaced, in IRF2BPL (V408, K410, K418, S433, K439), forming a 'belt' potentially oriented towards DNA (**Fig. 9a**, lower middle and right panels). Three positively charged Ks are strategically located at the ends (K410, K439) or at the center (K418) of this potential DNA binding interface (**Fig. 9a**, lower right panel). The functional relevance of the HCD is also supported by its high degree of phylogenetic conservation (**Supplementary Fig. 11b**), and by the fact that, within its core portion, the residues that are part of the predicted α-helical and β-strand structures are even more conserved than those in the random coil loops separating them (**Supplementary Fig. 11c,d**). These loops appear to be more tolerant to amino acid substitutions or even small deletions (e.g., four-residue gap in MBD1/2 in the alignment in **Fig 9b**).

Remarkably, four of the eight disease-related IRF2BPL missense mutations cause substitutions in the α/β-sandwich (P372R, R391C, K418N) or in the upstream $CC_2$ (P332A; **Fig. 10a**). K418N elides a charged residue in the β-hairpin that may have a key DNA-binding role. The actual ability of the α/β-sandwich domain to bind methylated vs. non-methylated DNA will have to be experimentally determined once potential consensus sequences are identified.

The presence of the $CC_2$ region upstream of the α/β-sandwich is reminiscent of a domain organization found in transcription factors undergoing CC-mediated dimerization upon DNA binding, such as FOXP2[54]. This suggests that $CC_2$ and the α/β-sandwich may constitute a functional unit, the HCD, that may represent a putative DNA-binding domain of IRF2BPL.

**A circular dichroism structural analysis confirms the presence of a β-hairpin structure in the HCD core**

To test experimentally whether the presumed DNA-binding β-hairpin in the HCD is folded as in the AF model, and not disordered as the flanking LCRs, we chemically synthesized a HCD core peptide (a.a. 395-448; **Fig. 9b**) and studied its secondary structure by circular dichroism (CD) in physiological saline buffer (**Fig. 9c,d**). In the AF model, this peptide contains 35% of β-strand with a 5-residue turn (9%), 22% of α-helix, and 33% of random coil. The deconvolution analysis of the peptide CD spectra revealed four structural components in proportions (33% β-strand, 13% turns, 15% α-helix, and 39% random coil) that are in excellent agreement with the AlphaFold model (**Fig. 9c,d**, upper panels). These components were quite stable upon heating from 5 to 75 °C (**Fig. 9c**, lower panel). In particular, the β-strand structure, i.e., the major structural component that was observed, was only marginally destabilized (-15%) by heating the peptide solution up to 75 °C, with a modest increase in the random coil component (+14.5%; **Fig. 9d**, lower panel).

Taken together, these findings indicate that the HCD contains in its core region folded structural elements, including a prominent, thermally stable β-strand component, in good quantitative agreement with the AlphaFold structural model of IRF2BPL.



**Disease-related missense mutations alter IRF2BPL LLPS behavior and subcellular localization**

Finally, we tested the potential impact of missense mutations on the LLPS behavior of IRF2BPL. We selected for this analysis two mutations (R372R and K418N) targeting the HCD, which represents the protein domain with the highest occurrence of missense mutations (cluster 4; **Fig. 1a** and **Fig. 10a,b**). K418N falls in the middle portion of the β-hairpin homologous to the DNA-binding interface in MBDs. We also studied a third mutation (F30L) targeting the N-terminal ZnF domain, which is known to be involved in DNA binding[27].

To test whether also these mutations may impact LLPS, like the truncating ones, we compared the condensation patterns of the three mutants, expressed in cells as GFP fusions for 24 h, with those of WT IRF2BPL (**Fig. 10c,d**; **Supplementary Fig. 12a**). As previously done for the truncating mutations, to quantitatively characterize the condensation patterns of WT and mutant IRF2BPL, we measured five

parameters, four of which varied significantly between WT and mutant IRF2BPL, i.e., the total cell area occupied by condensates, the only parameter that did not vary (**Fig. 10d**, first graph from the left; $F_{(3,39)}$ = 0.18, p=0.9, one-way ANOVA), the number of condensates per cell unit area (**Fig. 10d**, second graph; $F_{(3,39)}$ = 4.03, p<0.02, one-way ANOVA), the mean size of individual condensates (**Fig. 10d**, third graph; $F_{(3,51153)}$ = 87.960, p<0.001, one-way ANOVA), as well as the proportion of condensate area in the nucleus (**Fig. 10d**, fourth graph; $F_{(3,18)}$ = 29.01, p<0.001, one-way ANOVA), and the proportion of condensates in the nucleus (**Supplementary Fig. 12b**; $F_{(3,18)}$ = 10.53, p<0.001, one-way ANOVA).

Indeed, when these proteins were expressed in cells for 24 h (**Fig. 10c**, **Supplementary Fig. 12a**), the overall cellular area occupied by condensates was not significantly different for the three mutants and the WT protein. However, they displayed markedly altered patterns of condensation in terms of condensate number, size, and subcellular localization (**Fig. 10d**, **Supplementary Fig. 12b**). While the F30L and P372R mutants showed an increased mean condensate size (p<0.01, NK post hoc test for both mutants vs. WT), and a trend towards a reduction in condensate number, the K418N mutant displayed the most evident alterations of the condensation pattern. Indeed, this mutant showed a substantial loss of small intranuclear condensates, as well as a considerable reduction in condensate number (p<0.02, NK post hoc test vs. WT) with a notable increase in condensate size (p<0.001, NK post hoc test vs. WT). The most dramatic changes for these three mutants were observed in terms of subcellular localization. Indeed, all three mutants displayed marked mislocalization, with most of their condensates found in the cytoplasmic compartment (p<0.03, NK post hoc test for all mutants vs. WT), perhaps owing to an altered DNA binding capacity. The degree of mislocalization was again maximal for the K418N mutant. As for the cluster 2 and 3 truncating mutants, we did not observe major changes in condensate shape (**Supplementary Fig. 12c**).

These findings indicate that disease-related missense mutations can significantly alter the physiological condensation pattern of IRF2BPL. While some of the condensation phenotypes were less dramatically altered in comparison with what we observed with cluster 1 truncations, the substantial mislocalization of the enlarged mutant condensates to the cytoplasm suggested that also these mutants

may recruit WT IRF2BPL, thus triggering a fundamental mechanism in the molecular pathogenesis of NEDAMSS.

**Disease-related missense mutants can also lead to depletion of intranuclear WT IRF2BPL condensates**

To test whether missense mutants may recruit WT IRF2BPL, causing its mislocalization to the cytoplasm and the loss of its nuclear condensates, we used a similar approach as for the truncated mutants. We performed this analysis for the K418N mutant, which displayed the most dramatic alterations in LLPS behavior. This mutant was expressed for 24 h in NG108-15 neuronal cells as a GFP fusion, while GFP alone was expressed in control cultures. The subcellular distribution of the K418N mutant and WT IRF2BPL was studied using fluorescence microscopy (**Fig. 10e**) to visualize, in both transfected and non-transfected cells, the GFP green fluorescence of the mutant protein and the red fluorescence arising from the immunostaining of total IRF2BPL, either WT or mutant, as the available antibodies recognize both forms of the protein.

These experiments revealed that the number of red-fluorescent IRF2BPL-positive condensates in the nuclei of transfected cells expressing IRF2BPL-K418N was significantly reduced in comparison with neighboring non-transfected control cells (**Fig. 10e,f**; $F_{(1,107)} = 5.20$, $p<0.03$, two-way ANOVA, construct x transfection interaction; $p<0.04$, NK post hoc test). Thus, the expression of this missense mutant can deplete the nuclear pool of IRF2BPL condensates, as found for the truncated mutants. Such loss of nuclear IRF2BPL condensates was not observed in control cells expressing GFP alone (**Fig. 10f** $p = 0.78$, NK post hoc test). The number of nuclear IRF2BPL condensates was significantly lower in cells expressing IRF2BPL-K418N-GFP in comparison to cells expressing GFP alone (**Fig. 10g**; $p<0.01$, t-test).

Most of the IRF2BPL-positive red-fluorescent signal in transfected cells was also green-fluorescent, strongly suggesting that the WT protein was recruited into the mostly cytoplasmic mutant

condensates. Given the lack of WT-specific antibodies, other mechanisms (e.g., enhanced degradation) leading to the loss of nuclear IRF2BPL condensates cannot be ruled out in principle.

These findings indicate that the nuclear depletion of WT IRF2BPL, i.e., a key molecular mechanism of NEDAMSS, can also be triggered by missense mutations that are associated with aberrant LLPS and WT protein mislocalization, in a similar manner as observed for the truncating mutations.

## DISCUSSION

The results of our experiments identify the aberrant LLPS of mutant forms of IRF2BPL as the biophysical basis of the extranuclear recruitment and nuclear depletion of wild-type IRF2BPL, a fundamental event in the molecular pathogenesis of NEDAMSS. In turn, this process triggers the activation of downstream cellular and molecular mechanisms of the disease, including transcriptional dysregulation, with WNT1 upregulation, and neuronal cell dysfunction.

Specifically, we uncovered the ability of the IRF2BPL protein to physiologically undergo intra- and extra-nuclear LLPS, identifying the polyA/polyQ-bearing portion of its central region, and the flanking ZnF, as the key drivers of this process. Furthermore, we found how both truncating and non-truncating disease-related mutations lead to an aberrant LLPS behavior of IRF2BPL, with alterations in the morphology, ultrastructure, biochemical properties, dynamic behavior, and subcellular localization of its condensates. The aberrant, mostly cytoplasmic assemblies formed by mutant IRF2BPL sequester the WT protein, removing it from the nuclear and axonal compartments, triggering downstream disease mechanisms, with WNT1 upregulation, and lead to alterations in the electrophysiological properties of neuronal cells indicative of cellular damage.

These findings uncover the ability of IRF2BPL to undergo LLPS under physiological conditions and identify a key role for aberrant LLPS in the molecular pathogenesis of NEDAMSS and *IRF2BPL*-related disorders.

### A revised domain organization of the IRF2BPL protein

We have undertaken a systematic analysis of the IRF2BPL central region which, based on our findings, is not simply a 'variable' LCR[11]. These analyses led to a substantial redefinition of the domain structure of IRF2BPL. We found that the vast central region of the protein is characterized by a complex internal articulation of heterogeneous parts. Indeed, we showed how it contains three LCRs, with differential compositional features and structural propensity, one of which ($LCR_1$) plays a key role in LLPS, and a

central domain (HCD) with sequence/structure similarities to DNA-binding domains. Our findings support a revised IRF2BPL domain structure, with three distinct LCRs and three non-LCR domains (ZnF, HCD, RING).

This revised domain organization of the IRF2BPL protein can help rationalize the distribution of pathogenic mutation sites along the *IRF2BPL* gene. Indeed, 75% of the missense pathogenic mutations hit non-LCR domains, whereas most truncating mutations affect the LCRs. This differential distribution may be explained by the fact that single amino acid substitutions are more likely to impact structurally constrained domains rather than conformationally more flexible LCRs. Conversely, truncating mutations within LCRs lead to the production of variably truncated protein fragments from this intronless gene[27] that can gain non-physiological functions − like aberrant LLPS as we have found − contributing to the pathogenesis of the disease. These conclusions are consistent with previous analyses, showing that LCRs have greater mutational robustness than ordered protein domains and that truncating mutations are enriched in LCRs[37,38].

Consistent with these views, the HCD that we identified displays disease-related mutational patterns typical of highly constrained protein domains and a high degree of evolutionary conservation. Notably, a previous survey in the human population also found that one of the two 'constrained coding regions' in IRF2BPL falls within the HCD[55], supporting, together with our findings, the notion that this domain has an important functional role. Our experiments with missense mutants indicate that the HCD contributes to the regulation of IRF2BPL LLPS, possibly by mediating or modulating DNA binding. The possible loss of DNA binding induced by the K418N mutation may be responsible for the observed cytoplasmic mislocalization of IRF2BPL-K418N. Additional functions of the domain cannot be ruled out. Our CD structural analyses corroborate the AlphaFold model in showing that the HCD core forms a stable structural configuration similar to that which binds DNA in structurally related MBDs. Future experiments will have to fully characterize this domain, including its ability to bind methylated versus non-methylated DNA[56] and/or RNA.

**IRF2BPL physiologically undergoes LLPS driven by a ZnF domain and polyA/polyQ repeats**

We found that IRF2BPL undergoes LLPS in neuronal and non-neuronal cells, both in the nucleus and in the cytoplasm. Such condensation displays canonical hallmarks of an LLPS-driven process, in morphological, kinetic, and biochemical terms, as observed for other LLPS-prone polyA/polyQ proteins[16,20]. Remarkably, IRF2BPL condensates were also found in axons, neurites, and growth cones, indicating that the protein is trafficked in a cell-wide manner in neuronal cells and may have unknown functional roles in axonal and synaptic compartments.

We have mapped the IRF2BPL portion that is critical for LLPS-driven condensation to the N-terminal region (NTR), comprising the ZnF and the polyA/polyQ $LCR_1$. The condensation behavior of the IRF2BPL NTR appears to rely on the combination of the ZnF and $LCR_1$. The ZnF alone appears to be the trigger of the process as it can autonomously undergo LLPS, both in vitro and in the cellular context. A similar interplay between LCRs and flanking domains has been observed for other polyA and polyQ proteins. Indeed, the polyA/polyQ region driving RUNX2 condensation and aggregation displays a diffuse distribution in cells when expressed without the rest of the protein[13,20]. The condensation and aggregation of the N-terminal huntingtin (HTT) fragment, driven by an expanded polyQ tract, is enhanced or reduced, respectively, by an upstream helical peptide (N17) and a downstream polyproline tract[12,57,58]. As for these proteins, IRF2BPL condensation appears to rely on the interplay of flanking protein regions with different degrees of sequence complexity and LLPS propensity. It is interesting to note how the presence of proline-rich $LCR_2$ and $LCR_3$ in truncated forms of the protein (R331X and A708X, respectively) leads to less marked alterations in LLPS patterns in comparison with upstream truncations leaving only $LCR_1$, or portions of it, downstream of the ZnF. Thus, the presence of $LCR_2$ and $LCR_3$ may negatively regulate the LLPS propensity of the more N-terminal regions of IRF2BPL. Ultimately, the overall LLPS behavior of the entire protein may be determined by the functional interplay of its constituent domains which can have opposite effects on the process of phase separation.

**Aberrant LLPS in the molecular pathogenesis of NEDAMSS and *IRF2BPL*-related disorders**

Our findings indicate that disease-related missense and truncating mutations alter IRF2BPL physiological LLPS in the cellular context, leading to mislocalization of the WT protein produced by the normal allele. $LCR_1$ is the protein region most affected by truncating mutations (cluster 1) which lead to the highest degree of distortion of the intracellular LLPS pattern in comparison with that of WT IRF2BPL. Downstream truncations and missense mutations, while profoundly altering the LLPS behavior of the protein, had relatively less dramatic effects than cluster 1 truncations. Notably, this may correlate at the clinical level with the fact that missense mutations and distal truncations appear to cause milder phenotypes, as observed in the first described NEDAMSS patients[1,27]. However, careful genotype-phenotype correlations will need to be performed once a sufficiently high number of clinical cases have been characterized, to support these conclusions.

We have studied the impact of disease-related mutations on the physiological LLPS of IRF2BPL in the in vivo cellular environment. This approach was based both on technical considerations and current trends in the study of LLPS pathophysiology. Indeed, given the insolubility of the polyA/polyQ-rich IRF2BPL protein, it was not technically possible to perform in vitro LLPS on this protein or its disease-related mutants, although these could have provided additional support for the in vivo evidence, as we obtained for the isolated ZnF domain, i.e. the key physiological LLPS trigger in the protein. To our knowledge, IRF2BPL has not been previously purified in a soluble form, consistent with known technical obstacles in polyA/polyQ protein purification[47,48,49]. While the in vitro IRF2BPL production may represent a direction for future work, a growing body of evidence in the literature questions the relevance of LLPS assays performed under artificial in vitro conditions to protein behavior in the relevant cellular environment[59,60,61,62]. Nevertheless, we obtained in vitro evidence supporting the role of the more soluble ZnF as a key LLPS-driving domain of IRF2BPL in vivo. Notably, ZnF LLPS in vitro was observed using a physiological buffer without any highly artificial biochemical triggers such as crowding agents (e.g., PEG-8000), high ionic strength, or severely altered pH[62].

Truncated mutants, consisting of the ZnF and variable portions of the polyA/polyQ $LCR_1$, form condensates with altered localization, morphology, and ultrastructure in comparison with those formed

by WT IRF2BPL. The aberrant condensates have a mostly cytoplasmic localization, which may be explained by the lack of the NLS and by the fact that, based on our findings, the ZnF appears to carry a nuclear export signal (NES). Indeed, when expressed alone, the ZnF is cytoplasmic whereas ΔZnF-IRF2BPL is purely nuclear. Thus, the distribution of IRF2BPL may derive from the combined action of the NLS and the NES, and the action of the latter would prevail in the case of truncated mutants.

In morphological terms, the condensates formed by the truncated IRF2BPL mutants displayed, with the shortening of their length, progressively increasing ellipticity. This was paralleled at the ultrastructural level by the emergence of a fibrillary organization within the condensates. These condensates also coalesce more rapidly, display blunted FRAP, and are more resistant to disruption by 1,6-Hex and by the protein degradation machinery. Overall, the molecular assemblies of these fragments appear to drift from a liquid-like, transient condensation towards a more solid-like and persistent aggregated state, as frequently observed in neurodegenerative disorders[24]. A similar transition, with the initial LLPS-driven assembly of liquid-like condensates preluding to the consolidation of solid-like aggregates, has been observed for the polyQ-expanded N-terminal fragments of HTT[23], the molecular mediator of Huntington disease.

We have studied multiple cluster 1 $LCR_1$ truncation mutants (115, 122, and 172 residue-long) whose aberrant condensation phenotypes appeared overall to be inversely related to their length. Considering that the ZnF alone (a.a. 1-75) condensates at low levels, and that the Q116X mutant has the most pronounced aberrant condensation, the IRF2BPL 76-115 region appears to be the strongest LLPS-promoting subregion within $LCR_1$. This region overlaps with $CC_1$ (a.a. 76-136), containing the first polyA and part of the polyQ repeat. Previous studies on hybrid polyA/polyQ CCs indicate that their polyA section is considerably more stable than the polyQ one[13]. Thus, the loss of the $CC_1$ distal portion (a.a. 116-136, including 13 Qs) in Q116X may stabilize the polyA CC, enhancing its LLPS and subsequent more solid-like assembly. Such progressive $CC_1$ stabilization, arising from the gradual erosion of its relatively less stable C-terminal polyQ portion, may explain the length-dependent LLPS phenotypes displayed by the cluster 1 truncated mutants.

Beyond these length-based correlations for cluster 1 truncations, other structure-function interpretations may be given for the altered LLPS behavior of more distally truncated or missense mutants. Indeed, the overall LLPS behavior of the IRF2BPL protein conceivably relies on the interplay between its constituent domains, some of which are lost in truncated mutants. The variable loss of these domains may alter the balance between LLPS-promoting regions (ZnF+$LCR_1$) of the protein and downstream LCRs, or structured domains, with potential LLPS-limiting roles. For instance, RING domains can inhibit LLPS[63]. Moreover, the LLPS behavior of IRF2BPL may also be regulated by its DNA-binding capacity, which may be altered by missense mutations in domains known or predicted to bind DNA. For instance, the only known missense mutation targeting the ZnF, i.e., F30L, is predicted to alter zinc binding, based on our bioinformatics analyses. This may in turn alter the DNA binding ability of the protein[7], favoring the observed alterations in condensation patterns. Mutations in the HCD (e.g., K418N) may have a similar impact. Missense mutations in the RING domain, such as N770H, may also potentially dysregulate LLPS[63].

The aberrant condensates and aggregates formed by a mutant polyA/polyQ protein can sequester non-mutant protomers of the same protein, as well as other cellular proteins, causing their loss-of-function[12,13,14,15,16]. This phenomenon has also been recently observed for IRF2BPL truncated mutants which appear to recruit the nuclear WT protein, a process representing a key molecular dysfunction underlying the molecular pathogenesis of NEDAMSS[27]. These four mutants (E172X, Y173X, R188X, A708fs) contain the intact LLPS-driving region that we identified and displayed some differences in their ability to recruit WT IRF2BPL, which were tentatively attributed to the variable retention of the NLS in them[27].

We analyzed instead the effects of upstream mutations truncating within $LCR_1$, as they target the LLPS-driving region of IRF2BPL. We found that the cytoplasmic assemblies of all three mutants (Y173X, Q123X, and Q116X) recruit the WT protein, in agreement with Sinha Ray and coll. (2022)[27]. Our results provide mechanistic insight into this process. Indeed, they show that the degree of IRF2BPL recruitment is inversely related to the length of the N-terminal mutant fragment. This observation can be rationalized considering the relative contribution of the ZnF and the polyA/polyQ repeats to the

process. The ZnF, which is present in all mutants, is involved in the N-terminal interaction between IRF2BPL protomers[64]. The polyA/polyQ repeats may enhance this interaction as a function of their CC stability[12] which, for the IRF2BPL NTR, is inversely related to $LCR_1$ length. Our observations of a lesser, but still significant degree of distortion of the LLPS pattern of the A708X mutant, which also retained a more nuclear localization in comparison with mutants with upstream truncations, are also consistent with the findings of Sinha Ray and coll. (2022)[27]. Overall, these observations suggest that longer truncated forms of the protein may cause milder phenotypes than shorter ones. This hypothesis arising at the molecular and cellular level warrants future genotype-phenotype analyses once a larger number of clinical cases is studied.

We found that also missense mutations can alter IRF2BPL LLPS. The fact that some of them have a lesser overall impact on LLPS than truncating ones may also rationalize their milder phenotypic impact[27]. Thus, our findings indicate that aberrant LLPS may be a prevalent mechanism, triggered by both truncating and missense mutations, in the molecular pathogenesis of NEDAMSS and *IRF2BPL*-related disorders. Moreover, they suggest the possibility that the degree of LLPS disruption may correlate with phenotypic impact and disease severity, ultimately providing a basis to rationalize future genotype-phenotype correlations. This scenario does not rule out that loss-of-function mechanisms, caused by either missense or truncating mutations, may have a relevant role in the molecular pathogenesis of NEDAMSS, either autonomously or in combination with aberrant LLPS.

Dysfunctional LLPS has been implicated in developmental and neurodegenerative disorders, including polyQ- and polyA-expansion diseases[20,23]. Our findings indicate that LLPS and its dysregulation can play a fundamental role in the molecular pathogenesis of NEDAMSS and *IRF2BPL*-related disorders by inducing WT protein mislocalization and triggering downstream cellular and molecular mechanisms underlying these disorders.

NEDAMSS and *IRF2BPL*-related disorders have been attributed to the loss-of-function of the mutant IRF2BPL variants[1]. However, true haploinsufficiency may often not occur as, in patient-derived cells expressing truncated mutants, WT IRF2BPL protein levels can be normal or mildly reduced[27]. Thus, additional gain-of-function and dominant-negative mechanisms should be considered. Even if

not directly toxic[1], aberrant IRF2BPL condensates may gain cellular toxicity through the homotypic recruitment of WT IRF2BPL and, conceivably, through the heterotypic recruitment of other CC-prone cellular proteins[12,13,14,16,27,65].

The pathogenesis of NEDAMSS and *IRF2BPL*-related disorders has been hypothesized to derive also from neurodegenerative mechanisms, in addition to developmental ones[8,66,67]. The analogies that we found between the IRF2BPL mutants and protein fragments involved in polyQ diseases strongly support the hypothesis that *IRF2BPL*-related disorders and neurodegenerative diseases may share fundamental molecular mechanisms related to the formation of dysfunctional protein assemblies.

Ultimately, our findings identify aberrant protein LLPS as a key player in the molecular pathogenesis of NEDAMSS and *IRF2BPL*-related disorders, which lie at the intersection of neurodevelopmental and neurodegenerative diseases.

## METHODS

### Ethical statement

The research complies with the ethical regulations of the University of Turin, Vita-Salute San Raffaele University, and other institutions that participated in the study. Human neural progenitor cells (NPCs), were obtained from available induced pluripotent stem cells (iPSCs; OSRi005-A cell line, previously established by Iannielli et al. (2022)[68]). The line is deposited at https://hpscreg.eu/cell-line/OSRi005-A and is available for academic research. Its use has been authorized by Vita-Salute San Raffaele University (protocol authorization number: RF-2018-12366703).

### Bioinformatics characterization of human IRF2BPL

The primary sequences of human IRF2BPL and its orthologs in vertebrates were derived from the Uniprot (www.uniprot.org) and NCBI (www.ncbi.nlm.nih.gov/protein/) protein databases. The search for low-complexity protein regions (LCRs) in the IRF2BPL primary sequence, as plotted in **Fig. 1a**, ('LCR') was performed using the SEG algorithm[29] on PlaToLoCo[69]. The per-residue CC propensity for the IRF2BPL primary sequence was obtained using Paircoil2[70]. The polyA/polyQ CC-prone IRF2BPL region ($CC_1$) was identified using a Pscore threshold of 0.2 (1-Pscore $\geq$ 0.8 as in Fiumara et al. (2010)[12]; **Fig. 1a** and **Supplementary Fig**. 1b). $CC_2$ was identified using a threshold of 0.4 to include its initial part which was elided using the 0.2 threshold. The per-residue LLPS propensity (**Fig.** 1A) was calculated using FuzDrop[71], with a 0.6 pDP score threshold. The list of disease-related mutations, reported in **Supplementary Data 5** and **Figs. 1a**, **4a**, and **10b** was derived from literature sources available in October 2023. The AlphaFold2[72] structural model of human IRF2BPL was generated using ColabFold (v. 1.5.0; https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb)[73] on April 27, 2023 (Supplementary Data 1, model 3). The position of α-helical and β-strand structures in this model (**Fig. 1a**) was obtained using UCSF Chimera (version 1.14)[74]. The structural models of IRF2BP1 (Uniprot ID: Q8IU8) and IRF2BP2 (Uniprot ID: Q7Z5L9) were derived from the Alphafold

website (available at https://alphafold.ebi.ac.uk/entry/AF-Q8IU81-F1 and https://alphafold.ebi.ac.uk/entry/AF-Q7Z5L9-F1). The two latter models were last accessed on November 26, 2025. The pLDDT scores and the predicted aligned error (PAE) plots of the three models are reported in **Supplementary Figure 10**.

**Revised domain structure of human IRF2BPL**

Domain boundaries in our revised IRF2BPL domain structure (**Fig. 1a**) were derived by integrating information from the Pfam database, the AF model, and the Paircoil2 prediction, based on the Uniprot reference sequence (Q9H1B7). These boundaries are: ZnF (a.a. 1-75, comprising the ZnF and short flanking peptides), $LCR_1$ (a.a. 76-173, starting from the beginning of $CC_1$, encompassing the polyA/polyQ α-helix through the downstream α-helix of the second polyA repeat until Y173), $LCR_2$ (a.a. 174-329, starting from a short polyP stretch (a.a. 174-177) and comprising a long proline-rich random coil region until $CC_2$), HCD (a.a. 330-514, going from $CC_2$ to the end of the last structured element after the α/β-sandwich), $LCR_3$ (a.a. 515-706, from the end of $LCR_2$ to the beginning of the RING domain), and RING (a.a. 707-796, from the beginning of the RING domain as defined by InterPro (SSF57850, a,a, 707-783) to the C-terminal end of the protein, including a short terminal peptide, i.e. a.a.784-796). The amino acid composition of human IRF2BPL and its domains was calculated using Expasy ProtParam (https://web.expasy.org/protparam/)[75]. The composition of proteins of the entire human proteome was calculated on the Uniprot reference proteome using ad hoc Perl scripts[31]. For each IRF2BPL domain, we calculated the percent occurrence of each one of the 20 amino acids and their classes, i.e., small (sma; A, G), aliphatic (ali; I, L, V), aromatic (aro; F, Y, W), sulphurated (sul; C, M), positively charged (ch+; H, K, R), negatively charged (ch-; D, E), polar (pol; N, Q, S, T) and cyclic (cyc; P). We also calculated two per-residue scores expressing the local sequence simplicity (SIM) and repetitiveness (REP) in a 20-residue sliding window along the IRF2BPL primary sequence[30].

**Evolutionary analyses**

For evolutionary analyses, an alignment of IRF2BPL orthologous sequences belonging to 9 major clades of different stem age along the vertebrate lineage (Primates, Rodentia, Laurasiatheria, Atlantogenata, Marsupialia, Sauropsida, Lissamphibia, Actinopterygii, Chondrichthyes; IDs in **Supplementary Data 4**) was obtained using the NCBI Cobalt tool[76] (available at www.ncbi.nlm.nih.gov/tools). The entropy at each position of the alignment was calculated using Entropy-One (www.hiv.lanl.gov/content/sequence/ENTROPY). Clade stem ages were obtained from TimeTree[77] (available at http://timetree.org/), as divergence times (median) of each clade from Euarchontoglires. The stem age of Primates and Rodentia, within Euarchontoglires, was derived from the divergence time between Glires and Euarchonta[30] (stem ages: Primates, 87 mya; Rodentia, 87 mya; Laurasiatheria, 94 mya; Atlantogenata, 99 mya; Marsupialia, 160 mya; Sauropsida, 319 mya; Lissamphibia, 352 mya; Actinopterygii, 429 mya; Chondrichthyes, 462 mya). For each IRF2BPL ortholog, we calculated the i) percent occurrence of each amino acid, the ii) mean per-residue scores of local degree of sequence simplicity (SIM) and repetitiveness (REP) in a 20-residue sliding window around each residue, and the iii) total length of AARs (≥4 residues) of a given amino acid for all its constituent domains using previously developed Perl scripts[30,31]. Then, we calculated the mean values of these three classes of parameters across orthologs in each vertebrate clade and normalized them to those of the Chondrichthyes. Finally, we calculated the Pearson's r coefficients between the mean values of these parameters in each clade and clade stem ages. We also measured, for each IRF2BPL domain, an index of compositional variation across clades. Towards this aim, we first calculated, for each IRF2BPL domain, the mean percent occurrence of each amino acids across the species of each clade. Then, we calculated the coefficient of variation (CV, i.e., standard deviation/mean) across clades of the mean percent occurrence of each amino acid. Finally, we calculated, for each domain, the mean CV for the 20 amino acids.

**Analyses of *IRF2BPL* mutation sites**

A list of known IRF2BPL pathogenic mutations (**Supplementary Data 5**) was compiled from multiple publications[1,3,5,10,26,78] and from the 'The Leiden Open Variation Database', version 3.0[79] (LOVD;

www.lovd.nl/3.0/home; accessed on Oct. 15, 2023). The local frequency of disease-related mutation sites along the IRF2BPL primary sequence (**Figs. 4a** and **10b**) was plotted using a previously developed Perl script[80]. The potential impact of disease-related mutations of interest was predicted using MutPred2 for missense mutations (http://mutpred.mutdb.org)[35], MutPred-LOF (http://mutpred2.mutdb.org/mutpredlof/)[81] for nonsense and frameshift mutations, and MutPred-Indel (http://mutpred2.mutdb.org/mutpredindel/)[82] for in-frame indels. To predict the mutation tolerance at each IRF2BPL residue, we used MetaDome[32] (v. 1.0.1; available at https://stuart.radboudumc.nl/metadome). Per-residue pathogenicity scores for amino acid substitutions in IRF2BPL were obtained using AlphaMissense[33] (available at https://alphamissense.hegelab.org; see **Fig. 1a**, 'αM'). The position of IRF2BPL codon repeat frameshifting sites (CRSF) was derived from Ren et al. (2024)[83]. The translational readthrough probabilities at premature stop codons introduced by nonsense mutations in IRF2BPL were derived from Toledano et al. (2024)[84].



**In silico characterization of the IRF2BPL HCD domain**

The HCD domain was initially identified as a highly structured central region in the IRF2BPL AlphaFold model. Structural homology screenings were initially performed using DALI[85] (available at http://ekhidna2.biocenter.helsinki.fi/dali/), which indicated the murine MPN domain-containing protein (PDB: 7ydw-D)[86] as the closest match, in particular at a core region formed by a long β-hairpin and a neighbouring α-helix that may mediate DNA binding. A second DALI search using only this HCD core as the query indicated a structural similarity with multiple methyl-DNA-binding domains (MBDs) in human and non-human species (**Supplementary Data 7**). We matched the solved structures of some of these proteins, i.e. MBD2 (*G. gallus*; PDB: 2ky8), MBD4 (*M. musculus* PDB: 3vyq,), MBD5 (*A. thaliana*; PDB: 7feo), to the HCD core fragment (**Fig. 9a**, upper right panel) using Chimera[74] (v. 1.14). We included in this analysis a closely related structure of human MeCP2 bound to DNA to visualize the possible arrangement of an HCD-DNA complex (**Fig. 9a**). The Uniprot IDs of the proteins whose fragments are aligned with the HCD core domain in **Fig.** 9b are: Q9UIS9 (MBD1, *Homo sapiens*), Q9UBB5 (MBD2, *H. sapiens*), P51608 (MECP2, *H. sapiens*), Q9UIS9 (MBD1, *H. sapiens*), Q9UIS9

(MBD1, *H. sapiens*), Q9VHB7 (MBD2/3, *Drosophila melanogaster*). The secondary structure elements formed by the aligned primary sequences of human proteins were derived from their solved structures using the PDBsum tool (PDB IDs: 6d1t for MBD1, 6c2f for MBD2, 6c1y for MeCP2). For *Drosophila* MBD2/3 they were derived from the AlphaFold model related to the Uniprot Q9VHB7 sequence.

The residues forming the hydrophobic core in the aligned MBD domain sequences in **Fig.** 9b were derived from Wakefield et al. (1999)[87] and Hung et al. (2003)[88]. Those involved in DNA binding from Ohki et al. (2001)[89] and Hung et al. (2003)[88].

**Experimental characterization of the HCD core structure using circular dichroism (CD)**

A peptide encompassing the human IRF2BPL HCD core region (a.a. 395-448) was chemically synthesized (LifeTein LLC) with N-terminal acetylation, C-terminal amidation, and purity >95%. The peptide was dissolved in a saline buffer (100 mM NaCl, 20 mM phosphate buffer, pH 7.4)[12] to generate a stock solution (~3 mg/mL). This solution was slightly opalescent and peptide concentration was calculated based on peptide weight[30,51]. Before each measurement, the peptide was further diluted in saline buffer (0.2 mg/mL), transferred to a 1-mm quartz cuvette, and CD spectra (190–260 nm) were collected using a J-815 spectropolarimeter (Jasco). Blank spectra from the saline buffer alone were subtracted from the peptide spectra. The mean residue molar ellipticity (MRME or [θ]) was calculated as in Fiumara et al. (2010)[12]. To test structural stability, the samples were heated (from 5°C to 75°C; 10 °C/min) while recording ellipticity at 222 nm. Data were analyzed using Spectra Analysis (as part of the Spectra Manager suite; Jasco). Spectra deconvolution was performed using the Contin-LL algorithm[90] available on DichroWeb[91] (available at http://dichroweb.cryst.bbk.ac.uk).

**Plasmids and molecular cloning**

In all cloning procedures, polymerase chain reaction (PCR) products were obtained using the CloneAmp-HiFi PCR Premix (Clontech-Takara). The PCR primer oligonucleotides were designed with 15/18-bp extensions suitable for the In-fusion cloning system (Clontech-Takara). The sequence of these and other oligonucleotide primers used in this study (see below) is reported in the **Supplementary Data**

**8** file. PCR products were separated by agarose gel electrophoresis and purified using QIAquick (Qiagen) or NucleoSpin (Macherey-Nagel) kits. Purified PCR products were subcloned into linearized plasmids of interest using the In-fusion system. Plasmids were amplified using Stellar chemically competent cells (Clontech) and purified using QIAprep Spin Mini-/Maxi-prep (Qiagen), NucleoSpin plasmid Mini kit (Macherey-Nagel), or NucleoBond Xtra MIDI EF (Macherey-Nagel) kits. For all plasmids that were generated, the protein-coding sequences were verified by Sanger sequencing. The intronless *IRF2BPL* gene open-reading frame (ORF) was amplified by PCR, using the CloneAmp-HiFi PCR Premix from human genomic DNA extracted from HEK293F cells (ThermoFisher) using the QIAamp DNA Investigator Kit (QIAGEN), with primers bearing 15 bp extensions suitable for its subcloning into the pAcGFP-N1-In-Fusion-Ready plasmid (Clontech-Takara) using the In-Fusion cloning system to generate the p-IRF2BPL-GFP plasmid (forward primer: 5'-AAGGCCTCTGTCGACCACCATGTCGGCGGCGCAGGTGTCC
TCGTCC-3'; reverse primer: 5'-AGAATTCGCAAGCTTAGGGTCTCTCTCCTTTTTCACTTTAA
CATC-3'). A Kozak sequence fragment ('CACC') was included upstream of the IRF2BPL start codon. The amplified ORF (795 residues) contained 24 glutamines (Qs) in its polymorphic N-terminal polyQ repeat, corresponding to the most frequent allelic variant observed in the human population, in which the repeat varies between 20 and 31 residues[11]. Therefore, we used this prevalent allelic variant in our experiments. The Uniprot IRF2BPL protein sequence (Q9H1B7; 796 residues) corresponds to a 25Q variant bearing an additional Q residue. For simplicity, we referred in both the **Figure**s and the text to the Uniprot residue numbering, which is also used in the literature, in defining the IRF2BPL domain structure. Plasmids for the expression of IRF2BPL domains and mutant forms of the protein were generated using the In-fusion system. IRF2BPL fragments of interest were PCR-amplified using p-IRF2BPL-GFP as a template and suitable primers with 15 bp extension for their In-Fusion subcloning into the pAcGFP-N1-In-Fusion-Ready plasmid. A Kozak sequence fragment ('CACC') was included upstream of the start codon. A start codon was also added to express internal fragments of IRF2BPL. Using this approach, we generated plasmids for the expression of GFP-tagged fragments of interest, i.e. N-terminal region (NTR, a.a. 1-172), ZnF domain (a.a. 1-75), $LCR_1$ (a.a. 76-172) as well as NTR

(ΔNTD; a.a. 173-796) or ZnF (ΔZnF; a.a. 76-796) deletion mutants. Using the same approach, we generated the cluster 1 truncation mutants Y173X (equivalent to the NTR), Q123X (a.a. 1-122), and Q116X (a.a. 1-115), as well as truncation mutants mimicking cluster 2 (R331X) and cluster 3 (A708X) disease-related mutations. The ΔLCR1 internal deletion mutant, lacking a.a. 76-172, was obtained from two PCR-amplified fragments (a.a. 1-75 and a.a. 173-796) with suitable 18 bp extension for In-Fusion cloning. The QuikChange Multi Site-Directed Mutagenesis Kit (Agilent) was used to generate missense mutants, i.e. F30L, P372R and K418N, from the p-IRF2BPL-GFP plasmid with primers designed using the QuikChange Primer Design online tool (v. 2.23; available at https://www.agilent.com/store/primerDesignProgram.jsp).

To generate suitable plasmid vectors for optoDroplet induction experiments[46], PCR-amplified sequences encoding IRF2BPL fragments of interest (NTR, ZnF, $LCR_1$) were cloned into the pX-mCh-Cry2 plasmid[16] using the In-Fusion system.

To generate plasmid vectors for the production of recombinant IRF2BPL in *E. coli*, N-terminally tagged with 6xHis-GST (thrombin/TEV cleavable), the coding sequence of the protein was PCR-amplified from the p-IRF2BPL-GFP plasmid and subcloned into the pGAT3 vector (Addgene, #112589) using the In-Fusion system.



**Cell culture and neuronal differentiation**

Human neural progenitor cells (NPCs), obtained from available induced pluripotent stem cells (iPSCs; OSRi005-A line (https://hpscreg.eu/), previously established by Iannielli et al., 2022[68]), were dissociated with Accutase and plated on Matrigel-coated 6-well plates (300,000 cells per well) in NPC medium. After two days, the differentiation medium (Neurobasal medium (ThermoFisher) with Pen/Strep (ThermoFisher; 1%), B27 (ThermoFisher; 1:50), SU5402 (Sigma-Aldrich, 10 μM), PD0325901 (Sigma-Aldrich, 8 μM), and DAPT (Sigma-Aldrich, 10 μM)), was added and replaced daily for two days. At day 3, cells were detached either using Accutase or heating (10 min at 37 °C) to obtain a single-cell suspension. Cells were centrifuged, counted, and seeded at a density of 55,000 cells/$cm^2$ onto plates coated with poly-L-lysine/laminin/fibronectin (Sigma-Aldrich, 100 μg/ml,

2 µg/ml, 2 µg/ml, respectively) in neuronal maturation medium (Neurobasal A with 1X B-27 supplement, 2 mM glutamine, 1% Pen/Strep, BDNF (Peprotech, 20 ng/ml), ascorbic acid (Sigma-Aldrich, 100 nM), laminin (1 µg/µl), DAPT (10 µM), dibutyryl-cAMP (db-cAMP; Selleckchem, 250 µM)) supplemented with the ROCK inhibitor Y27632 (10 µM) for the first 24 h. The culture medium was replaced the next day to remove the ROCK inhibitor, and then half of the medium was replaced twice a week.

NG108-15 cells were obtained from the American Type Culture Collection (ATCC; catalogue # HB-12317; validated and tested for Mycoplasma contamination (negative) by ATCC) were maintained in Dulbecco's Modified Eagle's Medium (DMEM; Sigma) with 10% fetal bovine serum (FBS; Sigma) and 1X HAT supplement (ThermoFisher)[92]. To induce neuronal differentiation, NG108-15 cells were plated onto poly-D-lysine-coated glass coverslips in a 24-well plate and, after 24-48 h, the culture medium was replaced with a differentiation medium consisting of DMEM with 5% FBS, 1X HAT, 1 mM db-cAMP (Sigma), and 1 mM dexamethasone (Sigma), for 7-14 days[92]. Half of the differentiation medium was replaced on alternate days. For electrophysiological experiments, cells were plated in 35 mm cell culture dishes in maintenance medium. Four hours after plating, the medium was replaced with differentiation medium which was kept for 7–8 days. Half of the differentiation medium was replaced with fresh one on alternate days.

HEK293 cells (293-F; catalogue # 293-F; Thermo Fisher Scientific; validated and tested for Mycoplasma contamination (negative) by Thermo Fisher Scientific) were maintained following standard procedures at 37 °C with 5% $CO_2$ in DMEM (Thermo Fisher Scientific) supplemented with 10% FBS, 2mM L-glutamine, 100 units/mL penicillin, and 100 µg/mL streptomycin. The cells that were used for immunocytochemistry or plasmid transfection, were plated onto glass coverslips coated with poly-L-lysine (0.1 mg/mL, Sigma) in 24-well plates. For optoDroplet experiments, HEK293 cells were plated onto 50 mm plastic culture dishes with a central 14 mm glass coverslip (Mattek) coated with poly-L-lysine.

**Cell culture transfection**

Transfection of human neurons, NG108-15, and HEK293 cells was performed 24 h after plating using Lipofectamine 2000 (Thermo Fisher Scientific) following the manufacturer's protocol, using 0.25-1 μg of plasmid DNA for each coverslip in a single well of a 24-well plate with a 3:1 Lipofectamine (μl) to DNA (μg) ratio. Similar relative differences in terms of LLPS quantitation were observed between WT and mutant forms of IRF2BPL when transfecting with DNA concentrations in this range. Four hours after transfection, the culture medium of the NG108-15 cells was entirely replaced with fresh medium. Twenty-four hours after transfection, this medium was replaced with differentiation medium (half of which was replaced on alternate days). NG108-15 cells were fixed after 7-14 days of differentiation. HEK293 cells were fixed 24-48 h after transfection. The cells were fixed, after rinsing with PBS, using 4% PFA in PBS (pH 7.4) for 15-20 min at RT. After rinsing in PBS (3X), the coverslips were mounted onto microscopy slides using DAKO mounting medium (Agilent).

For electrophysiological experiments, NG108-15 cells were transfected at day 5-6 of differentiation with plasmid constructs for the expression of proteins of interest using Lipofectamine 2000. Electrophysiological recordings were performed after 24-48 hours at days 7-8 of differentiation.



**Immunocytochemistry (ICC)**

Cultured human neurons growing on glass coverslips were fixed in 4% paraformaldehyde (PFA) and 4% sucrose for 15 minutes. After rinsing (3X) with PBS, the cells were then permeabilized using 0.2% Triton X-100 in PBS, pH 7.4, for 1 minute, rinsed again in PBS (3X), and then incubated for 1 h in blocking solution containing 1% bovine serum albumin (BSA; Thermo Fisher Scientific). The cells were then incubated with primary antibodies, i.e., IRF2BPL (Abcam, cat. # ab221099; 1:1000) and MAP2 (Merck/Sigma-Aldrich, AB15452; 1:10000) for 2 hours at room temperature (RT) in PBS containing 1% BSA and 0.2% Triton X-100. After rinsing with PBS (3X), the cells were incubated with secondary fluorescently-conjugated antibodies (DyLight 649-conjugated goat anti-chicken IgG antibody (Rockland; cat. # 603-143-126), or DyLight 488-conjugated goat anti-Rabbit IgG (Thermo Fisher Scientific, cat. # 35552) for 1 hour at RT and rinsed with PBS. Finally, nuclei were stained with 2 mg/ml Hoechst (Thermo Fisher Scientific) in PBS and, after rinsing in PBS, the coverslips were

mounted on microscopy slides using the ProLong Glass Antifade mounting medium (Thermo Fisher Scientific).

ICC of NG108-15 and HEK293 cells was performed with a similar protocol with minor variations. NG108-15 cells were fixed at days 7-14 of differentiation and HEK293 after 24-48 h after plating. The cells were immunostained using the same anti-IRF2BPL primary antibody (Abcam, cat. # ab221099; 1:200) and anti-rabbit secondary antibodies fluorescently labelled with Alexa Fluor 488 (1:500) or 555 (1:1000; both from Thermo Fisher Scientific, cat. # A-11008 and # A-21428, respectively). The cells were incubated with DAPI (Thermo Fisher Scientific) for 15 min at RT to stain nuclei. Coverslips were then mounted onto microscopy slides using DAKO mounting medium (Agilent).

**Confocal fluorescence microscopy and super-resolution imaging (Airyscan, STED)**

Confocal and super-resolution imaging (Airyscan) of NG108-15 and HEK293 cells was performed using a ZEISS LSM 800 Airyscan microscope (Zeiss) with a 63X immersion objective using default Airyscan settings. Some confocal images of these cells were acquired using a Fluoview FV300 confocal microscope (Olympus). Multichannel fluorescence imaging, also in live-cell time-lapse experiments, was performed using a MICA imaging microhub (Leica Microsystems), equipped with 20X and 63X objectives. Images acquired with MICA were further processed using Thunder deconvolution (within Leica Application Suite X, LasX, v. 1.4.6; Leica Microsystems). Confocal and STED images of differentiated human neurons were obtained using a Leica TCS SP5 STED (Leica) inverted confocal/STED microscope. Samples were viewed with a 100X Apochromat NA = 1.4 oil-immersion objective. Excitation was provided by a white light laser and the 488 and 633 nm wavelengths were obtained by an acoustic-optical tunable filter (AOTF), detecting fluorescence in the 500–575 nm range for AlexaFluor 488 and the 650–750 nm range for AlexaFluor 647, with the pinhole set to 1 Airy size and pixel size set at 13 nm. Line scanning speed ranged from 1000 to 1400 Hz in standard acquisition mode, and line averages ranged from 96 to 128. In gated-STED mode, the 592 nm CW laser beam was superimposed at a typical power of 180 mW before the objective. In this modality, we also enabled a

second detector in the spectral range from 400 to 475 to collect the Hoechst fluorescence two-photon excited by the STED beam[93]. The detector gating time was set to 1.4 ns for STED imaging. Image analysis was carried out using FIJI software[94], applying background subtraction (rolling ball radius: 50 px) and Gaussian blur (radius: 1 px) to improve the signal-to-background and signal-to-noise ratios.

**Live cell imaging**

Short-term live-cell imaging of HEK293 cells expressing IRF2BPL-GFP was performed using a GE Healthcare DeltaVision Ultra fluorescence microscopy system. Time-lapse videos of up to 300 seconds were recorded capturing a frame every 2 seconds. Images were processed using FIJI software (NIH).

To perform longer (12 h) live-cell time-lapse imaging experiments to measure condensate coalescence dynamics, NG108-15 cells were plated at low density in 10-well glass bottom Cellview microslides (Greiner Bio-One) and differentiated for 5-6 days. On days 6-7, cells were co-transfected with plasmids encoding mCherry, as a whole-cell marker, and either IRF2BPL-GFP, IRF2BPL-Y173x-GFP, or IRF2BPL-Q123x-GFP. Imaging, using a MICA Microhub system (Leica), started ~24 h post-transfection and continued for 12 h at 37 °C in a 5% $CO_2$ atmosphere using an on-stage incubator (OkoLab). For each construct, 3-4 imaging fields, each containing at least 5-10 cells, were selected and imaged every 30 minutes for 12 h. Illumination of the cells was adjusted to prevent phototoxicity in repeated scans. The acquired image series were processed using Thunder deconvolution (Leica) and adjusted for brightness and contrast to highlight fluorescent condensates. The images were cropped to generate single-cell time-lapse series in which the number of condensates per cell was automatically quantified at each time point using an ad hoc CellProfiler pipeline. Only cells with more than 5 and fewer than 40 condensates at the beginning of the experiment were included in the analysis. These criteria excluded from the analysis cells with no further condensate coalescence potential or morphologically anomalous cells with unusually large size and a high number of condensates. To study coalescence dynamics, we quantified both the absolute and relative variation over time in the number of condensates per cell. In the latter case, for each cell, the number of condensates at each time point was normalized to the number of condensates at the beginning of the experiment. For each cell, 12 data

points were obtained from the 24-image series by averaging the two values measured at 30-minute intervals during each hour of the experiment.

**Liquid-liquid phase separation (LLPS) analysis**

The spontaneous formation of condensates by fluorescently tagged human IRF2BPL, its fragments, deletion mutants, and disease-related mutant forms was quantified by confocal fluorescence imaging in HEK293 and NG108-15 cells 24-48 h after transfection. Cell cultures expressing the different mutants were fixed in each experiment at the same time to allow quantitative comparisons between condensation patterns which evolve dynamically (see **Supplementary Movies 2-4**). Given the dynamic nature of LLPS and the coalescence dynamics of the condensates, cells expressing a LLPS-prone construct display relatively heterogeneous condensation patterns when fixed at a given time point. Thus, for quantitative imaging analyses, we used at least 5 Z-stack maximum intensity projections of fluorescence confocal images of large microscopy fields (233x233 μm, or 350x350 μm), each containing tens of cells. These images were converted into 8-bit images using ImageJ (NIH). The brightness and contrast of each image were alternatively adjusted to highlight either the fluorescent condensates, identified as discrete bright rounded puncta with an above-background fluorescence signal, or the entire profiles of the fluorescent nuclei/cells. These two sets of images were used to automatically quantify, for each microscopy field, the proportion of the fluorescent (GFP- or mCherry-positive) surface area occupied by condensates, the number of condensates per unit area of the fluorescent nuclei/cells, the size and eccentricity of each condensate, and the total area occupied by fluorescent nuclei/cells using ad hoc CellProfiler pipelines. To quantify the nuclear versus cytosolic localization of condensates, we generated two sets of images, one containing all condensates, and a second one in which cytosolic condensates were manually removed, based on their lack of overlap with DAPI-stained nuclear profiles, leaving only nuclear condensates. Using custom CellProfiler pipelines, we quantified both the total number and total area of condensates, as well as the number and area of nuclear condensates. From these parameters, for each field, we calculated the proportion of condensate area, or condensate number, found in the nuclear compartment. For each experimental group, we also analyzed the average area size

and eccentricity of condensates. In general, for each experimental group, we analyzed 10-20 microscopy fields from cell cultures derived from 3-6 transfections.

To define the sensitivity to 1,6-Hex of endogenous IRF2BPL, HEK293 cultures were exposed to 5% 1,6-Hex for 5 or 15 minutes[16,95]. Control experiments were performed using the same concentration of 2,5-Hexanediol which is known not to alter LLPS[45]. The cells were then rinsed with PBS and fixed in 4% PFA in PBS (15-20 min at RT) and processed for ICC using an anti-IRF2BPL primary antibody (Abcam; cat. # ab221099; lot # 1047602-7). Control cultures were treated in the same manner, except that 1,6-Hex was omitted. The proportion of the fluorescent area occupied by condensates and the number of condensates were measured as described above. The subcellular distribution of the protein was quantified by measuring the mean fluorescence intensity in nuclei (N), identified by DAPI staining, and in the cytoplasm (C), within a perinuclear ring of fixed thickness around the nuclei, using an ad hoc CellProfiler pipeline. The logarithm of the ratio between the two mean intensities (N/C) was used as a quantitative indicator of the subcellular distribution of the protein. Indeed, Log(N/C) is equal to 0 when the protein is evenly distributed between the nucleus and the cytoplasm and becomes greater or smaller than 0 when the protein is prevalently localized in the nucleus or the cytoplasm, respectively.

The same protocol was also used to quantify the sensitivity to 1,6-Hex and 2,5-Hex of exogenously expressed, GFP-tagged WT or mutant IRF2BPL. We also tested whether the 1,6-Hex effect on LLPS was reversible by adding an experimental group in which cells expressing IRF2BPL-GFP, before fixation and LLPS quantification, were incubated for an additional hour in normal culture medium after washing out 1,6-Hex.

To quantify acutely phototriggered LLPS through the optoDroplet system[46], HEK293 cells, grown onto poly-L-lysine coated glass coverslips and expressing either IRF2BPL fragments of interest fused to mCh-Cry2 or mCh-Cry2 alone as a control, were photoactivated 24-36 h after transfection. Photoactivation was obtained by illuminating the entire coverslip for 7 minutes using 488 nm blue light emitted by a led light array in a custom made apparatus, with a 10 cm working distance between the leds and the cell culture[16]. The intensity of the blue light, regulated through a potentiometer, was set at

a level that, in pilot experiments, was able to trigger the condensation of an LLPS-prone Cry2 mutant form (mCh-Cry2Olig; positive control) but not of mCh-Cry2 (negative control)[46]. Two minutes after photoactivation, the cells were rinsed in PBS and fixed with 4% PFA in PBS. Coverslips were mounted onto microscopy slides for confocal imaging to detect the mCherry fluorescence. Control cells were treated in the same manner, but photoactivation was omitted. LLPS-driven condensation was quantified, as described above, by calculating the relative proportion of the mCh-positive cell area occupied by condensates using an ad hoc CellProfiler pipeline.

**Correlative light and electron microscopy (CLEM)**

HEK293 cells were seeded on 35 mm dishes containing a gridded glass coverslip (MatTek) and transfected with plasmids to express IRF2BPL-GFP or GFP-tagged IRF2BPL disease-related mutants. Twenty-four hours after transfection, cells were fixed with 4% PFA, 1% glutaraldehyde in 0.1 M HEPES buffer (pH 7.4). To obtain additional references for a precise alignment of the fluorescence and EM images, cell cultures expressing IRF2BPL-GFP were also stained with MitoTracker Red CMXRos (200 nM, Thermo Fisher Scientific) and Hoechst 33342 (Thermo Fisher Scientific) to label mitochondria and nuclei, respectively. Confocal images of cells expressing IRF2BPL-GFP were acquired using a FluoVIEW 3000 RS system (Evident Scientific) equipped with a UPLSAPO 60X/1.3 Silicon objective. Z-stacks encompassing the entire cell were acquired with an optical section depth of 0.21 μm and the images were deconvolved using Huygens Software (v. 24.1; Scientific Volume Imaging; SVI) before alignment with the matching EM images. The cell cultures were then postfixed with reduced osmium (1% $OsO_4$, 1.5% potassium ferrocyanide in 0.1 M cacodylate buffer, pH 7.4) for 1 h on ice. After several washes in ultrapure water, they were incubated in 0.5% uranyl acetate overnight at 4°C. Samples were then dehydrated with increasing ethanol concentrations, embedded in epoxy resin, and polymerized for 48 h at 60 °C. Ultrathin serial sections (70 nm thickness) were obtained using an UC7 ultramicrotome (Leica), collected on formvar-carbon coated copper slot grids, stained with uranyl acetate and Sato's lead solutions, and observed under a Talos L120C transmission electron microscope

(FEI, ThermoFisher) operating at 120 kV. Images were acquired with a Ceta CCD camera (FEI, ThermoFisher) and aligned with confocal images using the Fiji macro-plugin Big Warp.

To highlight periodicities in the EM images, we performed a Fast Fourier Transform (FFT) analysis using Image-J (NIH). Each image was converted to 8-bit format and a corresponding Fourier frequency domain image was obtained using the 'FFT' command. Thresholding of the latter image was applied to highlight the main frequency components. Inverse Fast Fourier Transform (iFFT) processing was then performed using the iFFT command to obtain the final images (see **Fig. 6** and **Supplementary Fig. 8**) that were pseudo-colored using the 'Red Hot' Image-J lookup table.

**Fluorescence recovery after photobleaching (FRAP)**

To perform FRAP experiments, HEK293 cells were plated at low density in 8-well glass-bottom microslides (IBIDI) and transfected with plasmids to express WT IRF2BPL-GFP, IRF2BPL-Y173X-GFP, or IRF2BPL-Y123X-GFP. Live-cell FRAP experiments were performed ~24h after transfection, on a point scanning confocal microscope TCS SP8 (Leica) equipped with an HCX PL APO 63X/1.4 oil objective, an adaptive focus control, and a stage incubator (OkoLab) to maintain temperature at 37 °C in a 5% $CO_2$ atmosphere. Each individual cell nucleus was imaged for 5 iterations (zoom 5x, pinhole 5AU), then a WT or mutant (Y173X, Q123X) condensate within a 3 $\mu m^2$ circular region of interest (ROI) was bleached, with 100% laser power at 488 nm (Argon laser at 70% initial power, 1 iteration). We selected for these experiments, intranuclear condensates of similar, relatively small size for both the WT and mutant forms of IRF2BPL. For the mutant proteins, we selected the smaller condensates formed by these proteins early on after transfection, before massive condensate coalescence occurs. Indeed, we observed that FRAP did not occur in those few extremely large condensates formed by mutant IRF2BPL at later points after transfection, as no soluble, diffusible protein appeared to be present around the bleached condensates. After bleaching, images were taken every 2.3 sec for 2 min and every 15 sec for 4 min to monitor fluorescence recovery. In the images acquired for each condensate that was studied, before and after photobleaching, fluorescence intensities of the photobleached ROI (ROI1), of a second ROI encompassing the entire nucleus (ROI2), and a third ROI encompassing the

non-fluorescent background region (ROI3), were measured using the LasX software (Leica). Data were analyzed using the EasyFrap tool[96] (v. R2016a (9.0.1); available at https://easyfrap.vmnet.upatras.gr), which subtracts both background fluorescence (ROI3) and compensates for imaging-related fluorescence decay (ROI2). The final FRAP curve was obtained from EasyFRAP using double normalization and was fitted to a double exponential curve to calculate the mobile fraction ($M_f$) and the recovery half-time ($t_{1/2}$). The FRAP experiments were performed in parallel for WT and mutant (Y173X, Q123X) IRF2BPL, as reported in **Fig. 5e**. The curve relative to WT IRF2BPL from this graph is also shown by itself, for clarity, in **Fig. 2k**.

**Cycloheximide (CHX) chase experiment**

HEK293 cells, plated on 6-well plastic culture plates (VWR) and expressing GFP-tagged proteins of interest, or GFP alone, for 36 hours, were treated for 12 hours with either 0.05 μg/mL CHX (Merck), dissolved in DMSO, or with DMSO alone as a control. The cells were then harvested in ice-cold PBS. Whole-cell extracts were prepared in lysis buffer (2% w/v sodium dodecyl-sulfate (SDS) and 12% v/v β-mercaptoethanol) and separated on Nupage 3-8% Tris-acetate gels (Thermo Fisher Scientific). Immunoblotting was performed using anti-GFP HRP-conjugated primary antibodies (Proteintech, cat. # HRP-66002).



**In vitro LLPS assay**

To test the ability of IRF2BPL to undergo LLPS also in vitro, we attempted to express the protein in *E. coli* of two different strains (BL21 and Rosetta) and purify it. Bacterial cells were transformed with the pGAT3-IRF2BPL vector, transferred into LB medium, and used to inoculate variable volumes of Terrific Broth (TB) medium. Antibiotic selection was maintained during all steps. After reaching $OD_{600}$ = 0.6 at 37°C with shaking incubation (180 rounds per minute (RPM); 0.35 g), bacterial cultures were transferred to different incubating temperatures (16, 20, and 25 °C) and gene expression was induced by adding IPTG (0.1 or 1 mM). Protein purification was attempted under different conditions, using either one of the two available tags, i.e. GST or 6xHis, by means of GST-trap and His-trap columns

(Cytiva), respectively. For GST-trap purification, cells were lysed in a phosphate buffer (50 mM KPi pH 7.4, 200 mM NaCl) with 50 µM ZnSO4 to increase ZnF domain stability, resuspended in buffer supplemented with protease inhibitor cocktail (Roche), and sonicated. The soluble protein fraction was separated through centrifugation (100,000 g 4°C, 45 minutes) and loaded on a previously equilibrated GST-trap fast flow column (1 mL resin, Cytiva). Elution was performed with 50 mM KPi (pH 7.4), 200 mM NaCl, 50 µM ZnSO4, 20 mM reduced glutathione, and 1 mM DTT. For His-trap purification, cells were resuspended in 50 mM KPi (pH 7.4), 200 mM NaCl, 50 µM ZnSO4 supplemented with protease inhibitor cocktail and 15 mM imidazole. After sonication and centrifugation 100,000 g at 4°C for 45 min), the soluble fraction was loaded on a pre-equilibrated His-trap High Performance 5 mL column (Cytiva). The bound protein was eluted with 400 mM imidazole in the same buffer. His-trap purification was also attempted under denaturing conditions using urea as a chaotropic agent. Cells were resuspended in a buffer containing 50 mM TrisHCl (pH 7.5), 500 mM NaCl, 4 M urea, and 15 mM imidazole, and disrupted by stirring for 2 hours. After centrifugation to remove insoluble debris, the supernatant was loaded on a His-gravitrap column 1 mL resin (Cytiva). Elution was achieved by adding 500 mM imidazole to the buffer. All eluates were separated into fractions using size exclusion chromatography (SEC).

Under all purification and elution conditions that were tested, we were not able to reliably detect protein products of the expected molecular weight of tagged IRF2BPL upon electrophoretic separation of bacterial lysates, or eluted fractions, and subsequent gel staining to detect proteins. Consistent with the hypothesis that aggregation was responsible for the lack of penetration of the protein in a soluble form into the separating gels, we found, using direct immunoblotting of lysates and eluted fractions onto nitrocellulose (dot-blot; see below and **Supplementary Fig. 6**), that the protein was indeed expressed by bacteria and that smaller amounts of it were present in some eluted fractions. Light microscopy analyses of these fractions confirmed the presence of large amounts of aggregated material (**Supplementary Fig. 6**).

To overcome these issues, we sought to chemically synthesize the key LLPS-driving region of IRF2BPL, i.e. the ZnF domain, which displays a lower condensation tendency in the cellular context,

as compared to the full-length protein, and may therefore be more amenable to in vitro studies. A peptide encompassing the ZnF region (a.a. 1-75) of human IRF2BPL was chemically synthesized (LifeTein, LLC) with an N-terminal FITC tag linked through a 6-aminohexanoic acid (ahx) spacer, C-terminal amidation, and purity >90%. The peptide was dissolved in saline buffer (100 mM NaCl, 20 mM phosphate buffer, pH 7.4[12], with the addition of 0.1 mM $ZnCl_2$, to generate a 100 µM solution. The solution was spun to remove insoluble debris and 30 µL aliquots of the cleared peptide solution were dispensed into an 18-well glass-bottom plate (Ibidi) for imaging using the MICA imaging system (Leica). Imaging commenced a few minutes post-plating and continued up to 12-24 h. Images were acquired using a 63X oil-immersion objective, capturing both transmitted light DIC images, to visualize rounded droplet-like condensate structures, and green fluorescence from FITC to confirm that the observed condensates were formed by the peptide. In physiological buffer at room temperature, condensation of the peptide started to become visible after 1-2 h and further increased up to 24 h of incubation, similar to what previously found with other proteins in vitro[97].



**Dot blot**

Crude bacterial lysates and size-exclusion chromatography (SEC) fractions of protein eluates were directly applied onto a nitrocellulose membrane (1 µl per sample). The membrane was allowed to air dry for ~10 minutes at room temperature (RT), blocked in 2.5% non-fat dry milk in Tris-buffered saline with 0.1% Tween-20 (1×TBS-T) for 1 hour at RT with gentle agitation. After blocking, the membrane was washed three times for 10 min with 1×TBS-T with gentle shaking and then incubated overnight at 4°C with a primary antibody against IRF2BPL (Abcam, cat. # ab221099; 1:1000) in blocking solution (2.5% milk in 1× TBS-T). The membrane was washed three times with 1×TBS-T (10 min per wash) and then incubated for 1 hour at RT with an HRP-conjugated secondary antibody (Thermo Fisher Scientific, cat. # 31460; 1:3000) in blocking solution. After incubation, the membrane underwent three additional 10-min washes with 1×TBS-T. Membrane-bound secondary antibodies were detected using the Enhanced Chemiluminescence (ECL) reagent (Promega) and images were captured using the ChemiDoc imaging system (Bio-Rad).

**Generation of a human IRF2BPL knockout (KO) cell line**

To knockout the *IRF2BPL* gene in HEK293T cells, CRISPR RNA (crRNA) was designed using the IDT Alt-R CRISPR-Cas9 guide RNA design tool (Integrated DNA Technologies, available at https://eu.idtdna.com; AltR1/rCrC rUrCrG rUrArG rUrUrG rArCrG rCrArA rCrCrG rGrUrU rUrUrA rGrArG rCrUrA rUrGrC rU/AltR2). Chemically synthesized crRNA (2 nmol) and tracrRNA (5 nmol) were resuspended in IDTE buffer to a final concentration of 40 µM each. Equimolar volumes of crRNA and tracrRNA were then combined and annealed by incubation at 95 °C for 5 minutes and then gradually cooled to room temperature to form a duplex sgRNA complex. Annealing was performed on the same day as transfection to ensure RNA stability and optimal editing efficiency. HEK293T cells were then reverse-transfected with a Cas9-expressing plasmid (pX330-gRNA, #158973; Addgene) using Lipofectamine 3000 (Thermo Fisher Scientific), following the manufacturer's protocol, and seeded in six-well plates on the same day. Approximately 24 hours later, upon reaching ~80% confluency, cells were transfected with the pre-annealed crRNA:tracrRNA duplex using Lipofectamine RNAiMAX (Thermo Fisher Scientific), according to the manufacturer's instructions. A subset of transfected cells was lysed for western blot analysis to confirm IRF2BPL protein depletion. The cells were harvested in ice-cold PBS. Whole-cell extracts were prepared in RIPA buffer (10mM Tris-HCl, 150 mM NaCl, 0.5 mM EDTA, 0.1% SDS, 1% Triton, 1% Deoxycholate) and separated on Nupage 4-12% Bis-Tris gels (Thermo Fisher Scientific). Immunoblotting was performed using anti-IRF2BPL (ab221099, Abcam; 1:1000) and anti-GAPDH (Santa Cruz Biotechnology, cat. # sc32233; 1:1000) antibodies.

**Generation of a human cell line bearing a pathogenic Q127X mutation**

To generate a HEK293T mutant cell line bearing an IRF2BPL pathogenic mutation, we introduced the heterozygous IRF2BPL c.379C>T transition using CRISPR7Cas9 genome editing. This pathogenic nonsense mutation has been identified in patients (ClinVar NM_024496.4(IRF2BPL):c.379C>T) and leads to a Q127X protein truncation (cluster 1). Using the CRISPOR[98] online tool (v. 5.2; available at http://crispor.tefor.net), an oligo pair (produced by Sigma-Aldrich; Fwd: 5'-

CACCGCAACAGCTCAACCACGTTGA-3'; Rev :5'-AAACTCAACGTGGTTGAGCTGTTGC- 3') encoding the 20 nt guide sequences, and including overhangs for ligation into the BsmBI site, was identified to target the genomic region encompassing the IRF2BPL c.379C>T mutation site. The oligo pair was annealed and ligated into the LV-U6-filler-crRNA-EF1α-Blasticidin vector[99]. The pCAG-spCas9-P2A-Puromycin expression vector[99] was employed for transient expression of spCas9 in HEK293T cells. Additionally, a single-stranded oligodeoxynucleotide (ssODN) repair template (5'-AGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAAT AGCTCAATCACGTTGATGGTTCCAGCAAGCCTGCGGTGCTGGCGGCCCC-3') was designed to introduce the c.379C>T mutation along with a synonymous c.387C>T change, aimed at abolishing the BstXI restriction site, enabling downstream genotyping via restriction fragment length polymorphism (RFLP) analysis. HEK293T cells were seeded in six-well plates and transfected at >80% confluency with sgRNA-Blast and spCas9-Puro constructs and ssODN using Lipofectamine LTX (Thermo Fisher Scientific), according to the manufacturer's instructions. Following transfection, cells were subjected to dual selection with puromycin (1 μg/ml; Sigma-Aldrich) and blasticidin (10 μg/ml; Thermo Fisher Scientific). Successfully co-transfected cells were subjected to fluorescence-activated cell sorting (FACS) for single-cell clone isolation. The FACS sequential gating/sorting strategies are illustrated in **Supplementary Figure 13**. Genomic DNA (gDNA) of the clones was extracted by incubating HEK293T cells for 4 h at 55 °C in a digestion buffer mix (100 mM Tris–HCl pH 8.0, 200 mM NaCl, 5 mM EDTA, and 0.5% SDS, supplemented with proteinase K). Then, genomic DNA was purified using isopropanol, washed with 70% ethanol, and resuspended in nuclease-free water. The genomic regions flanking the mutation site were amplified by PCR using 100 ng of genomic DNA as a template and suitable primers (Fwd: 5'-GACTTCTCGGAACCCGTATG-3'; Rev: 5'-CGTCTGCGGTAGCAGGTT-3'). PCR products were analyzed on 1% agarose gels and purified (Wizard SV gel and PCR Clean-Up System, Promega). For RFLP genotyping, 300 ng of the purified PCR products were incubated with 10U of BstXI enzyme in NEB buffer r3.1 (New England Biolabs) overnight. Edited clones were confirmed by Sanger sequencing (Eurofins).

**RNA extraction and quantitative gene expression analysis**

WT and mutant (KO and Q127X) HEK293 WT cells were plated on 6-well plates and, 48h later, total RNA was isolated cells using TRIzol Reagent (Thermo Fisher Scientific), following the manufacturer's protocol. RNA was reverse transcribed using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara). WNT1 RNA levels were measured by quantitative real-time PCR (qPCR) using Luna Universal qPCR Master Mix (New England Biolabs) on a CFX96 Connect Thermal Cycler (Bio-Rad). Gene expression levels were normalized to GAPDH mRNA. The primers for WNT1 were: Fwd, 5'–CTGTCCTGCCTCCTCATC–3'; Rev, 5'-GGACCCAGCACAATAAATAGTT–3'. The primers for GAPDH were: Fwd, 5'-TTCGACAGTCAGCCGCATCTTCTT–3'; Rev, 5'– GCCCAATACGACCA AATCCGTTGA–3'.

**Electrophysiological recordings**

Whole-cell patch-clamp recordings were performed using an AxoClamp 2A amplifier and Clampex software (within the pClamp software suite; v. 11.1; Molecular Devices). Differentiated NG108-15 neuronal cells transfected with plasmids to express IRF2BPL-GFP, IRF2BPL-Y173X-GFP, IRF2BPL-Q123X-GFP, or GFP alone (used as a negative control) were recorded at room temperature. Cells were continuously perfused with an external solution containing 125 mM NaCl, 2.5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 1.25 mM $NaHCO_3$, and 10 mM glucose (pH 7.4). Patch pipettes (10–20 MΩ resistance) were filled with an internal solution containing 130 mM potassium gluconate, 2 mM $MgCl_2$, 0.5 mM EGTA, and 10 mM HEPES (pH 7.2). The osmolarity of these solutions was adjusted to 290 mOsm/l with sucrose. Recordings were acquired in current-clamp mode. The resting membrane potential (RMP) was measured immediately upon achieving the whole-cell configuration. To test cell excitability, cells were depolarized by delivering a series of 200-600 ms depolarizing pulses of increasing intensity, from 0 to +300 pA, in 50 pA steps which were delivered at 15-second intervals. Voltage responses were analyzed by measuring both the peak membrane potential, relative to the RMP, and the ΔV between the same peak and the steady-state potential (sp) reached in the remaining part of the depolarizing step. For

each cell, we assessed the depolarizing current threshold that triggered an active response (>5 mV). The proportion of cells exhibiting peak depolarizations ≥70 mV was also quantified. Input resistance was assessed by delivering 50 pA hyperpolarizing pulses (200 ms) and measuring the associated membrane potential change. Data were analyzed using Clampfit software (Molecular Devices).

**Software**

Images were processed and analyzed using Image-J/Fiji (v. 2.17.0) [94], CellProfiler[100] (v. 4.2.8; available at https://cellprofiler.org/), Photoshop Elements (v. 11; Adobe), and InkScape (v.1.4.2; available at https://inkscape.org). Protein structures reported in .pdb file format (derived either from AlphaFold or the Protein Data Bank (PDB), available at www.rcsb.org) were visualized and analyzed using Chimera software (version 1.14)[74]. Coiled-coil structure predictions were obtained using Paircoil2[70] (available at https://cb.csail.mit.edu/paircoil2). Shannon entropy in primary sequence alignments of IRF2BPL orthologs was calculated using EntropyOne software (HIV Sequence Database), available at https://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy_one.html/. Per-residue LLPS propensity scores for protein primary sequences were obtained using FuzDrop[40] (available at https://fuzdrop.bio.unipd.it/predictor). SIM and REP scores were calculated using software developed by Vaglietti et al. (2025)[30]. The DNA sequencing chromatogram shown in **Fig. 8a** was generated using Chromas software (v. 2.6.6; Technelysium) and graphically modified using Photoshop Elements 11. Data were processed and analyzed using Excel (as part of the Office 2024 suite; Microsoft Corporation) and Prism (v.10; GraphPad). Statistical analyses were performed using Excel and Statistica (v.13.6.0; TIBCO).

**Statistics and reproducibility**

All experiments, including those for which representative images are shown in **Figs. 1f**, **2a,b,c,e,f,h**, **3g**, **4b**, **7d,g**, **8b,e**, and **10c**, as well as in **Supplementary Figs. 5a,b**, **6a,b**, **7a,c**, **8a,b**, **9a,b, and 12a,** have been performed at least in three independent replicates with similar results. Student's t-test, one- or two-way analysis of variance (ANOVA) followed by Newman-Keuls post hoc test, which includes

a multiple comparison correction, and Fisher's exact test were performed, where appropriate, as indicated in the Results section and the figure legends. All tests were two-sided, when possible. A value of $p \leq 0.05$ was considered statistically significant in all instances.

## DATA AVAILABILITY

Unless otherwise stated, the data generated in this study are provided in the article, supplementary, and source data file. Source data are provided with this paper.

The Uniprot identifiers (IDs) of the primary sequences of IRF2BPL and its paralogs IRF2BP1 and IRF2BP2 are, respectively:

Q9H1B7 [https://www.uniprot.org/uniprotkb/Q9H1B7/entry]

Q8IU81 [https://www.uniprot.org/uniprotkb/Q8IU81/entry]

Q7Z5L9 [https://www.uniprot.org/uniprotkb/Q7Z5L9/entry].

The ClinVar database ID of the IRF2BPL-Q116X mutation is:

1801711 [https://www.ncbi.nlm.nih.gov/clinvar/variation/1801711/?oq=1801711&m=NM_024496.4(IRF2BPL):c.346C%3ET%20(p.Gln116Ter))].

The Protein Data Bank [PDB; https://www.rcsb.org/] IDs of solved structures of proteins of interest for the DALI structural analyses are reported in **Supplementary Data 7**, together with the corresponding primary sequence IDs of the Uniprot database [https://www.uniprot.org].

The Leiden Open Variation Database [LOVD, version 3.0; https://www.lovd.nl/] IDs of pathogenic mutations of interest are listed in **Supplementary Data 5**.

The National Center for Biotechnology Information (NCBI) database [https://www.ncbi.nlm.nih.gov/protein/] IDs for the primary sequences of vertebrate IRF2BPL orthologs of interest are listed in **Supplementary Data 4**.

## REFERENCES

1. Marcogliese, P. C., Shashi, V., Spillmann, R. C., Stong, N., Rosenfeld, J. A., Koenig, M. K., Martínez-Agosto, J. A., Herzog, M., Chen, A. H., Dickson, P. I., Lin, H. J., Vera, M. U., Salamon, N., Graham, J. M., Ortiz, D., Infante, E., Steyaert, W., Dermaut, B., Poppe, B., … Zheng, A. (2018). IRF2BPL Is Associated with Neurological Phenotypes. The American Journal of Human Genetics, 103(2), 245–260.

2. Tran Mau-Them, F., Guibaud, L., Duplomb, L., Keren, B., Lindstrom, K., Marey, I., Mochel, F., Van Den Boogaard, M. J., Oegema, R., Nava, C., Masurel, A., Jouan, T., Jansen, F. E., Au, M., Chen, A. H., Cho, M., Duffourd, Y., Lozier, E., Konovalov, F., … Vitobello, A. (2019). De novo truncating variants in the intronless IRF2BPL are responsible for developmental epileptic encephalopathy. Genetics in Medicine, 21(4), 1008–1014.

3. Prilop, L., Buchert, R., Woerz, S., Gerloff, C., Haack, T. B., & Zittel, S. (2020). IRF2BPL mutation causes nigrostriatal degeneration presenting with dystonia, spasticity and keratoconus. Parkinsonism & related disorders, 79, 141-143.

4. Costa, C., Oliver, K. L., Calvello, C., Cameron, J. M., Imperatore, V., Tonelli, L., ... & Prontera, P. (2023). IRF2BPL: a new genotype for progressive myoclonus epilepsies. Epilepsia, 64(8), e164-e169.

5. Gardella, E., Michelucci, R., Christensen, H. M., Fenger, C. D., Reale, C., Riguzzi, P., ... & Rubboli, G. (2023). IRF2BPL as a novel causative gene for progressive myoclonus epilepsy. Epilepsia, 64(8), e170-e176.

6. Kelleher, K. J., Sheils, T. K., Mathias, S. L., Yang, J. J., Metzger, V. T., Siramshetty, V. B., Nguyen, D.-T., Jensen, L. J., Vidović, D., Schürer, S. C., Holmes, J., Sharma, K. R., Pillai, A., Bologa, C. G., Edwards, J. S., Mathé, E. A., & Oprea, T. I. (2023). Pharos 2023: An integrated resource for the understudied human proteome. Nucleic Acids Research, 51(D1), D1405–D1416.

7. Heger, S., Mastronardi, C., Dissen, G. A., Lomniczi, A., Cabrera, R., Roth, C. L., Jung, H., Galimi, F., Sippell, W., & Ojeda, S. R. (2007). Enhanced at puberty 1 (EAP1) is a new transcriptional regulator of the female neuroendocrine reproductive axis. Journal of Clinical Investigation, 117(8), 2145–2154.

8. Marcogliese, P. C., Dutta, D., Ray, S. S., Dang, N. D. P., Zuo, Z., Wang, Y., Lu, D., Fazal, F., Ravenscroft, T. A., Chung, H., Kanca, O., Wan, J., Douine, E. D., Network, U. D., Pena, L. D. M., Yamamoto, S., Nelson, S. F., Might, M., Meyer, K. C., … Bellen, H. J. (2022). Loss of IRF2BPL impairs neuronal maintenance through excess Wnt signaling. Science Advances, 8(3), eabl5613.

9. Yokoyama, A., Kouketsu, T., Otsubo, Y., Noro, E., Sawatsubashi, S., Shima, H., Satoh, I., Kawamura, S., Suzuki, T., Igarashi, K., & Sugawara, A. (2021). Identification and Functional Characterization of a Novel Androgen Receptor Coregulator, EAP1. Journal of the Endocrine Society, 5(11), bvab150.

10. Mancini, A., Howard, S. R., Cabrera, C. P., Barnes, M. R., David, A., Wehkalampi, K., ... & Dunkel, L. (2019). EAP1 regulation of GnRH promoter activity is important for human pubertal timing. Human molecular genetics, 28(8), 1357-1368.

11. Rampazzo, A., Pivotto, F., Occhi, G., Tiso, N., Bortoluzzi, S., Rowen, L., Hood, L., Nava, A., & Antonio Danieli, G. (2000). Characterization of C14orf4, a Novel Intronless Human Gene Containing a Polyglutamine Repeat, Mapped to the ARVD1 Critical Region. Biochemical and Biophysical Research Communications, 278(3), 766–774.

12. Fiumara, F., Fioriti, L., Kandel, E. R., & Hendrickson, W. A. (2010). Essential Role of Coiled Coils for Aggregation and Activity of Q/N-Rich Prions and PolyQ Proteins. Cell, 143(7), 1121–1135.

13. Pelassa, I., Corà, D., Cesano, F., Monje, F. J., Montarolo, P. G., & Fiumara, F. (2014). Association of polyalanine and polyglutamine coiled coils mediates expansion disease-related protein aggregation and dysfunction. Human Molecular Genetics, 23(13), 3402–3420.

14. Hosp, F., Gutiérrez-Ángel, S., Schaefer, M. H., Cox, J., Meissner, F., Hipp, M. S., Hartl, F.-U., Klein, R., Dudanova, I., & Mann, M. (2017). Spatiotemporal Proteomic Profiling of Huntington's Disease Inclusions Reveals Widespread Loss of Protein Function. Cell Reports, 21(8), 2291–2303.

15. Kwon, M. J., Han, M. H., Bagley, J. A., Hyeon, D. Y., Ko, B. S., Lee, Y. M., ... & Jan, Y. N. (2018). Coiled-coil structure-dependent interactions between polyQ proteins and Foxo lead to dendrite pathology and behavioral defects. Proceedings of the National Academy of Sciences, 115(45), E10748-E10757.

16. Vaglietti, S., Villeri, V., Dell'Oca, M., Marchetti, C., Cesano, F., Rizzo, F., Miller, D., LaPierre, L., Pelassa, I., Monje, F. J., Colnaghi, L., Ghirardi, M., & Fiumara, F. (2023). PolyQ length-based molecular encoding of vocalization frequency in FOXP2. iScience, 26(10), 108036.

17. Almeida, B., Fernandes, S., Abreu, I. A., & Macedo-Ribeiro, S. (2013). Trinucleotide Repeats: A Structural Perspective. Frontiers in Neurology, 4.

18. Fang, X., Wang, L., Ishikawa, R., Li, X., Fiedler, M., Liu, F., ... & Dean, C. (2019). The Arabidopsis coiled coil protein FLL2 promotes formation of liquid droplets to enhance polyadenylation.

19. Langdon, E. M., Qiu, Y., Ghanbari Niaki, A., McLaughlin, G. A., Weidmann, C. A., Gerbich, T. M., ... & Gladfelter, A. S. (2018). mRNA structure determines specificity of a polyQ-driven phase separation. Science, 360(6391), 922-927.

20. Basu, S., Mackowiak, S. D., Niskanen, H., Knezevic, D., Asimi, V., Grosswendt, S., Geertsema, H., Ali, S., Jerković, I., Ewers, H., Mundlos, S., Meissner, A., Ibrahim, D. M., & Hnisz, D. (2020). Unblending of Transcriptional Condensates in Human Repeat Expansion Disease. Cell, 181(5), 1062-1079.e30.

21. Chen, T. C., & Huang, J. R. (2020). Musashi-1: an example of how polyalanine tracts contribute to self-association in the intrinsically disordered regions of RNA-binding proteins. International journal of molecular sciences, 21(7), 2289.

22. Alberti, S., Gladfelter, A., & Mittag, T. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell, 176(3), 419–434.

23. Peskett, T. R., Rau, F., O'Driscoll, J., Patani, R., Lowe, A. R., & Saibil, H. R. (2018). A Liquid to Solid Phase Transition Underlying Pathological Huntingtin Exon1 Aggregation. Molecular Cell, 70(4), 588-601.e6.

24. Alberti, S., & Dormann, D. (2019). Liquid–liquid phase separation in disease. Annual review of genetics, 53, 171-194.

25. Chen, Y., Wan, Y., Pei, X., Wang, T., Ma, Z., & Chen, L. (2023). GATA3 ZnF2-defective mutant condensation underlies type I IFN-activating in breast cancer. bioRxiv, 2023-05.

26. Horovitz, D. D. G., de Faria Domingues de Lima, M. A., Pires, L. D. C., Campos Araujo, A. D. Q., & Vargas, F. R. (2023). Neurological Phenotypes of IRF2BPL Gene Variants: A Report of Four Novel Variants. Journal of Central Nervous System Disease, 15, 11795735231181467.

27. Sinha Ray, S., Dutta, D., Dennys, C., Powers, S., Roussel, F., Lisowski, P., Glažar, P., Zhang, X., Biswas, P., Caporale, J. R., Rajewsky, N., Bickle, M., Wein, N., Bellen, H. J., Likhite, S., Marcogliese, P. C., & Meyer, K. C. (2022). Mechanisms of IRF2BPL-related disorders and identification of a potential therapeutic strategy. Cell Reports, 41(10), 111751.

28. Wang, Y., Ke, Z., Li, Y., Qiu, M., Liu, J., Yang, Z., ... & Chen, S. (2024). De novo variants of IRF2BPL result in developmental epileptic disorder. Orphanet Journal of Rare Diseases, 19(1), 121.

29. Wootton, J. C., & Federhen, S. (1993). Statistics of local complexity in amino acid sequences and sequence databases. Computers & chemistry, 17(2), 149-163.

30. Vaglietti S, Boggio Bozzo S, Ghirardi M, Fiumara F. (2025) Divergent evolution of low-complexity regions in the vertebrate CPEB protein family. Front Bioinform. 5:1491735.

31. Pelassa I, Cibelli M, Villeri V, Lilliu E, Vaglietti S, Olocco F, Ghirardi M, Montarolo PG, Corà D, Fiumara F. (2019) Compound Dynamics and Combinatorial Patterns of Amino Acid Repeats Encode a System of Evolutionary and Developmental Markers. Genome Biol Evol. 11(11):3159-3178.

32. Wiel L, Baakman C, Gilissen D, Veltman JA, Vriend G, Gilissen C. (2019) MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. Hum Mutat. 40(8):1030-1038.

33. Cheng J, Novati G, Pan J, Bycroft C, Žemgulytė A, Applebaum T, Pritzel A, Wong LH, Zielinski M, Sargeant T, Schneider RG, Senior AW, Jumper J, Hassabis D, Kohli P, Avsec Ž. (2023) Accurate proteome-wide missense variant effect prediction with AlphaMissense. Science. 381(6664):eadg7492.

34. Tordai, H., Torres, O., Csepi, M., Padanyi, R., Lukács, G. L., & Hegedűs, T. (2024). Analysis of AlphaMissense data in different protein groups and structural context. Scientific data, 11(1), 495.

35. Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam HJ, Mort M, Cooper DN, Sebat J, Iakoucheva LM, Mooney SD, Radivojac P. (2020) Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nat Commun. 11(1):5918.

36. Mirzaa GM, Chong JX, Piton A, Popp B, Foss K, Guo H, Harripaul R, Xia K, Scheck J, Aldinger KA, Sajan SA, Tang S, Bonneau D, Beck A, White J, Mahida S, Harris J, Smith-Hicks C, Hoyer J, … Bamshad MJ. (2020) De novo and inherited variants in ZNF292 underlie a neurodevelopmental disorder with features of autism spectrum disorder. Genet Med. 22(3):538-546.

37. Walter, J., Charon, J., Hu, Y., Lachat, J., Leger, T., Lafforgue, G., ... & Michon, T. (2019). Comparative analysis of mutational robustness of the intrinsically disordered viral protein VPg and of its interactor eIF4E. PloS one, 14(2), e0211725.

38. Bayrak, C. S., Stein, D., Jain, A., Chaudhary, K., Nadkarni, G. N., Van Vleck, T. T., ... & Itan, Y. (2021). Identification of discriminative gene-level and protein-level features associated with pathogenic gain-of-function and loss-of-function variants. The American Journal of Human Genetics, 108(12), 2301-2318.

39. Vacic V, Markwick PR, Oldfield CJ, Zhao X, Haynes C, Uversky VN, Iakoucheva LM. (2012) Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. PLoS Comput Biol. 2012;8(10):e1002709.

40. Hardenberg M, Horvath A, Ambrus V, Fuxreiter M, Vendruscolo M. Widespread occurrence of the droplet state of proteins in the human proteome. PNAS (2020);117(52):33254-33262.

41. Nirenberg, M., Wilson, S., Higashida, H., Rotter, A., Krueger, K., Busis, N., ... & Adler, M. (1983). Modulation of synapse formation by cyclic adenosine monophosphate. Science, 222(4625), 794-799.

42. Han, H. Q., Nichols, R. A., Rubin, M. R., Bähler, M., & Greengard, P. (1991). Induction of formation of presynaptic terminals in neuroblastoma cells by synapsin IIb. Nature, 349(6311), 697-700

43. Zhong, Z. G., Noda, M., Takahashi, H., & Higashida, H. (1999). Overexpression of rat synapsins in NG108-15 neuronal cells enhances functional synapse formation with myotubes. Neuroscience letters, 260(2), 93-96.

44. Rebane, A. A., Ziltener, P., LaMonica, L. C., Bauer, A. H., Zheng, H., López-Montero, I., ... & Ernst, A. M. (2020). Liquid–liquid phase separation of the Golgi matrix protein GM130. FEBS letters, 594(7), 1132-1144.

45. Lin Y, Mori E, Kato M, Xiang S, Wu L, Kwon I, McKnight SL. (2016) Toxic PR Poly- Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. Cell. 167(3):789-802.e12.

46. Shin, Y., Berry, J., Pannucci, N., Haataja, M. P., Toettcher, J. E., & Brangwynne, C. P. (2017). Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. Cell, 168(1–2), 159-171.e14.

47. Perutz MF, Johnson T, Suzuki M, Finch JT. (1994) Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. Proc Natl Acad Sci U S A. 91(12):5355-8.

48. Temussi PA, Masino L, Pastore A. (2003) From Alzheimer to Huntington: why is a structural understanding so difficult? EMBO J. 22(3):355-61.

49. Nanajkar N, Sahoo A, Matysiak S. (2024) Unraveling the Molecular Complexity of N-Terminus Huntingtin Oligomers: Insights into Polymorphic Structures. J Phys Chem B. 128(32):7761-7769.

50. Tsubotani K, Maeyama S, Murakami S, Schaffer SW, Ito T. (2021) Taurine suppresses liquid-liquid phase separation of lysozyme protein. Amino Acids. 53(5):745-751.

51. Lilliu, E., Villeri, V., Pelassa, I., Cesano, F., Scarano, D., & Fiumara, F. (2018). Polyserine repeats promote coiled coil-mediated fibril formation and length-dependent protein aggregation. Journal of Structural Biology, 204(3), 572–584.

52. Pelassa, I., & Fiumara, F. (2015). Differential Occurrence of Interactions and Interaction Domains in Proteins Containing Homopolymeric Amino Acid Repeats. Frontiers in Genetics, 6.

53. Mienville JM, Pesold C. (1999) Low resting potential and postnatal upregulation of NMDA receptors may cause Cajal-Retzius cell death. J Neurosci. 19(5):1636-46.

54. Li, S., Weidenfeld, J., & Morrisey, E. E. (2004). Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. Molecular and cellular biology, 24(2), 809-822.

55. Havrilla, J. M., Pedersen, B. S., Layer, R. M., & Quinlan, A. R. (2019). A map of constrained coding regions in the human genome. Nature genetics, 51(1), 88-95.

56. Laget, S., Joulie, M., Le Masson, F., Sasai, N., Christians, E., Pradhan, S., ... & Defossez, P. A. (2010). The human proteins MBD5 and MBD6 associate with heterochromatin but they do not bind methylated DNA. PloS one, 5(8), e11982.

57. Tam, S., Spiess, C., Auyeung, W., Joachimiak, L., Chen, B., Poirier, M. A., & Frydman, J. (2009). The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation. Nature structural & molecular biology, 16(12), 1279-1285.

58. Bhattacharyya, A., Thakur, A. K., Chellgren, V. M., Thiagarajan, G., Williams, A. D., Chellgren, B. W., ... & Wetzel, R. (2006). Oligoproline effects on polyglutamine conformation and aggregation. Journal of molecular biology, 355(3), 524-535.

59. Holehouse AS, Alberti S. (2025) Molecular determinants of condensate composition. Mol Cell. 85:290-308.

60. Hedtfeld M, Dammers A, Koerner C, Musacchio A. A validation strategy to assess the role of phase separation as a determinant of macromolecular localization. (2024) Mol Cell. 84:1783-1801.e7.

61. Poudyal M, Patel K, Gadhe L, Sawner AS, Kadu P, Datta D, Mukherjee S, Ray S, Navalkar A, Maiti S, Chatterjee D, Devi J, Bera R, Gahlot N, Joseph J, Padinhateeri R, Maji SK. (2023) Intermolecular interactions underlie protein/peptide phase separation irrespective of sequence and structure at crowded milieu. Nat Commun. 14:6199.

62. Patel CK, Singh S, Saini B, Mukherjee TK. (2022) Macromolecular Crowding-Induced Unusual Liquid-Liquid Phase Separation of Human Serum Albumin via Soft Protein- Protein Interactions. J Phys Chem Lett. 13:3636-3644.

63. Chen C, Guo C, Fang K, Lin C, Luo Z. (2025) RNF219 RING Finger Domain Mutants Drive Phase Separation to Encapsulate CCR4-NOT and Promote Cell Proliferation. Cell Prolif. e70072.

64. Yeung, K. T., Das, S., Zhang, J., Lomniczi, A., Ojeda, S. R., Xu, C. F., ... & Samuels, H. H. (2011). A novel transcription complex that selectively modulates apoptosis of breast cancer cells through regulation of FASTKD2. Molecular and cellular biology, 31(11), 2287-2298.

65. Vaglietti, S., & Fiumara, F. (2021). PolyQ length co-evolution in neural proteins. NAR Genomics and Bioinformatics, 3(2), lqab032.

66. Ginevrino, M., Battini, R., Nuovo, S., Simonati, A., Micalizzi, A., Contaldo, I., Serpieri, V., & Valente, E. M. (2020). A novel IRF2BPL truncating variant is associated with endolysosomal storage. Molecular Biology Reports, 47(1), 711–714.

67. Antonelli, F., Grieco, G., Cavallieri, F., Casella, A., & Valente, E. M. (2022). Adult onset familiar dystonia-plus syndrome: A novel presentation of IRF2BPL-associated neurodegeneration. Parkinsonism & Related Disorders, 94, 22-24.

68. Iannielli, A., Luoni, M., Giannelli, S. G., Ferese, R., Ordazzo, G., Fossati, M., ... & Broccoli, V. (2022). Modeling native and seeded Synuclein aggregation and related cellular dysfunctions in dopaminergic neurons derived by a new set of isogenic iPSC lines with SNCA multiplications. Cell Death & Disease, 13(10), 881.

69. Jarnot, P., Ziemska-Legiecka, J., Dobson, L., Merski, M., Mier, P., Andrade-Navarro, M. A., ... & Gruca, A. (2020). PlaToLoCo: the first web meta-server for visualization and annotation of low complexity regions in proteins. Nucleic acids research, 48(W1), W77-W84.

70. McDonnell, A. V., Jiang, T., Keating, A. E., & Berger, B. (2006). Paircoil2: improved prediction of coiled coils from sequence. Bioinformatics, 22(3), 356-358.

71. Vendruscolo, M., & Fuxreiter, M. (2022). Sequence determinants of the aggregation of proteins within condensates generated by liquid-liquid phase separation. Journal of Molecular Biology, 434(1), 167201.

72. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. nature, 596(7873), 583-589.

73. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. (2022) ColabFold: making protein folding accessible to all. Nat Methods. 19(6):679-682.

74. Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C., & Ferrin, T. E. (2006). Tools for integrated sequence-structure analysis with UCSF Chimera. BMC bioinformatics, 7, 1-10.

75. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; Protein Identification and Analysis Tools on the Expasy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607

76. Papadopoulos, J. S., & Agarwala, R. (2007). COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics, 23(9), 1073-1079.

77. Hedges, S. B., Dudley, J., and Kumar, S. (2006). TimeTree: a public knowledgebase of divergence times among organisms. Bioinformatics 22, 2971–2972.

78. Qian, X. H., Liu, X. Y., Zhu, Z. Y., Wang, S. G., Song, X. X., Chen, G., ... & Cao, L. (2021). Neurodevelopmental disorder caused by a truncating de novo variant of IRF2BPL. Seizure, 84, 47-52

79. Fokkema, I. F., Kroon, M., López Hernández, J. A., Asscheman, D., Lugtenburg, I., Hoogenboom, J., & den Dunnen, J. T. (2021). The LOVD3 platform: efficient genome-wide sharing of genetic variants. European Journal of Human Genetics, 29(12), 1796-1803.

80. Marchetti, C., Vaglietti, S., Rizzo, F., Di Nardo, G., Colnaghi, L., Ghirardi, M., & Fiumara, F. (2021). Heptad stereotypy, S/Q layering, and remote origin of the SARS-CoV-2 fusion core. Virus evolution, 7(2), veab097.

81. Pagel KA, Pejaver V, Lin GN, Nam HJ, Mort M, Cooper DN, Sebat J, Iakoucheva LM, Mooney SD, Radivojac P. (2017) When loss-of-function is loss of function: assessing mutational signatures and impact of loss-of-function genetic variants. Bioinformatics. 33(14):i389-i398.

82. Pagel KA, Antaki D, Lian A, Mort M, Cooper DN, Sebat J, Iakoucheva LM, Mooney SD, Radivojac P. (2019) Pathogenicity and functional impact of non-frameshifting insertion/deletion variation in the human genome. PLoS Comput Biol. 15(6):e1007112.

83. Ren G, Gu X, Zhang L, Gong S, Song S, Chen S, Chen Z, Wang X, Li Z, Zhou Y, Li L, Yang J, Lai F, Dang Y. (2024) Ribosomal frameshifting at normal codon repeats recodes functional chimeric proteins in human. Nucleic Acids Res. 52(5):2463-2479.

84. Toledano I, Supek F, Lehner B. (2024) Genome-scale quantification and prediction of pathogenic stop codon readthrough by small molecules. Nat Genet. 56(9):1914-1924.

85. Holm, L., & Laakso, L. M. (2016). Dali server update. Nucleic acids research, 44(W1), W351-W355.

86. Yang M., Li X., Tian Z., Ma L., Ma J., Liu Y., Shang G., Liang A., Wu W., Chen Z. (2023) Structures of MPND Reveal the Molecular Recognition of Nucleosomes. Int J Mol Sci. 24(4):3368.

87. Wakefield, R. I., Smith, B. O., Nan, X., Free, A., Soteriou, A., Uhrin, D., ... & Barlow, P. N. (1999). The solution structure of the domain from MeCP2 that binds to methylated DNA. Journal of molecular biology, 291(5), 1055-1065.

88. Hung, M. S., & Shen, C. K. J. (2003). Eukaryotic methyl-CpG-binding domain proteins and chromatin modification. Eukaryotic cell, 2(5), 841-846.

89. Ohki, I., Shimotake, N., Fujita, N., Jee, J. G., Ikegami, T., Nakao, M., & Shirakawa, M. (2001). Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. Cell, 105(4), 487-497.

90. Sreerama, N., & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Analytical biochemistry, 287(2), 252-260.

91. Miles AJ, Ramalli SG, Wallace BA. (2022) DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. Protein Sci. 31(1):37-46.

92. Mitchell, P. J., Hanson, J. C., Quets-Nguyen, A. T., Bergeron, M., & Smith, R. C. (2007). A quantitative method for analysis of in vitro neurite outgrowth. Journal of neuroscience methods, 164(2), 350-362.

93. Bianchini, P., & Diaspro, A. (2012). Fast scanning STED and two-photon fluorescence excitation microscopy with continuous wave beam. Journal of Microscopy, 245(3), 225-228.

94. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., ... & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. Nature methods, 9(7), 676-682.

95. Ulianov, S. V., Velichko, A. K., Magnitov, M. D., Luzhin, A. V., Golov, A. K., Ovsyannikova, N., Kireev, I. I., Gavrikov, A. S., Mishin, A. S., Garaev, A. K., Tyakht, A. V., Gavrilov, A. A., Kantidze, O. L., & Razin, S. V. (2021). Suppression of liquid–liquid phase separation by 1,6-hexanediol partially compromises the 3D genome organization in living cells. Nucleic Acids Research, 49(18), 10524–10541.

96. Koulouras, G., Panagopoulos, A., Rapsomaniki, M. A., Giakoumakis, N. N., Taraviras, S., & Lygerou, Z. (2018). EasyFRAP-web: a web-based tool for the analysis of fluorescence recovery after photobleaching data. Nucleic acids research, 46(W1), W467-W472.

97. Kanaan NM, Hamel C, Grabinski T, Combs B. (2020) Liquid-liquid phase separation induces pathogenic tau conformations in vitro. Nat Commun. (2020) 11(1):2809.

98. Concordet, J. P., & Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic acids research, 46(W1), W242-W245.

99. Giannelli SG, Luoni M, Castoldi V, Massimino L, Cabassi T, Angeloni D, Demontis GC, Leocani L, Andreazzoli M, Broccoli V. (2018) Cas9/sgRNA selective targeting of the P23H Rhodopsin mutant allele for treating retinitis pigmentosa by intravitreal AAV9.PHP.B-based delivery. Hum Mol Genet. 27(5):761-779.

100. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A (2021). CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics, 22 (1), 433. . PMID: 34507520 PMCID: PMC8431850.

101. Lei, M., Tempel, W., Chen, S., Liu, K., & Min, J. (2019). Plasticity at the DNA recognition site of the MeCP2 mCG-binding domain. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1862(9), 194409.

**ACKNOWLEDGEMENTS**

This work was supported by Fondazione Cariplo-Fondazione Telethon (Grant GJC22078 to F.F.), the Italian Ministry for University and Research (MIUR; Grant Prin-2022 20228BS45E to L.C.), and the University of Turin (RiLo-22/RiLo-23 to F.F). We thank ALEMBIC, the Advanced Light and Electron Microscopy Bioimaging Center of the IRCCS San Raffaele Scientific Institute, for providing the facilities for performing part of the microscopy experiments. Support from the *Rita Levi Montalcini* Department of Neuroscience and the Department of Chemistry (Project CH4.0; CUP: D13C22003520001) of the University of Turin under the MUR program "Dipartimenti di Eccellenza 2023–2027" is also acknowledged.



**AUTHOR CONTRIBUTIONS**

FF and LC conceived the study and designed the experiments; SV and FF performed bioinformatics analyses and structural predictions; AI produced human neural progenitor cell (NPC)-derived neuron cultures; PB, VB, LC performed STED super resolution microscopy experiments; MD and SBB performed Airyscan super resolution microscopy experiments; MD, SBB, CM, AC, SRa, performed confocal microscopy experiments; MN, LC, VB, and AL performed CLEM experiments; LC, VB, AL performed FRAP experiments; MD and SBB performed 1,6-Hex assays; MD, SBB, and CM perfomed molecular cloning experiments; RP, MD, FF, and SRoa designed and built an experimental apparatus for optoDroplet experiments; MD, SBB, and CM performed optoDroplet assays; SBB performed in vitro LLPS assays; SBB performed live-cell condensate coalescence dynamics experiments; FC and FF performed FFT analyses on EM images; PM, and LC performed cycloheximide pulse-chase assays; MD performed experiments of IRF2BPL recruitment; RI, CDL, GF, SS, ML performed CRISPR Cas9 gene editing experiments; PM, LC, RI, CDL, GF performed western blot experiments; DR performed qPCR experiments; MD and FJM performed electrophysiological experiments; SBB and FF performed circular dichroism; SBB, IC, AG, FS, DB, SRos, GDN performed protein expression and purification

from *E. coli*; SBB performed dot blot experiments; EA and MC advised on the design of qPCR primers; MD, SBB, SV, FF, LC, PB, FC, FJM analyzed the data; MD, SBB, SV, FF, LC, FC prepared figures; FF and LC wrote the original manuscript draft; all authors reviewed the manuscript; FF, LC, MG, SRoa provided supervision and project oversight; FF and LC administered the project; FF and LC provided key resources and funding; FF and LC supervised this work and contributed equally as senior authors.

**COMPETING INTERESTS**

The authors have no competing interests to declare.

## FIGURE LEGENDS

**Figure 1. The IRF2BPL central region contains four heterogeneous parts**

**a.** Known IRF2BPL domain structure with ZnF and RING domains, the poly-A (A) and -Q (Q), PEST, and NLS motifs in shades of gray. Colored bar: revised domain structure with LCRs (1-3; orange, yellow, and cyan) and the HCD (green). Lollipop plots: disease-related truncating (lower) and non-truncating (upper) mutations. Alpha-helices and β-strands predicted by AlphaFold are reported below (α/β). $LCR_1$ and HCD contain CC-prone α-helices ($CC_{1/2}$) predicted by Paircoil2. Below: LLPS propensity graph reporting FuzDrop per-residue pDP scores, identifying four major (1-4; pDP ≥ 0.60, red line) and three smaller LLPS-prone regions (asterisks). Below, entropy plot (Ent) along a sequence alignment of 456 IRF2BPL orthologs (gaps not represented). Below, plot of MetaDome per-residue tolerance scores to replacement (color scale from red (intolerant) to blue (tolerant)). Bottom, upper heatmap: AlphaMissense (αM) pathogenicity scores for amino acid replacements at each position (increasing pathogenicity from blue to red). Lollipop plot: position of known pathogenic substitutions. Lower plot: AlphaMissense overall per-residue pathogenicity scores. Color scale from blue (benign) to red (pathogenic). **b.** AF IRF2BPL model (domain colors as *a*). **c.** Pie charts highlighting the biased composition of LCRs (Q+A: black, P+Q+S: dark gray, other amino acids: light gray). **d.** Mean per-position entropy for each domain in an alignment of 456 IRF2BPL orthologs. Box plots report mean (cross), median (line), 25-75 percentile (limits), and 10-90 percentile (whiskers). One-way ANOVA: $F_{(5,1115)}=64.031$, $p<0.0001$, n= 88 (ZnF), 133 ($LCR_1$), 231 ($LCR_2$), 204 (HCD), 357 ($LCR_3$), and 108 (RING) alignment positions. Asterisks indicate significant differences vs. non-LCR domains ($p<0.05$, NK test). **e.** Significantly different distribution of truncating (n=31) and non-truncating (n=8) mutations across LCRs (red) and non-LCR domains (gray). ($p = 0.0098$, Fisher's exact (FE) test). **f.** Confocal and STED images of NPC-derived human neurons after immunocitochemistry (ICC) with α-IRF2BPL and α-MAP2 antibodies (65 DIV). Left and middle panels: IRF2BPL (green) and MAP2 (red) signals. Scale bars: 2.5 and 5 μm, respectively. A nucleus (yellow box) is magnified in the STED image on the right (scale bar: 0.5 μm). *a,c-e*: source data are provided as a Source Data file.

**Figure 2. Intracellular IRF2BPL foci display features of LLPS-driven condensates**

**a.** Left: confocal images of a human NPC-derived neuron after ICC for IRF2BPL (green) and MAP2 (red). Blue: DAPI. Scale bar: 20 μm. Upper right panel: overlay (scale bar: 5 μm). Below: STED of nucleus (n) and axon (a; yellow boxes; scale bars: 5 (n), 2.5 (a) μm). Arrowheads: condensates **b.** Upper central panel. NG108-15 cell after ICC for IRF2BPL. The somatic signal is saturated. Flanking panels: nucleus (n), axon (a), and growth cone (c) details. Arrowheads: condensates. Scale bar: 20 μm. **c.** Left panel: as in *b*, for HEK293 cells with intra- (arrowheads) and extra-nuclear (arrows) condensates. (blue: DAPI; scale bar: 20 μm). **d.** Time-lapse imaging ($t_0$-$t_3$; 2 s intervals, HEK293 cell) of coalescing IRF2BPL-GFP condensates (arrowheads; scale bar in **Supplementary Movie 1**). **e.** CLEM of an IRF2BPL-GFP-expressing HEK293 cell. Fluorescence (first), EM (second), and overlay (third) images. Condensates (IRF2BPL-GFP; arrowheads) are in green, nuclei (DAPI) in blue, mitochondria (MitoTracker; arrow) in red. Fluorescent condensates correspond to membraneless protein assemblies (yellow boxes; bottom panels. Scale bars: 2.5 (larger images), 0.1 (lower left panel), and 0.2 (lower right panel) μm. **f.** Confocal images of HEK293 cells treated with 1,6-Hex or 2,5-Hex, or untreated (control) and immunostained (ICC) to reveal IRF2BPL (green; blue: DAPI; scale bar: 10 μm). **g.** Effect of 1,6-Hex and 2,5-Hex on condensation. Values normalized to the control group. One-way ANOVA: $F_{(2,36)}$=10.549, p=0.00025, n=13 microscopy fields for all groups. Asterisks indicate significant differences ($p<0.05$, NK test; also in *i*,*j*; n.s.: non-significant). Box plots (also in *i*,*j*) report mean (cross), median (line), interquartile (limits), and 10-90 percentile (whiskers). **h.** Upper row: as in *f*, for cells fixed after 5' or 15' of 1,6-Hex treatment (blue: DAPI; scale bar: 40 μm). Lower row: magnified details (IRF2BPL only; scale bar: 10 μm). **i.** Effect of 1,6-Hex on condensation, as in *h*. Data presented as in *g*. One-way ANOVA: $F_{(2,20)}$=8.6752, p=0.0019, n= 9 (control), 9 (5'), and 5 (15'). **j.** IRF2BPL translocation, as in *h*, as log ratio between nuclear (N) and cytoplasmic (C) signals. One-way ANOVA: $F_{(2,20)}$=63.925, $p<0.00001$, n= 9 (control), 9 (5'), and 5 (15') microscopy fields. **k.** IRF2BPL-GFP FRAP. Upper panel: time-lapse images of photobleached (p) and control (c) condensates. Lower panel: FRAP time course ($t_{1/2}$=50 s; $M_f$=83%). *g-k*: source data are provided as a Source Data file.

**Figure 3. IRF2BPL NTR drives LLPS through the combined action of the ZnF domain and $LCR_1$**

**a.** Upper left panel: scheme of WT IRF2BPL and its IRF2BPL-NTR (a.a.1-172) and IRF2BPL-ΔNTR (a.a. 173-796) fragments. Lower panel: confocal images of HEK293 cells expressing GFP-tagged IRF2BPL, its fragments, or GFP alone (control). Blue: DAPI (as in *b* and *c*). Scale bar: 10 μm. Upper right panel: proportion of GFP-positive cell area occupied by condensates. Data normalized to WT group. One-way ANOVA: $F_{(3,99)}$=45.681, p<0.0001, n= 29 (WT), 38 (NTR), 19 (ΔNTR), 17 (control). Asterisks indicate significant differences (p<0.05, NK test; as in *b*, *c*, *e*). Box plots (also in *b*, *c*) report mean (cross), median (line), interquartile (limits), and 10-90 percentile (whiskers). **b.** As in A, for the NTR and its subparts, i.e., ZnF (a.a. 1-75) and LCR1 (a.a. 76-172). Scale bar: 10 μm. Data normalized to NTR group. One-way ANOVA: $F_{(3,37)}$=856.24, p<0.0001, n=10 (NTR), 13 (ZnF), 13 ($LCR_1$), 5 (control). **c**. As in A, for WT full-length IRF2BPL, IRF2BPL-ΔZnF (a.a. 76-796) and IRF2BPL-ΔLCR1. Scale bar: 10 μm. Data normalized to WT group. One-way ANOVA: $F_{(2,34)}$=126.41, p<0.00001, n= 10 (WT), 14 (ΔZnF), 13 (ΔLCR1). **d.** OptoDroplet system[46]:illumination (488 nm) triggers the condensation of an LLPS-prone peptide of interest (POI) in fusion with mCherry (mCh) and a cryptochrome fragment (Cry2). **e-f.** Graph (*e*) reporting the proportion (%; mean ± SEM) of mCh-positive cell area occupied by condensates in cultures expressing the indicated constructs that were either photoactivated (488 nm) or not (*f*; scale bar: 10 μm). Two-way ANOVA: $F_{(3,77)}$=20.804, p<0.00001, n=12 microscopy fields for all groups except for control groups (n=7 and 6). **g.** DIC and fluorescence imaging of FITC-labelled ZnF peptide droplets formed in vitro in saline buffer. Top left image: low-magnification image. Arrowhead: condensate. Scale bar: 20 μm. The square boxed in yellow is magnified in the second panel. Scale bar: 4 μm. The corresponding DIC and overlay images are in the third and fourth panels. Bottom row: higher magnification images of condenstes. Scale bar: 2.5 μm. *a-c,-e*: source data are provided as a Source Data file.

**Figure 4. Disease-related truncated mutants of IRF2BPL undergo aberrant LLPS**

**a.** Upper panel: schematic representation of known disease-related protein truncation sites (lollipop plot, 'sites') and their local frequency profile (using a 21-residue sliding window) along the IRF2BPL

primary sequence. Truncating variants group into three clusters within, or at the ends, of LCRs (cluster 1-3, horizontal gray bars). Most truncations belong to cluster 1, upper line with arrowhead. We generated constructs to express truncated IRF2BPL variants (bars below), i.e. Y173X, Q123X, Q116X, R331X, and A708X as fusion proteins with GFP. Lower panel: sequence alignments for representative cluster 1-3 pathogenic IRF2BPL variants. Asterisks above or below the colored blocks mark the beginning of frameshifted regions in some mutants. **b.** Confocal images of HEK293 cells expressing GFP tagged full-length WT IRF2BPL or its indicated mutant forms for 24h. DAPI-stained nuclei are in blue. Scale bar: 10 μm. **c.** Box plots showing (from left to right), the relative GFP-positive cell area occupied by condensates (first graph from the left; one-way ANOVA: $F_{(3,58)}$=15.892, p<0.00001, n=16 (WT), 15, (Y173X), 15 (Q123X),16 (Q116X) microscopy fields), the number of condensates per GFP-positive cell unit area (second graph; one-way ANOVA: $F_{(3,58)}$=19.674, p<0.00001, n=16 (WT), 15, (Y173X), 15 (Q123X),16 (Q116X) microscopy fields), the average condensate size (third graph; one-way ANOVA: $F_{(3,45523)}$=266.29, p<0.00001, n=16362 (WT), 13087, (Y173X), 8599 (Q123X), 7479 (Q116X) condensates), and the proportion (%) of nuclear condensates (fourth graph; one-way ANOVA: $F_{(3,20)}$=367.84, p<0.00001, n=6 microscopy fields per group) in HEK293 cell expressing the indicated proteins for 24h. In the first three graphs, values are normalized to those of the WT group. Asterisks indicate statistically significant differences (p<0.05, NK test). Box plots report mean (cross), median (line), interquartile (limits), and 10-90 percentile (whiskers). **d.** Histograms showing the distribution of condensates, formed by the indicated GFP-tagged proteins, binned by eccentricity (0.05 bins). Note how the proportion of condensates with higher eccentricity increases progressively with the shortening of $LCR_1$ in the different mutants. *a,c,d*: source data are provided as a Source Data file.

**Figure 5. Disease-related truncating mutations enhance coalescence, reduce sensitivity to 1,6-Hex, and blunt FRAP of IRF2BPL condensates**

**a.** Five frames (3 h intervals) of time-lapse experiments highlighting the coalescence dynamics of condensates of GFP-tagged WT and mutant IRF2BPL. Scale bars: 10 μm. **b.** First graph from the left: number of condensates per cell, from 24 to 36 hours post-transfection. Second graph: initial (24 h) and final (36 h) number of condensates. A two-way ANOVA (construct: $F_{(2,380)}$=29.309, p<0.00001, n=102

(WT), 57 (Y173X), 34 (Q123X) cells) showed how the reduction in condensate number occurred for all constructs but was more pronounced for the mutants, as also highlighted in the two graphs on the right (data normalized to 24 h; one-way ANOVA: $F_{(2,190)}$=4.7065, p=0.01, n=102 (WT), 57 (Y173X), 34 (Q123X) cells; asterisks indicate statistically significant differences (p<0.05, NK test). Box plots (also in *d*) report mean (cross), median (line), interquartile (limits), and 10-90 percentile (whiskers). **c.** Confocal fluorescence images of HEK293 cultures expressing WT, Y173X, Q123X, or Q116X IRF2BPL, that were treated (+) or not (-) with 1,6-Hex for 15 min. Scale bar: 10 μm. **d.** Box plots showing the relative GFP-positive cell area occupied by condensates (first graph from the left; one-way ANOVA: $F_{(3,34)}$=20.476, p<0.00001, n=11 (WT), 13 (Y173X), 9 (Q123X), 5 (Q116X) microscopy fields), the number of condensates per GFP-positive cell unit area (second graph; one-way ANOVA; $F_{(3,34)}$=24.177, p<0.00001, n as above), and the individual condensate size (third graph; one-way ANOVA: $F_{(3,34)}$=10.559, p=0.00005, n per group as above) in HEK293 cultures, expressing the indicated constructs, that were treated (+) or not (-) with 1,6-Hex for 15 min as in *c* (asterisks indicate statistically significant differences; p<0.05, NK test). **e.** FRAP dynamics for condensates of WT and mutant IRF2BPL. For each construct, photobleached (p) and control (c) condensates are shown before (pre) and at the indicated times after photobleaching, as in **Fig. 2K**. **f.** Differential FRAP time course for WT (same data as in **Fig. 2K**), Y173X, and Q123X condensates. *b*,*d*,*f*: source data are provided as a Source Data file.



**Figure 6. Altered morphological and ultrastructural features of condensates formed by disease-related IRF2BPL mutants**

Ultrastructure of HEK293 cells expressing either WT IRF2BPL (left column) or two disease-related mutants, i.e., Y173X (middle column) or Q123X (right column), as GFP fusions, as revealed by CLEM (upper row; scale bars: 5 μm). Ns label cell nuclei. The second and third rows display details of the EM signal at different levels of magnification (scale bars: 1 μm for WT and Q123X, 0.5 μm for Y173X). The third row presents image details corresponding to the 0.5x0.5 μm yellow squares in the second row. Note how WT IRF2BPL appears to accumulate isotropically in rounded condensates, whereas the two

truncated mutants form fibrillary structures within more eccentric, elongated condensates. The fibrils are better highlighted upon fast Fourier transform (FFT) image processing (last three rows; see also **Supplementary Fig. 8**; see Methods). The fourth row displays pseudocolored inverse FTT (iFTT) images derived from the corresponding EM images shown in the third row. The associated Fourier space representation images (FFT) are in the bottom row. The even spacing of the parallel fibrillary profiles (arrowheads) within the Y173X and Q123X mutant condensates is highlighted by white dots. Moreover, the plots in the fifth row presents graphs of fibril spacing measured along the inclined white segments shown in the pseudocolored images of the fourth row. Note how the WT condensate does not display any regular internal periodicity, whereas both Y173X and Q123X condensates display a regular pattern of parallel fibrillary profiles, with a mean spacing ($\bar{s}$) of 16.4 and 16.8 nm, respectively.

**Figure 7. Pathogenic truncated forms of IRF2BPL display lower degradation rates and their aberrant condensates recruit WT IRF2BPL in a length-dependent manner**

**a.** Immunoblot (α-GFP) of lysates of HEK293 cells expressing GFP-tagged WT or mutant IRF2BPL, or GFP alone (control) and treated with CHX (0.05 μg/ml) or vehicle alone (control). **b.** Quantification of experiments shown in *a*. For each construct, data are normalized to the respective control group (one-way ANOVA: $F_{(3,8)}$=9.2558, p=0.0055, n=3 per group; asterisks indicate significant differences, p<0.05 NK test). **c.** As in B, data normalized to the WT group (one-way ANOVA: $F_{(2,6)}$=7.1490, p=0.025, n=3 per group; asterisks indicate significant differences, p<0.05 NK test). **d.** NG108-15 cells expressing GFP-tagged IRF2BPL mutants (green), or GFP alone (**Supplementary Fig. 9A**), immunostained to detect endogenous WT IRF2BPL (red; blue: DAPI). Nuclei of transfected cells ($n_t$) are depleted in WT IRF2BPL. Two nuclei are highlighted in white. Arrowheads: cytoplasmic condensates of mutant proteins (but not GFP alone; **Supplementary Fig. 9a**) recruit WT IRF2BPL. Scale bar: 10 μm. **e.** Number of condensates per nuclear unit area in cells expressing the indicated constructs vs. non-transfected control cells in the same cultures. Data normalized to those of control cells for each coverslip (two-way ANOVA: $F_{(3,3335)}$=15.683, p<0.00001, n=501 (Y173X,$n_c$), 500 (Y173X,$n_t$), 957 (Q123X,$n_c$), 403 (Q123X,$n_t$), 296 (Q116X,$n_c$), 168 (Q116X,$n_t$), 381 (GFP,$n_c$), 137 (GFP,$n_t$) cells; asterisks indicate

significant differences, p<0.05 NK test). Box plots (also in *f*) report mean (cross), median (line), interquartile (limits), and 10-90 percentile (whiskers). **f.** Efficiency of the three mutants (panels *d*,*e*) in reducing WT IRF2BPL intranuclear condensates in a length-dependent manner (two-way ANOVA and asterisks as in *e*; data normalized to the GFP group). **g.** Lower right image: As in *d*, for human NPC-derived neurons. Scale bar: 10 μm. Lower middle image: MAP2 immunostaining (white). Scale bar: 10 μm. All other images are magnifications of a single neuron. Scale bar (as for the upper right image): 5 μm. Y173X condensates recruit WT IRF2BPL. **h.** As in *d*, for a NG108-15 cell axon. Axonal Y173X condensates (arrowhead) recruit endogenous IRF2BPL. Scale bar: 5 μm. *b*,*c*,*e*,*f*: source data are provided as a Source Data file.



**Figure 8. IRF2BPL disease-related mutants undergoing aberrant LLPS trigger cellular and molecular disease mechanisms**

**a.** CRISPR/Cas9-generated monoallelic c.379C>T mutation in HEK293 cells which express IRF2BPL-Q127X together with the WT protein. **b.** Confocal images of WT IRF2BPL immunostaining in Q127X (right) and isogenic control (left) cells. WT IRF2BPL is largely recruited to cytoplasmic condensates. Scale bars: 40 (upper) and 10 (lower) μm. **c.** qPCR showed a significant WNT1 upregulation in Q127X cells (asterisk: p=0.0049, *t*-test, n=3 per group; mean ± SEM). **d**. Western immunoblotting showing the absence of IRF2BPL staining in lysates of CRISPR/Cas9-edited IRF2BPL KO HEK293 cells (right lane) but not in isogenic WT cells. **e.** IRF2BPL immunostaining shows the lack of IRF2BPL expression in KO, but not in WT, cells. Scale bars: 40 μm. **f.** qPCR shows a significant upregulation of WNT1 in KO cells (asterisk: p=0.012, *t*-test, n=5 per group; mean ± SEM). **g.** Upper panels: current clamp recordings of voltage responses to depolarizing current steps (0-300 pA) delivered to NG108-15 cells expressing WT or mutant IRF2BPL, or GFP (control). Cells expressing Y173X or Q123X display and impaired spiking ability. **h-j.** Quantification, in neuronal cells expressing the indicated constructs, of resting membrane potential (RMP; panel *h*; one-way ANOVA: $F_{(3,86)}$=4.5456, p=0.0052, n=25 (WT), 21 (Y173X), 18 (Q123X), 26 (GFP) cells), and peak spike amplitude measured either from RMP (panel *i*; $F_{(3,86)}$=3.7546, p=0.013, n as in *h*) or from the steady-state potential (*sp*) after the peak (ΔV peak-*sp*;

panel *j*; one-way ANOVA: $F_{(3,85)}$=4.5359, p=0.0053, n= 25 (n as in *h* except for GFP (n=25)). Values normalized to the WT group mean (RMP: -74.5 mV; spike amplitude: 93.1 mV from RMP and 24.2 mV from *sp*). Asterisks: significant differences, p<0.05 NK test (also in *k*). Box plots report mean (cross), median (line), interquartile (limits), and 10-90 percentile (whiskers). **k.** Cells with depolarization peak <70 mV are more numerous in the Y173X and Q123X groups (FE test: p=0.0061 (WT vs.173X). p=0.01 (WT vs. Q123X), n= 25 (WT), 21 (Y173X), 18 (Q123X) cells) but not in the GFP group (n=26; p=0.22, FE test). *c-d,f,h-k*: source data are provided as a Source Data file.

**Figure 9. Circular dichroism identifies a folded core of the HCD as in the AF structural model**

**a.** Upper left panel: AF structural model (**Fig. 1a**) detail highlighting the HCD (green). A CC-prone helix ($CC_2$) precedes an α/β-sandwich with structural homology to methyl-DNA-binding domains. The domain core (upper middle panel), with a long β-hairpin ($\beta_2$-$\beta_3$ strands of the β-sheet) and a helix (α) is similar to the DNA-binding interface of human MBD and MeCP2 proteins. Upper right panel: structural matching of the HCD core with the DNA-binding region of human MBD-2, -4, -5 and MeCP2 from solved PDB structures. Lower left panel: HCD core (green) superimposed to the DNA-bound portion of MeCP2 (PDB: 6C1Y)[101]. The β-hairpin enters the DNA major groove. Lower middle panel: MeCP2 is omitted, highlighting in red five HCD residues that may bind DNA (see panel B), as highlighted in the lower right panel. **b.** Structural (upper panel) and primary sequence alignment of the HCD core (a.a. 389-461) with corresponding regions of human MBD1, MBD2, and MeCP2, and of *Drosophila* MBD2/3 (X stands for the species-specific insert 'ANNNASSNNNSSATASSNNNNN'). **c.** CD spectra (mean residue molar ellipticity (MRME), 190–250 nm range) of an HCD core peptide (a.a. 395-448). The spectrum at physiological temperature (37°C, saline buffer; top panel) indicates a mixture of secondary structures, with a random coil (rc)-related negative peak (~200 nm) and a shoulder around 220 nm indicative of structured components. The stability of the structure was tested at 5, 25, 50, and 75 °C (bottom panel). **d.** Upper graph: relative proportion of α-helical (α), β-strand (β), turn, and random coil (rc) conformations in the HCD core peptide as predicted by AF and as measured at 37 °C using CD. The CD results are consistent with the AF model. Lower graph: proportion of the two prevalent

structural components (e.g., β and rc) in CD spectra measured at increasing temperatures. *c*,*d*: source data are provided as a Source Data file.

**Figure 10. Disease-related missense mutations can lead to aberrant LLPS and depletion of intranuclear condensates formed by WT IRF2BPL**

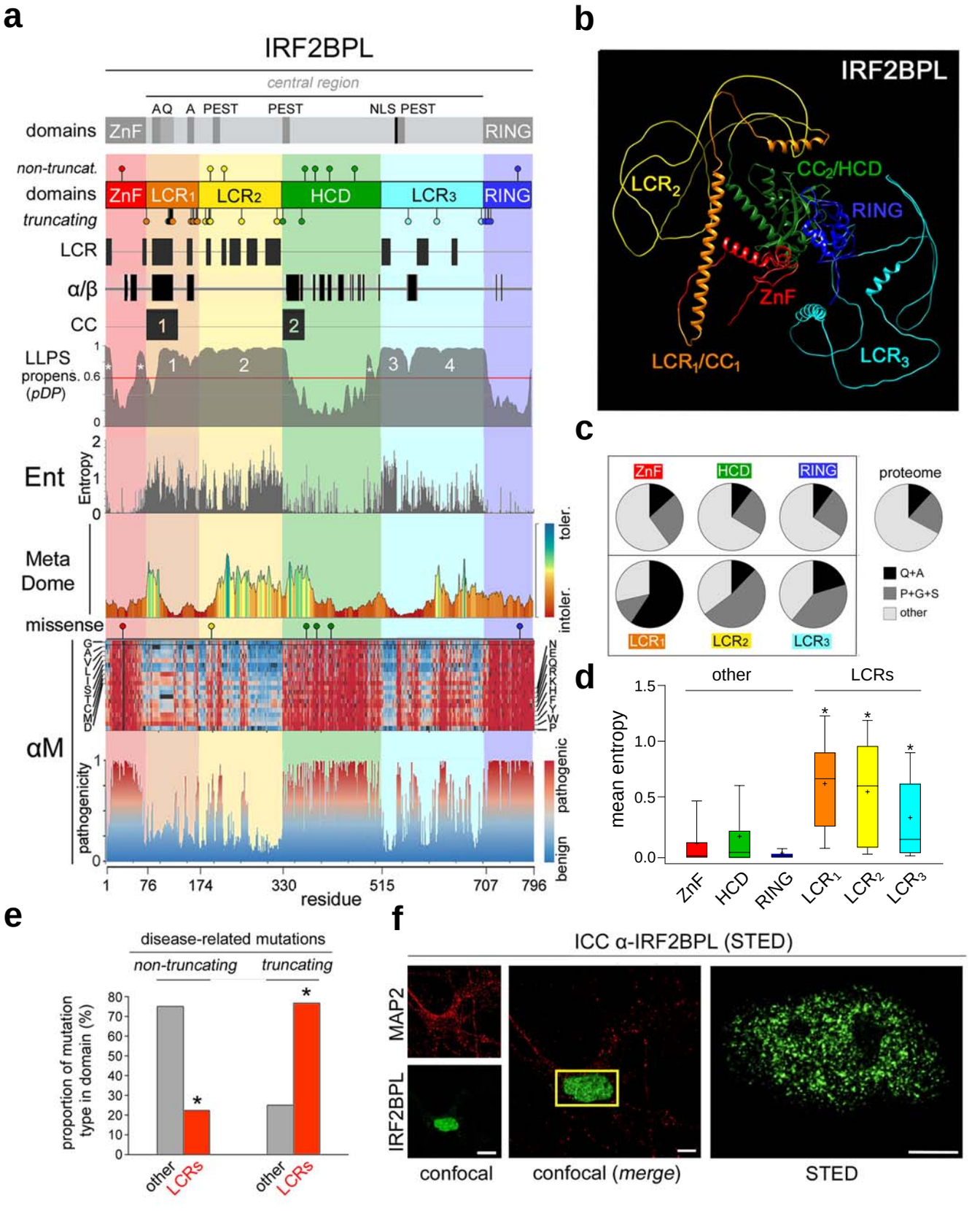
**a.** AF HCD model highlighting three pathogenic substitutions. **b.** Sites (lollipop plot) of disease-related non-truncating mutation ('NTMs') and their frequency profile (51-residue window). Half of them affect the HCD (cluster 4). Lower panel: IRF2BPL mutants (F30L, P372R, K418N) whose LLPS was experimentally characterized. **c.** HEK293 cells expressing GFP-tagged WT or mutant IRF2BPL (green; blue: DAPI). Scale bar: 10 μm. **d.** GFP-positive cell area occupied by condensates (first graph from the left, one-way ANOVA: $F_{(3,39)}$=0.18116, p=0.90, n=11 (WT), 11 (F30L), 10 (P372R), 11 (K418N) microscopy fields), number of condensates per GFP-positive cell unit area (second graph, $F_{(3,39)}$=4.0358, p=0.013, n as in first graph), average condensate size (third graph, $F_{(3,51153)}$=87.960, p=0.0001, n=20419 (WT), 13835 (F30L), 9749 (P372R), 7154 (K418N) condensates) and proportion of condensate area (%) in the nucleus (fourth graph, $F_{(3,18)}$=29.016, p<0.00001, n=6 (WT), 5 (F30L), 5 (P372R), 6 (K418N) microscopy fields) in HEK293 cells (24 h of expression). In the first three graphs, values are normalized to the WT group. Asterisks: significant differences (p<0.05 NK test; also in *f*,*g*). Box plots (also in *f*,*g*) report mean (cross), median (line), interquartile (limits), and 10-90 percentile (whiskers). **e.** NG108-15 cells, expressing either IRF2BPL-K418N-GFP or GFP, immunostained to reveal total IRF2BPL (WT and mutant, in red; blue: DAPI). Some nuclear profiles are highlighted in white. Arrowhead: recruitment of the WT protein to a cytoplasmic condensate in a K418N-expressing cell. Scale bar: 10 μm. **f.** Mean number of condensates per nuclear unit area in cells expressing ($n_t$) IRF2BPL-K418N-GFP or GFP alone, in comparison with neighboring non-transfected control cells ($n_c$) in the same cultures. Data normalized to the mean value in control cells. Two-way ANOVA: $F_{(1,107)}$=5.2052, p=0.024, n= 37 (K418N,$n_c$), 17 (K418N,$n_t$), 36 (GFP,$n_c$), 21 (GFP,$n_t$). **g.** Mean number of condensates per nuclear unit area in cells alternatively expressing IRF2BPL-K418N-GFP or GFP alone ($n_t$). Data

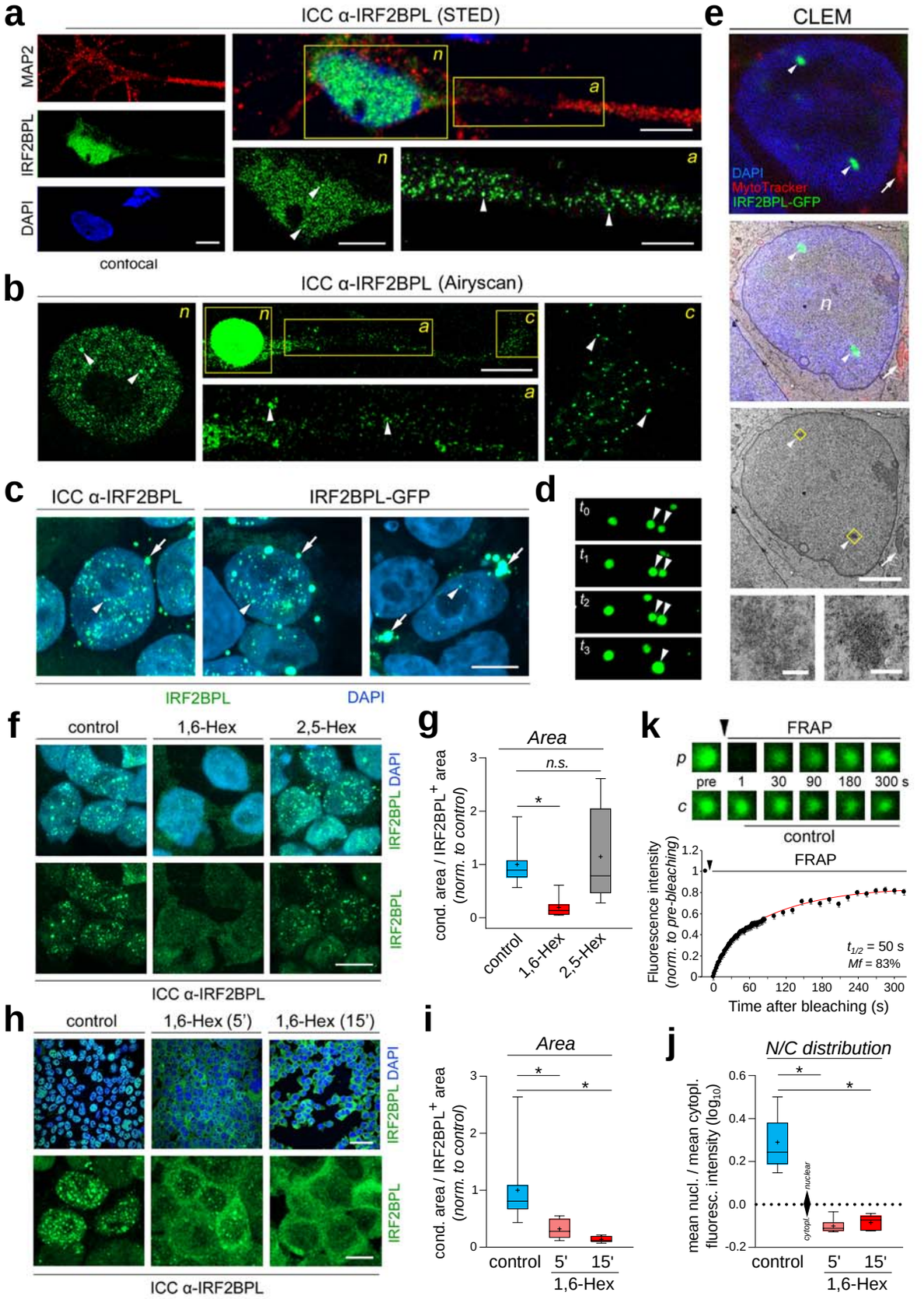
first normalized to non-transfected controls for both groups and then to the GFP $n_t$ group. Asterisk: p=0.0029, t-test, n=17 (K418N), 21 (GFP). *b,d,f,g*: source data are provided as a Source Data file.

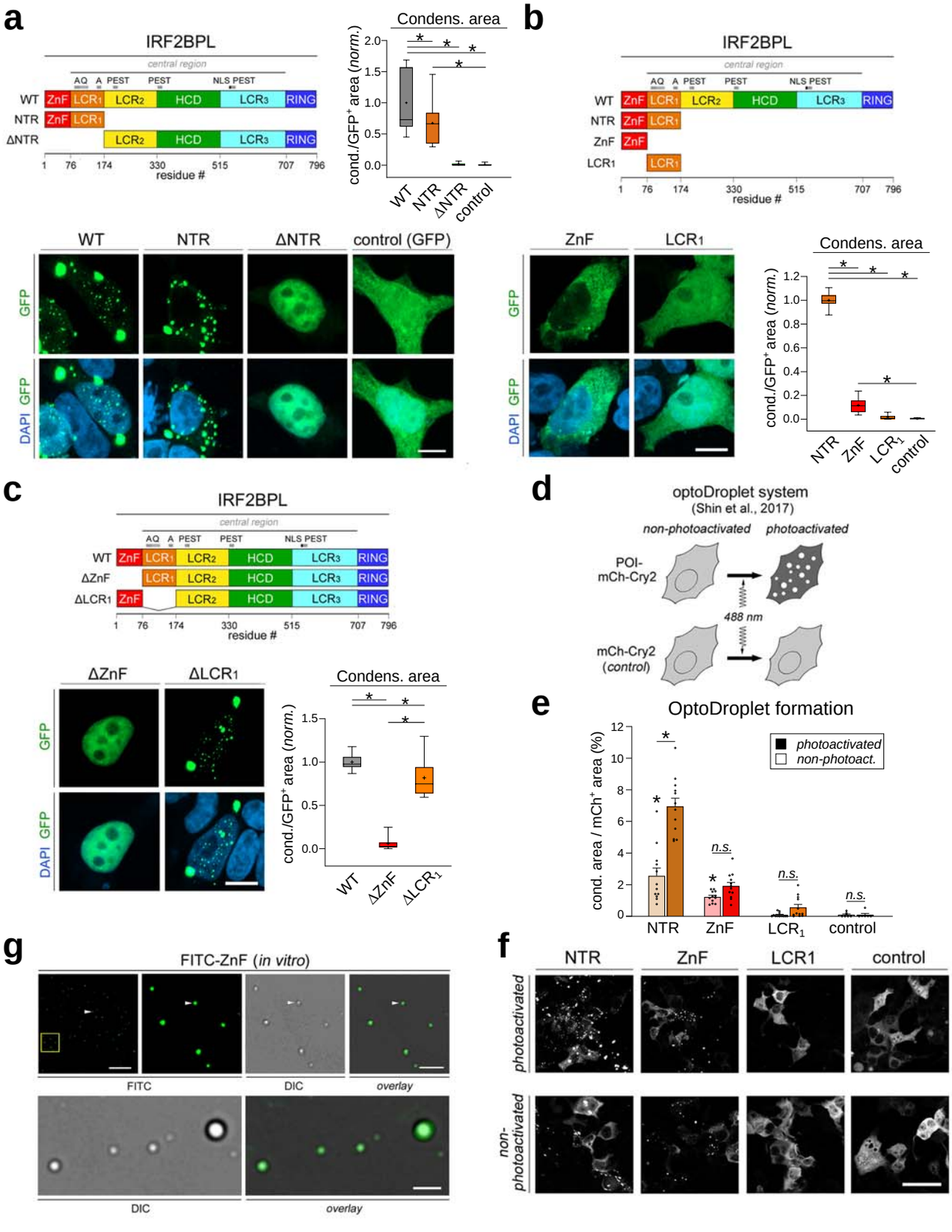
**Editor's Summary**

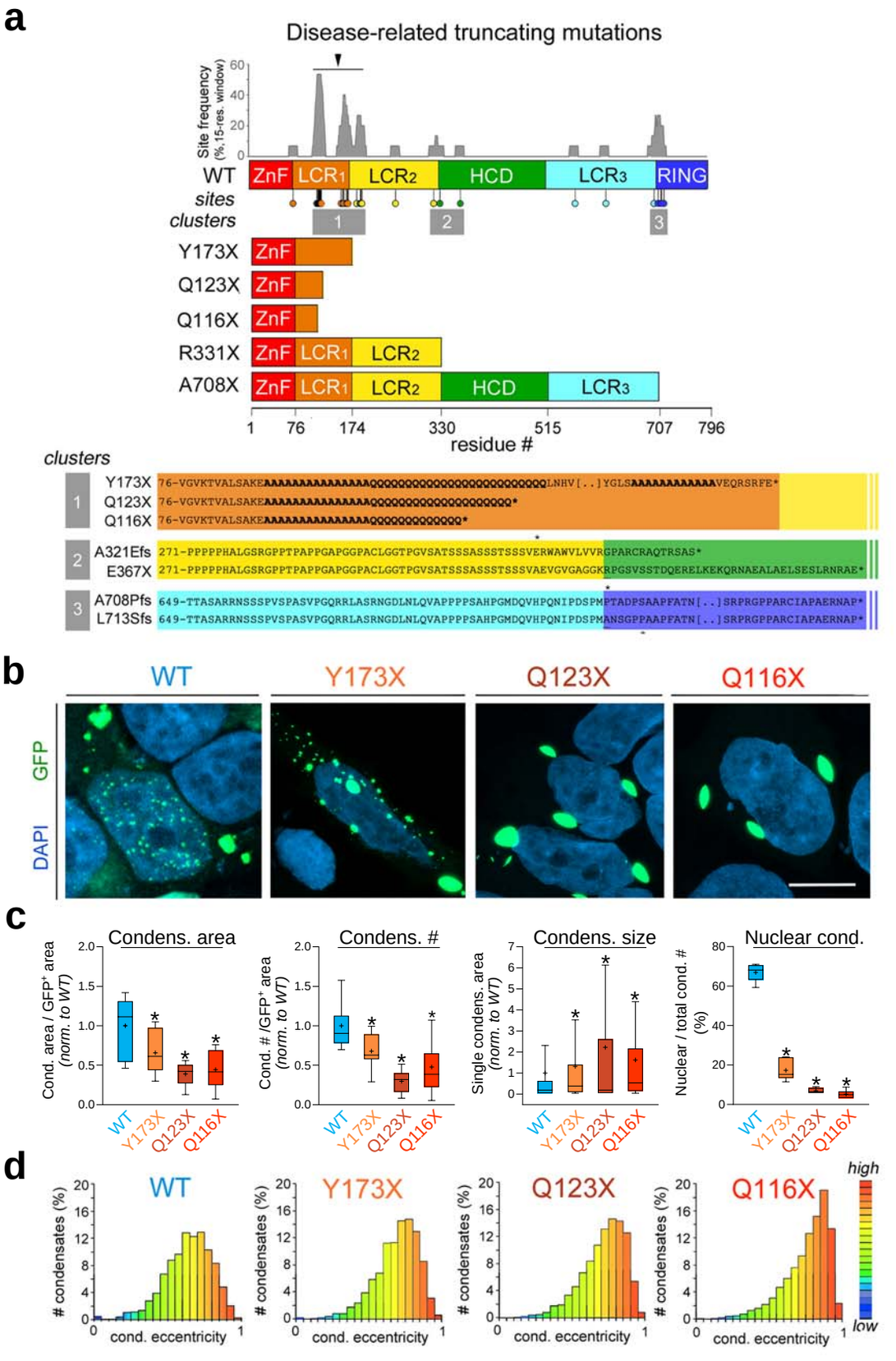
This article shows how mutations causing the neurodevelopmental syndrome NEDAMSS and related disorders are associated with the aberrant intracellular condensation of the IRF2BPL protein, which triggers cellular and molecular disease mechanisms.

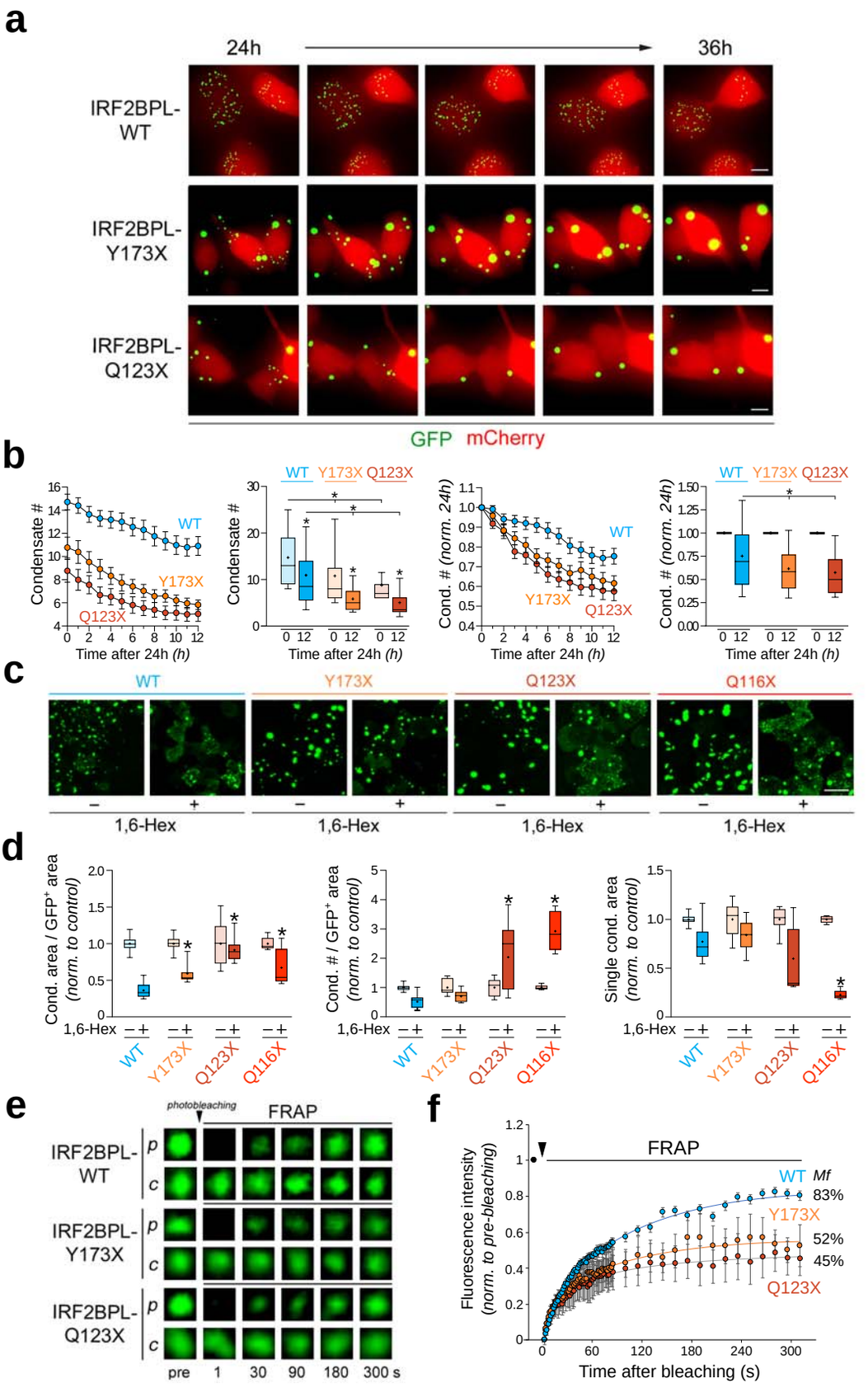
**Peer review information:** *Nature Communications* thanks Tiratha Raj Singh and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

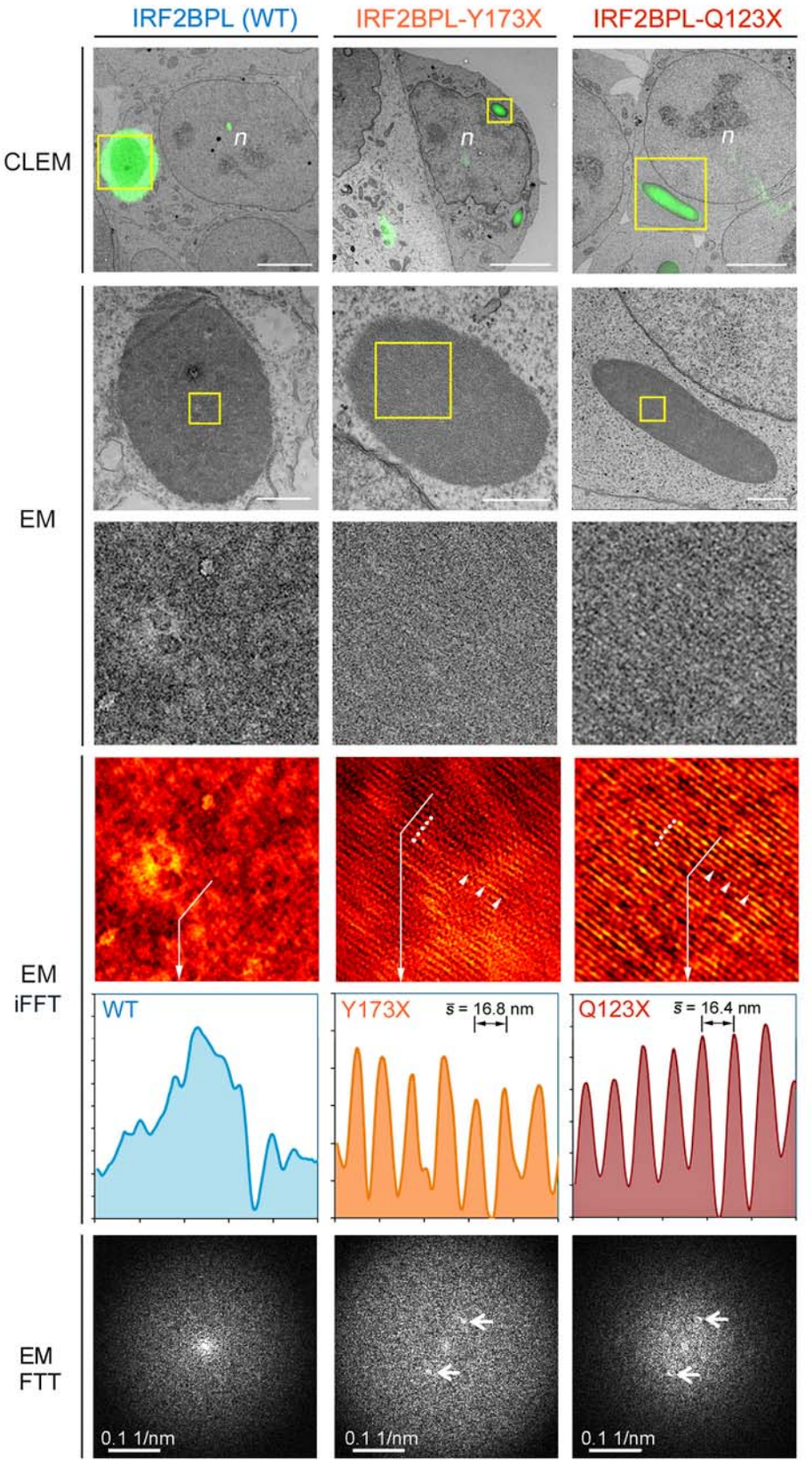

a
IRF2BPL
central region
AQ A PEST PEST NLS PEST
domains
ZnF RING
non-truncat.
domains
ZnF LCR1 LCR2 HCD LCR3 RING
truncating
LCR
α/β
CC
1
2
LLPS
propens.
(pDP)
1 2 3 4
Ent
Entropy
Meta
Dome
missense
αM
pathogenicity
toler.
intoler.
pathogenic
benign
1 76 174 330 515 707 796
residue
b
IRF2BPL
LCR2
CC2/HCD
RING
ZnF
LCR1/CC1
LCR3
c
ZnF
HCD
RING
proteome
LCR1
LCR2
LCR3
Q+A
P+G+S
other
d
other
LCRs
mean entropy
1.5
1.0
0.5
0.0
ZnF
HCD
RING
LCR1
LCR2
LCR3
e
disease-related mutations
non-truncating
truncating
proportion of mutation type in domain (%)
80
70
60
50
40
30
20
10
0
other
LCRs
other
LCRs
f
ICC α-IRF2BPL (STED)
MAP2
IRF2BPL
confocal
confocal (merge)
STED

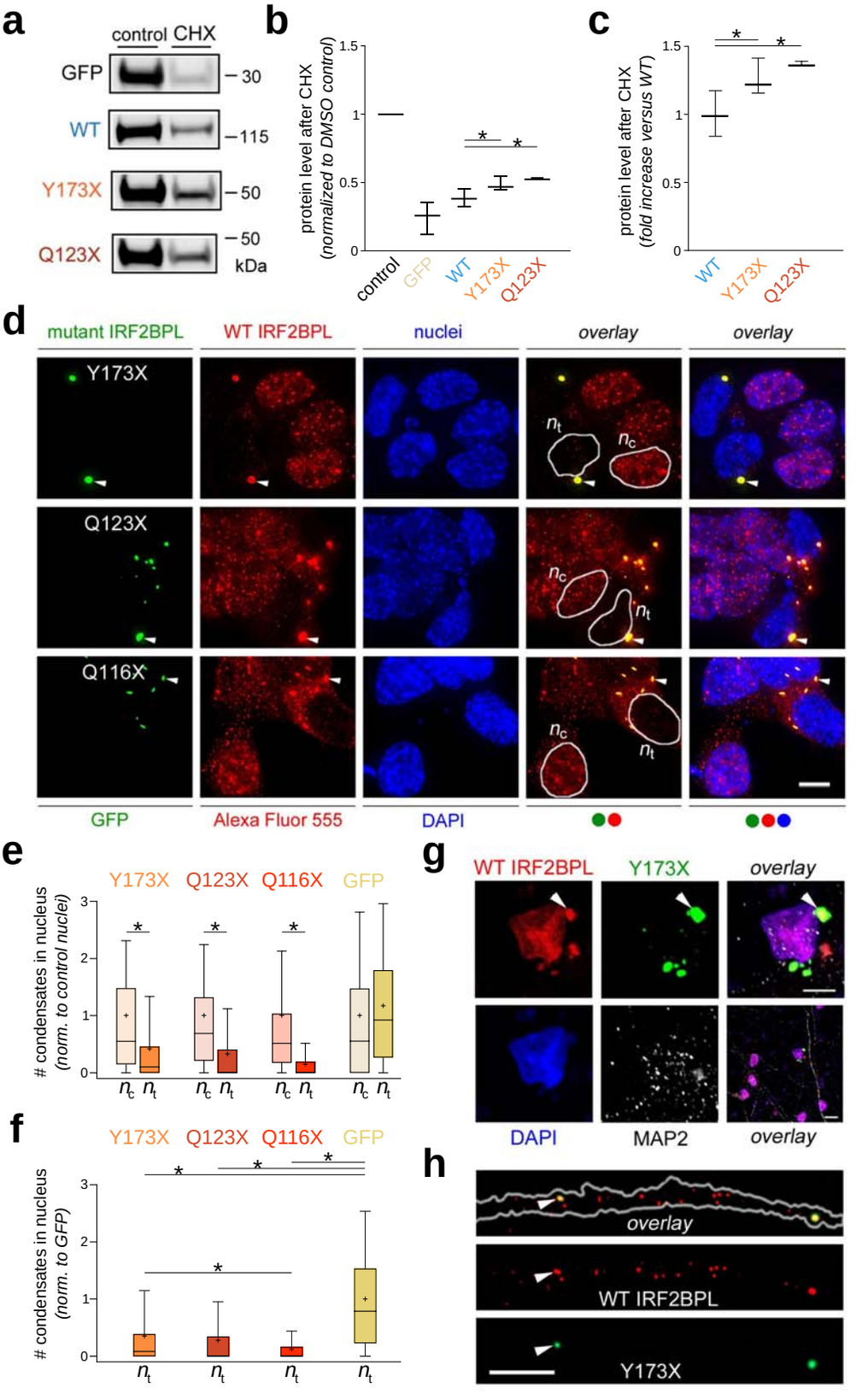

a
ICC α-IRF2BPL (STED)
MAP2
IRF2BPL
DAPI
contocal
n
a
b
ICC α-IRF2BPL (Airyscan)
c
e
CLEM
DAPI
MytoTracker
IRF2BPL-GFP
c
ICC α-IRF2BPL
IRF2BPL-GFP
IRF2BPL
DAPI
d
$t_0$
$t_1$
$t_2$
$t_3$
f
control
1,6-Hex
2,5-Hex
IRF2BPL DAPI
IRF2BPL
ICC α-IRF2BPL
g
Area
n.s.
cond. area / IRF2BPL+ area (norm. to control)
control
1,6-Hex
2,5-Hex
k
FRAP
p
c
pre
1
30
90
180
300 s
control
FRAP
Fluorescence intensity (norm. to pre-bleaching)
$t_{1/2}$ = 50 s
Mf = 83%
Time after bleaching (s)
h
control
1,6-Hex (5')
1,6-Hex (15')
IRF2BPL DAPI
IRF2BPL
ICC α-IRF2BPL
i
Area
cond. area / IRF2BPL+ area (norm. to control)
control
5'
15'
1,6-Hex
j
N/C distribution
mean nucl. / mean cytopl. fluoresc. intensity ($log_{10}$)
cytopl.
nuclear
control
5'
15'
1,6-Hex

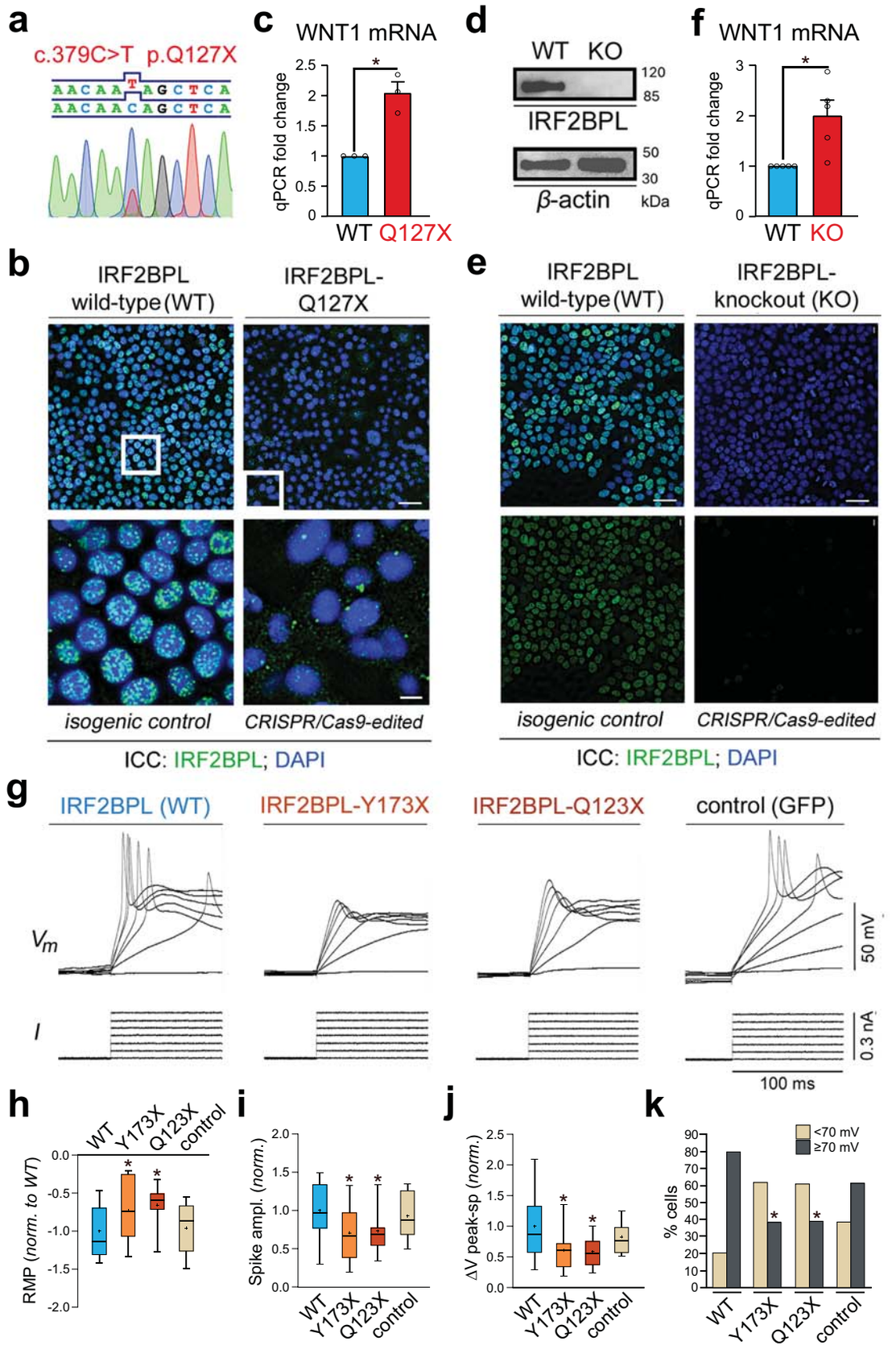

a
IRF2BPL
central region
AQ A PEST PEST NLS PEST
WT ZnF LCR1 LCR2 HCD LCR3 RING
NTR ZnF LCR1
ΔNTR LCR2 HCD LCR3 RING
1 76 174 330 515 707 796
residue #
Condens. area
cond./GFP+ area (norm.)
WT NTR ΔNTR control
WT NTR ΔNTR control (GFP)
GFP
DAPI GFP
b
IRF2BPL
central region
AQ A PEST PEST NLS PEST
WT ZnF LCR1 LCR2 HCD LCR3 RING
NTR ZnF LCR1
ZnF ZnF
LCR1 LCR1
1 76 174 330 515 707 796
residue #
ZnF LCR1
GFP
DAPI GFP
Condens. area
cond./GFP+ area (norm.)
NTR ZnF LCR1 control
c
IRF2BPL
central region
AQ A PEST PEST NLS PEST
WT ZnF LCR1 LCR2 HCD LCR3 RING
ΔZnF LCR1 LCR2 HCD LCR3 RING
ΔLCR1 ZnF LCR2 HCD LCR3 RING
1 76 174 330 515 707 796
residue #
ΔZnF ΔLCR1
GFP
DAPI GFP
Condens. area
cond./GFP+ area (norm.)
WT ΔZnF ΔLCR1
d
optoDroplet system
(Shin et al., 2017)
non-photoactivated photoactivated
POI-mCh-Cry2
488 nm
mCh-Cry2 (control)
e
OptoDroplet formation
photoactivated
non-photoact.
cond. area / mCh+ area (%)
n.s.
NTR ZnF LCR1 control
g
FITC-ZnF (in vitro)
FITC DIC overlay
DIC overlay
f
NTR ZnF LCR1 control
photoactivated
non-photoactivated

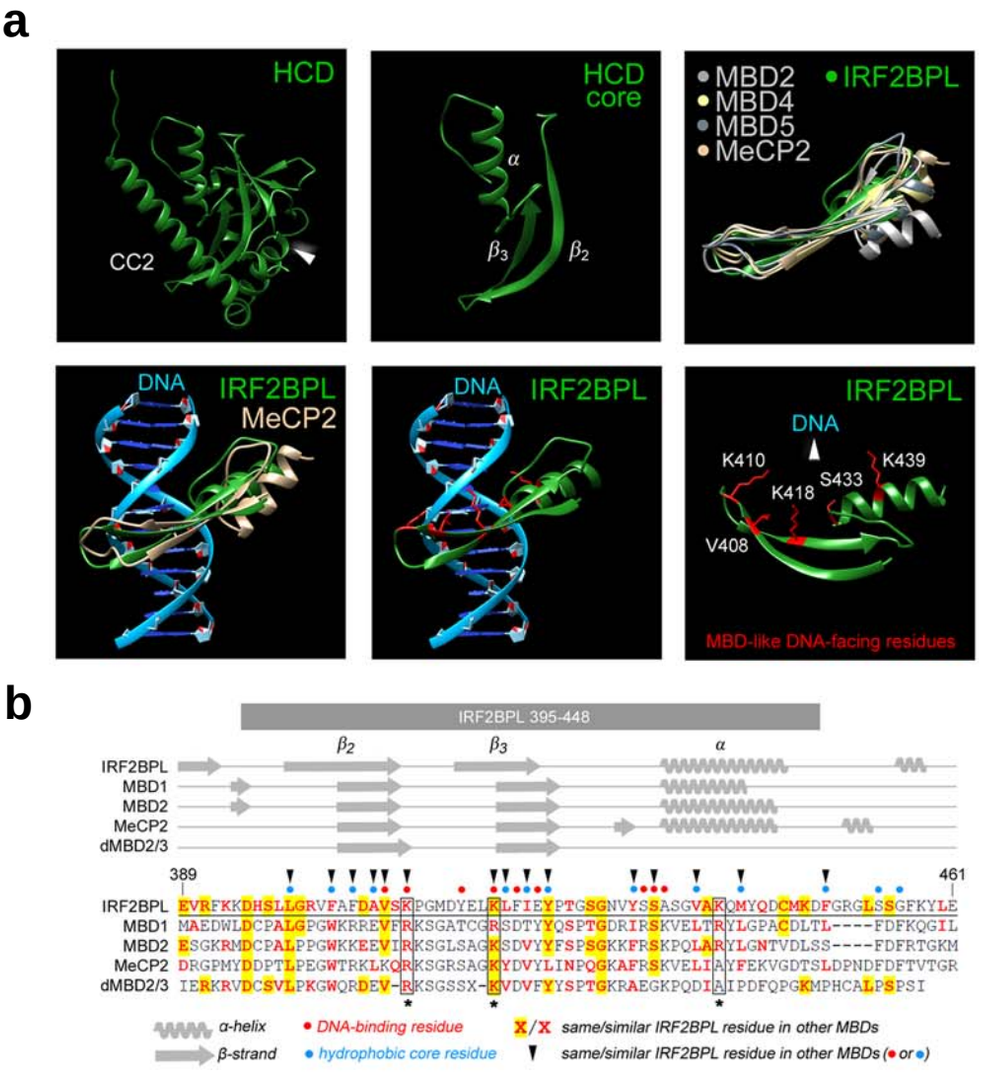

a
Disease-related truncating mutations
Site frequency (%, 15-res. window)
WT
ZnF
LCR1
LCR2
HCD
LCR3
RING
sites
clusters
Y173X
Q123X
Q116X
R331X
A708X
residue #
clusters
A321Efs
E367X
A708Pfs
L713Sfs
b
WT
Y173X
Q123X
Q116X
GFP
DAPI
c
Condens. area
Condens. #
Condens. size
Nuclear cond.
Cond. area / GFP+ area (norm. to WT)
Cond. # /GFP+ area (norm. to WT)
Single condens. area (norm. to WT)
Nuclear / total cond. # (%)
d
# condensates (%)
cond. eccentricity
high
low

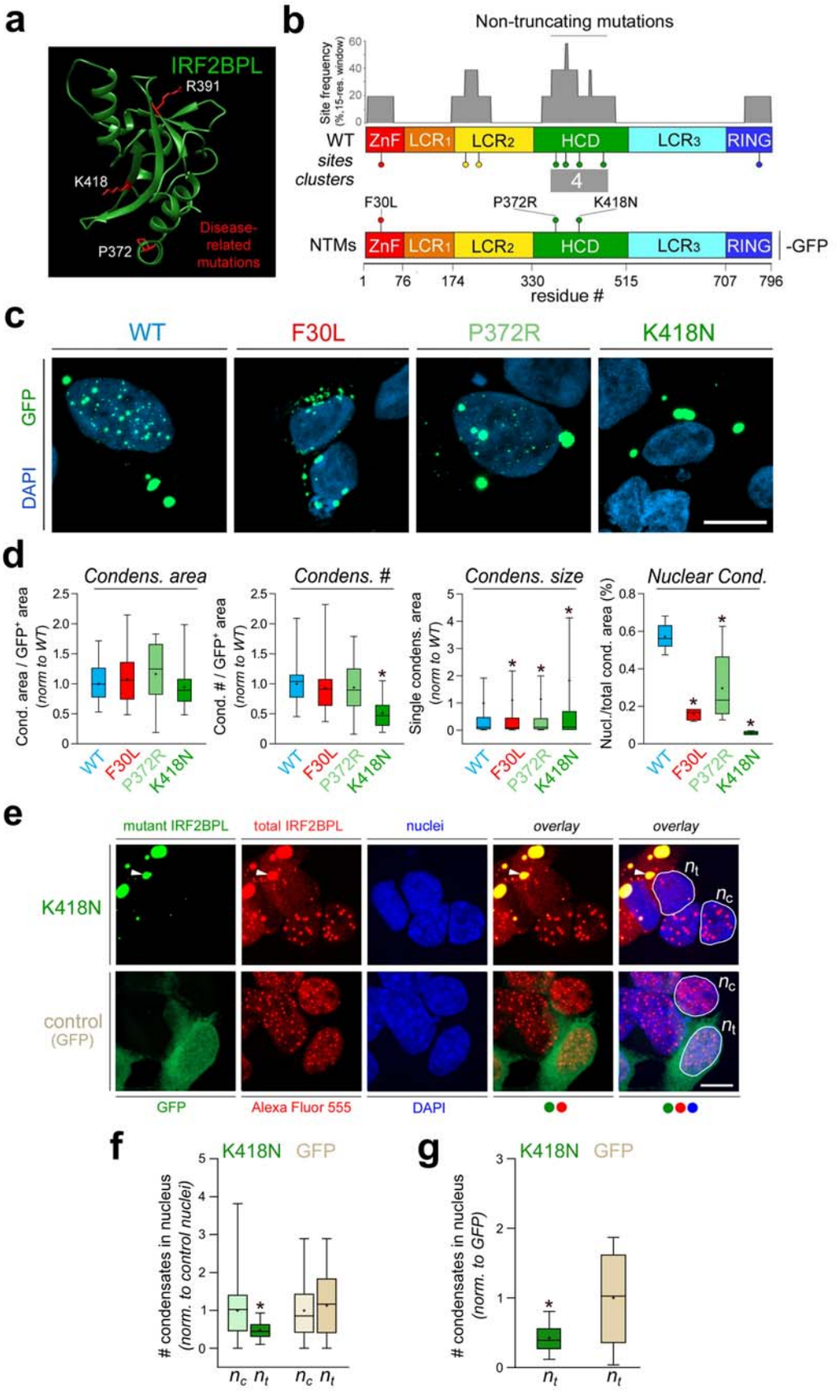

a
24h
36h
IRF2BPL-WT
IRF2BPL-Y173X
IRF2BPL-Q123X
GFP mCherry
b
Condensate #
Time after 24h (h)
WT
Y173X
Q123X
Cond. # (norm. 24h)
c
Q116X
1,6-Hex
d
Cond. area / GFP+ area (norm. to control)
Cond. # / GFP+ area (norm. to control)
Single cond. area (norm. to control)
e
photobleaching
FRAP
p
c
pre
1
30
90
180
300 s
f
Fluorescence intensity (norm. to pre-bleaching)
Time after bleaching (s)
Mf
83%
52%
45%


IRF2BPL (WT)
IRF2BPL-Y173X
IRF2BPL-Q123X
CLEM
EM
EM iFFT
EM FTT
n
WT
Y173X
Q123X
s̄ = 16.8 nm
s̄ = 16.4 nm
0.1 1/nm
0.1 1/nm
0.1 1/nm


a
control CHX
GFP
WT
Y173X
Q123X
30
115
50
50
kDa
b
protein level after CHX
(normalized to DMSO control)
control
GFP
WT
Y173X
Q123X
c
protein level after CHX
(fold increase versus WT)
WT
Y173X
Q123X
d
mutant IRF2BPL
WT IRF2BPL
nuclei
overlay
overlay
Y173X
Q123X
Q116X
GFP
Alexa Fluor 555
DAPI
e
Y173X Q123X Q116X GFP
# condensates in nucleus
(norm. to control nuclei)
f
Y173X Q123X Q116X GFP
# condensates in nucleus
(norm. to GFP)
g
WT IRF2BPL
Y173X
overlay
DAPI
MAP2
overlay
h
overlay
WT IRF2BPL
Y173X


a
c.379C>T p.Q127X
AACAATAGCTCA
AACAACAGCTCA
c
WNT1 mRNA
qPCR fold change
WT Q127X
d
WT KO
120
85
IRF2BPL
50
30
β-actin
kDa
f
WNT1 mRNA
qPCR fold change
WT KO
b
IRF2BPL wild-type (WT)
IRF2BPL-Q127X
isogenic control
CRISPR/Cas9-edited
ICC: IRF2BPL; DAPI
e
IRF2BPL wild-type (WT)
IRF2BPL-knockout (KO)
isogenic control
CRISPR/Cas9-edited
ICC: IRF2BPL; DAPI
g
IRF2BPL (WT)
IRF2BPL-Y173X
IRF2BPL-Q123X
control (GFP)
Vm
I
50 mV
0.3 nA
100 ms
h
WT
Y173X
Q123X
control
RMP (norm. to WT)
i
Spike ampl. (norm.)
WT
Y173X
Q123X
control
j
ΔV peak-sp (norm.)
WT
Y173X
Q123X
control
k
<70 mV
≥70 mV
% cells
WT
Y173X
Q123X
control


a
HCD
CC2
HCD core
α
β3
β2
MBD2
MBD4
MBD5
MeCP2
IRF2BPL
DNA
IRF2BPL
MeCP2
DNA
IRF2BPL
IRF2BPL
DNA
K410
K418
S433
K439
V408
MBD-like DNA-facing residues
b
IRF2BPL 395-448
β2
β3
α
IRF2BPL
MBD1
MBD2
MeCP2
dMBD2/3
389
461
IRF2BPL EVRFKKDHSLLGRVFAFDAVSKPGMDYELKLFIEYPTGSGNVYSSASGVAKQMYQDCMKDFGRGLSSGFKYLE
MBD1 MAEDWLDCPALGPGWKRREVFRKSGATCGRSDTYYQSPTGDRIRSKVELTRYLGPACDLTL----FDFKQGIL
MBD2 ESGKRMDCPALPPGWKKEEVIRKSGLSAGKSDVYYFSPSGKKFRSKPQLARYLGNTVDLSS----FDFRTGKM
MeCP2 DRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTSLDPNDFDFTVTGR
dMBD2/3 IERKRVDCSVLPKGWQRDEV-RKSGSSX-KVDVFYYSPTGKRAEGKPQDIAIPDFQPGKMPHCALPSPSI
α-helix
β-strand
DNA-binding residue
hydrophobic core residue
X/X same/similar IRF2BPL residue in other MBDs
same/similar IRF2BPL residue in other MBDs (● or ●)
c
IRF2BPL395-448
MRME (deg·cm2·dmol-1)
37 °C
IRF2BPL395-448
MRME (deg·cm2·dmol-1)
5 °C
25 °C
50 °C
75 °C
200 210 220 230 240 250
wavelength (nm)
d
% secondary struct.
AF
CD
α
β
turn
rc
% secondary struct.
rc
β
5 25 50 75
temperature (°C)


a
IRF2BPL
R391
K418
P372
Disease-related mutations
b
Non-truncating mutations
Site frequency (%,15-res. window)
WT
sites
clusters
ZnF
LCR1
LCR2
HCD
LCR3
RING
4
F30L
P372R
K418N
NTMs
-GFP
1
76
174
330
515
707
796
residue #
c
WT
F30L
P372R
K418N
GFP
DAPI
d
Condens. area
Cond. area / GFP+ area (norm to WT)
Condens. #
Cond. # / GFP+ area (norm to WT)
Condens. size
Single condens. area (norm to WT)
Nuclear Cond.
Nucl./total cond. area (%)
e
mutant IRF2BPL
total IRF2BPL
nuclei
overlay
overlay
K418N
control (GFP)
nt
nc
GFP
Alexa Fluor 555
DAPI
f
K418N
GFP
# condensates in nucleus (norm. to control nuclei)
nc
nt
g
K418N
GFP
# condensates in nucleus (norm. to GFP)
nt