

Intrinsically disordered N-terminal and structured DNA-binding domains jointly regulate progesterone receptor transcriptional condensates

Priyanka Sharma

`priyanka.sharma@inserm.fr`

Institut de Pharmacologie et de Biologie Structurale, IPBS <https://orcid.org/0000-0003-3890-3590>

Louna Raluy

Institut de Pharmacologie et de Biologie Structurale, IPBS

Dominik Saul

<https://orcid.org/0000-0002-0673-3710>

Catalina Romero-Aristizabal

Centre de Regulació Genòmica (CRG)

Borja Requena

ICFO - Institut de Ciències Fotòniques <https://orcid.org/0000-0003-1904-4137>

Lara I. de Cucalon

Centre de Regulació Genòmica (CRG)

Robyn Kosinsky

Robert Bosch Center for Tumor Diseases, Stuttgart, Germany.

Nadine Youssef Youssef

Institut de Pharmacologie et de Biologie Structurale, IPBS <https://orcid.org/0009-0005-5508-9393>

Maciej Lewenstein

ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology

<https://orcid.org/0000-0002-0210-7800>

Xavier Salvatella

Institute for Research in Biomedicine <https://orcid.org/0000-0002-8371-4185>

Gorka Munoz-Gil

University of Innsbruck <https://orcid.org/0000-0001-9223-0660>

Juan A. Torreno-Pina

ICFO – Institut de Ciències Fotòniques

Article

Keywords:

Posted Date: June 2nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9703632/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: **Yes** there is potential Competing Interest. X.S. is founder and scientific advisor of Nuage Therapeutics.

Intrinsically disordered N-terminal and structured DNA-binding domains jointly regulate progesterone receptor transcriptional condensates

Louna Raluy¹, Dominik Saul^{2, 3}, Catalina Romero-Aristizabal⁴, Borja Requena⁵, Lara I. de Llobet Cucalon⁴, Robyn Laura Kosinsky³, Nadine Youssef¹, Carla Garcia-Cabau⁶, Maciej Lewenstein^{5,7}, Xavier Salvatella^{6,7}, Gorka Muñoz-Gil⁸, Maria F. Garcia-Parajo^{5,7,*}, Juan A. Torreno-Pina^{5,9,*}, Priyanka Sharma^{1,*}

¹Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, Toulouse, France.

²Division of Endocrinology, Mayo Clinic, Rochester, MN 55905, USA
Robert and Arlene Kogod Center on Aging, Mayo Clinic, Rochester, MN 55905, USA.

³Robert Bosch Center for Tumor Diseases, Stuttgart, Germany.

⁴Centre de Regulació Genòmica (CRG), The Barcelona Institute of Science and Technology (BIST), Dr. Aiguader 88, Barcelona, Spain.

⁵ICFO-Institut de Ciències Fotòniques, The Barcelona Institute for Science and Technology (BIST), 08860 Barcelona, Spain.

⁶Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Barcelona, Spain.

⁷ICREA (Institució Catalana de Recerca i Estudis Avançats), Pg. Lluís Companys 23, 08010 Barcelona, Spain.

⁸University of Innsbruck, Department for Theoretical Physics, Technikerstr. 21a, A-6020 Innsbruck, Austria.

⁹Present address: Department of Macromolecular Structure, Centro Nacional de Biotecnología-CSIC, C/ Darwin, 3, 28049 Cantoblanco Madrid, Spain.

*Co-corresponding authors:

Maria F. Garcia-Parajo (maria.garcia-parajo@icfo.eu)

Juan A. Torreno-Pina (ja.torreno@cnb.csic.es)

Priyanka Sharma (priyanka.sharma@ipbs.fr)

Abstract

Transcriptional condensates (TCs) are dynamic, membrane-less assemblies of proteins and nucleic acids that form at specific genomic loci to control gene expression. The primary drivers of TC formation and function are long stretches of polypeptide chains that lack well-defined tertiary structure, known as the intrinsically disordered regions (IDRs) of Transcription Factors (TFs). However, the role of well-structured domains such as the DNA-binding domain (DBD) in condensate formation and downstream regulation remains poorly understood. Here, we investigated whether the DBD of the Progesterone Receptor (PR), a model TF regulated by hormone binding, contributes to the formation of PR TCs. By combining Single Particle Tracking (SPT) with deep-learning-based single-trajectory analysis, we demonstrated that both the IDR and the DBD are required for the formation of PR condensates. Furthermore, transcriptomic analysis showed that mutations of either of the two domains result in the expression of a distinct set of genes, with functional assays corroborating these findings and linking domain-specific transcription to differential effects on cell proliferation and migration. Together, our results show that domains beyond IDRs also play a role in TC formation and function and uncover domain-specific regulation of transcriptional and oncogenic programs.

Introduction

Cells exploit compartmentalization as a key regulatory mechanism to control non-equilibrium biochemical processes. While classic examples involve lipid-membrane-bound organelles, recent studies have proposed biomolecular condensation as a general principle of membrane-less compartmentalization (1). The resulting assemblies, known as biomolecular condensates, can transiently and specifically recruit biomolecules, enabling the spatial and temporal organization of biochemical reactions, signal transduction, and transcriptional regulation in the cell. Importantly, this dynamic organization allows cells to respond rapidly to environmental or signaling cues that regulate the stability of condensates. The aberrant biomolecular condensation of specific biomolecules due to mutations, mis-splicing, or overexpression has also been associated with different diseases and, in cancer, with the dysregulation of gene expression, the disruption of cellular homeostasis, and oncogenesis (2–5). Understanding how aberrant condensation plays a role in cancer can therefore highlight new opportunities for targeted therapeutic interventions.

Transcription factors (TFs) are nuclear proteins that regulate the initiation and maintenance of gene expression. Over the past decade, it has become clear that TF activity can arise through their capacity to assemble into high-order species, perhaps by undergoing a phase transition, leading to the formation of clusters or transcriptional condensates (TCs). Such TCs can then facilitate the recruitment of components of the transcriptional machinery, including RNA Polymerase II and the mediator complex (6–12). In addition to this important activity, TCs can also locally reorganize chromatin enhancer hubs, further enabling a precise control of transcription plasticity in cancer (13–19). From a molecular and a mechanistic perspective, it is well-established that the intrinsically disordered regions (IDR) of TFs play a central role in their ability to (I) cluster or condense and (II) recruit interacting molecules (12, 20–24). In contrast, the role of the structured DNA-binding domain (DBD), despite noteworthy *in vitro* work using ensemble measurements, has been less explored (25).

Nuclear receptors (NR) constitute a family of TFs that exclusively trigger transcription only upon exposure to specific steroid hormones (26–28). This tunability makes them ideal candidates to investigate how TCs form and contribute to transcriptional activity in a time-resolved fashion. Indeed, we and others have recently shown that the progesterone receptor (PR), the estrogen receptor (ER), the glucocorticoid receptor (GR), and the androgen receptor (AR) are all able to form condensates only upon binding to their relevant hormone (8, 10, 23, 29–33). While some of these studies have focused on the intrinsically disordered N-terminal domain (NTD) of NRs, which plays a key role in TC formation and transactivation, little is known about the role of the structured DBD in controlling the formation and maintenance of TCs. Most importantly, whether and how the IDR and the DBD synergistically control the properties of TCs and the transcriptional program of NRs remain unknown.

Here, we addressed these unresolved questions by using PR as a model system and by applying Single Particle Tracking (SPT), a powerful single-molecule-based method that enables the study of individual TFs and their interactions with biologically plausible components such as chromatin in a minimally invasive manner in living cells (22, 31, 33, 34). In addition, several SPT-associated data analysis methods have been proposed to infer different subpopulations based on the lateral diffusion of TFs. One prominent

example is the “Spot-On” analysis, which is based on the fitting of lateral displacements using a kinetic mixture model and allows distinguishing between subpopulations of molecules (e.g., “bound” vs “free”) (35). Recently, machine learning (ML) has also extended the possibilities of classifying diffusive behaviors of TFs using short trajectories. Indeed, using ML, we previously identified two different diffusing subpopulations of PR and, furthermore showed that the low mobility population corresponded to PR molecules diffusing within a condensate (29).

By relying on distinct PR mutants combined with SPT and different SPT analyses, including machine learning, we here unveiled the role of the structured DBD along with the intrinsically disordered NTD in the formation and regulation of TCs. Bulk transcriptomics and functional experiments further revealed that mutations in these two domains result in the expression of strikingly distinct sets of genes, both linked to oncogenic pathways. Taken together, our results provide additional evidence that domains beyond the IDRs contribute to biological condensate formation and function.

Results

Both the NTD and the DBD control the slowly diffusing population of PR

To examine how the NTD and the DBD control the spatiotemporal behavior of PR in living cells at the single molecule level, we generated two PR mutants in living MCF7 breast cancer cells: (I) one lacking the complete intrinsically disordered N-terminal domain of the PR (“ Δ NTD”) and (II) another with point mutations in the DNA binding region of PR that obliterate its affinity for DNA (for simplicity, we will refer to this mutant as “ Δ DBD”) (**Fig. 1A**). To avoid any side interactions between the mutants and the untagged endogenously expressed PR, we used a Knockdown and Rescue (pKAR) system where the endogenous receptor is knocked down using RNA interference and instead expresses a rescue RNAi-resistant PR (**Supplementary Fig. 1A-B**) (36).

To enable SPT, we designed PR constructs in which we introduced a SNAP tag at the N-terminus of WT PR, Δ NTD, or Δ DBD (**Fig. 1A-B**) (29, 37). We then performed SPT using single-molecule sensitive highly inclined thin illumination (HILO) of SNAP-Janelia Fluor 549 tagged PR molecules in the cell nucleus (**Fig. 1C**). SPT videos were acquired after incubating the cells (WT, Δ DBD, or Δ NTD) with the PR ligand, a synthetic progesterone derivative R5020, at 10nM for 60 min. These experimental conditions were used since, in our previous report, they were shown to induce PR condensate formation (29). In order to visualize fast dynamic events such as DNA-binding, we recorded the SPT videos with a framerate of 15 ms. After identifying the centroid position of each single PR molecule with a localization precision of ~ 25 nm, we reconnected each position as a function of time to obtain individual trajectories that report on the diffusion of PR and its mutants.

We first analyzed the SPT trajectories by extracting the instantaneous diffusion coefficient (D_{2-4}) by linearly fitting the 2nd-4th points of the mean-square displacement (MSD) curve for each trajectory and generated distributions of all the D_{2-4} values from hundreds of trajectories from multiple cells (**Fig. 1D-E**, **Supplementary Fig. 1C**). Our results showed an increase of D_{2-4} values for both mutants compared to WT, with median values of $D_{2-4} = 0.024 \mu\text{m}^2/\text{s}$ (WT), $0.132 \mu\text{m}^2/\text{s}$ (Δ DBD), and $0.078 \mu\text{m}^2/\text{s}$ (Δ NTD). Moreover, compared to the WT histogram, which largely shows a single distribution

centered around low D values, both mutants exhibit binomial distributions, suggesting the existence of different mobility sub-populations, in particular, the emergence of an increased diffusing population. Thus, these results reveal that compared to WT, both the Δ DBD and Δ NTD mutants exhibited increased diffusion, indicating that both the DBD and the NTD influence the low mobility of PR molecules.

To extract quantitative information on these apparently different subpopulations in the lateral diffusion of PR and the mutants, we applied the “Spot-On” analysis. In our case, the data was best fitted to a two-population model, which we further considered as “free” and “bound” sub-populations (35). Remarkably, the bound fraction extracted for WT is $77.9 \pm 0.6\%$, which is significantly reduced to $42.3 \pm 2.3\%$ and $48.1 \pm 1.1\%$ for Δ DBD and Δ NTD, respectively (**Fig. 1F**, **Supplementary Fig. 1D**). Moreover, the reduced bound sub-population of the PR mutants (Δ DBD or Δ NTD) showed an increased mobility compared to WT, as indicated by the diffusion coefficients of the bound fraction: $0.018 \pm 0.001 \mu\text{m}^2/\text{s}$, $0.042 \pm 0.004 \mu\text{m}^2/\text{s}$, and $0.034 \pm 0.002 \mu\text{m}^2/\text{s}$ for WT, Δ DBD, and Δ NTD, respectively (**Fig. 1G**). Altogether, D_{2-4} and Spot-On analysis indicate that both the DBD and the NTD control the low-mobility population and the bound fraction of PR.

Deep learning single-trajectory analysis identifies the NTD and DBD-dependent PR diffusion within condensates

The relatively short duration of our SPT trajectories represents a challenge in characterizing the dynamics of the PR. To gain further resolution in the available data, we use STEP, a deep-learning-based method to predict the anomalous diffusion exponent and diffusion coefficient for every frame of an input trajectory and extract segments of constant diffusive properties (38) (**Supplementary Fig. 2**). We used the value of anomalous exponent α (see Materials and Methods for details) as a measure of diffusion behavior, with $\alpha=1$ corresponding to Brownian diffusion and $\alpha < 1$ to anomalous diffusion (*i.e.* subdiffusion). We generated 2D histograms of the anomalous exponents and diffusion coefficients for each of the detected segments of WT and the two mutants (Δ DBD and Δ NTD). In the case of the WT, the 2D histogram shows one main cluster centered around D values of $0.1 \mu\text{m}^2/\text{s}$ and exhibiting α values below 0.5 (**Fig. 2A**), which we previously identified as a diffusion within a PR condensate (29). In contrast, the two PR-mutants show a shift toward higher D and α values, concomitant with an overall reduction in the fraction of the low-mobility component. To further uncover these changes in the distributions, we computed the difference between WT and Δ DBD/ Δ NTD probability distribution functions across the plane (α , D), showing a remarkable shift in the distributions toward higher values in both variables (**Fig. 2B**).

Next, we clustered the segments of each distribution by means of the k-clustering algorithm and obtained two different clusters that we associate with diffusion inside and outside of condensates (**Supplementary Fig. 3**) (29). By computing the percentage of segments in each population, we obtained that WT contains $84.5 \pm 0.3\%$ and $15.5 \pm 0.3\%$ of diffusion inside and outside of condensates, respectively (**Fig. 2C**). Importantly, Δ DBD ($51.5 \pm 5.3\%$) and Δ NTD ($60.1 \pm 2.6\%$) show a remarkable reduction in the population that diffuse inside condensates, indicating that both protein domains contribute to the formation of PR condensates.

Finally, we computed the angles between successive steps of each classified segment of the two populations and the three constructs. We plotted the distribution of the angles for each population, showing that while the population of diffusion outside of condensates has an isotropic angle distribution, the population of diffusion within condensates displays an expected increase at 180° (**Fig. 2D**) (29). Taken together, STEP identifies a (I) population of PR molecules diffusing within condensates and (II) that this population is strongly dependent on both the DBD and the NTD.

NTD and DBD distinctly regulate the transcription programs

Our results so far indicate that the NTD and the DBD contribute similarly to the capacity of the PR to form condensates, but do not provide information on how they influence the transcriptional activity of the receptor. To this end, we performed mRNA-sequencing in cells expressing WT and the two PR mutants (Δ DBD or Δ NTD) under progesterin exposure. A principal component analysis (PCA) of mRNA-sequencing data showed a distinct clustering of the results obtained for WT, Δ DBD, and Δ NTD before and after R5020 induction (**Supplementary Fig. 4**). Upon progesterin treatment, we identified 2,293 differentially expressed genes (DEGs) in cells expressing WT, 2,418 DEGs in cells expressing Δ NTD, and only 1,165 DEGs in Δ DBD cells, reflecting a markedly attenuated response to progesterin. In total, this analysis identified 3,291 DEGs, including 773 DEGs (23.5%) that were found in all three cases (WT, Δ DBD and Δ NTD) (**Fig. 3A**). In addition, WT and Δ NTD shared additional 839 DEGs, bringing the total overlap between the two to 1,612 (52%) (**Fig. 3A**). Similar comparison between WT and Δ DBD cells yielded 77 additional DEGs for a total overlap of 850 (32.6%), and 123 DEGs when comparing Δ NTD and Δ DBD cells, for a total of 896 DEGs (33.3%) (**Fig. 3A**). Furthermore, 192 genes (16.5% of the total 1,165 DEGs) were uniquely regulated in Δ DBD cells, while a larger set of 683 genes (28.2% of 2,418 DEGs) was uniquely regulated in the Δ NTD cells, suggesting a more prominent role for the NTD in fine-tuning transcriptional outputs. Moreover, 604 genes (26.3% of 2,293 DEGs) of unique genes were expressed in WT cells (**Fig. 3A**). Overall, these results indicate that inactivating DNA binding results in more profound changes in the transcriptional response to progesterin and suggest that Δ NTD and Δ DBD may affect a distinct set of cellular processes.

To get a more granular view of processes differentially affected by the two mutants, we performed Gene Ontology analysis (**Fig. 3B**). The genes uniquely affected in Δ NTD cells were significantly enriched in developmental as well as fundamental cellular processes, such as cell adhesion, cellular localization, and RNA metabolic pathways, which are frequently deregulated during oncogenesis. This suggests that the NTD is critical for the optimal expression of PR-dependent genes that help maintain proper cellular programs. In comparison, genes uniquely regulated in Δ DBD cells were enriched for molecular functions related to signal transduction and metabolic processes. Meanwhile, genes uniquely regulated in WT cells were enriched for cellular processes, such as cell migration, cell motility, cell organization, and negative regulation of metabolic processes and pathways, which were largely absent from both mutants. These findings suggest that mutations in either of the domains (NTD and DBD) regulate a distinct set of cellular processes via controlling the expression levels of unique sets of genes. Moreover, disruption of either the NTD or the DBD compromises PR's ability to drive expression of genes associated with oncogenic cell migration in MCF7, a breast cancer cell line, highlighting that both domains play a role in PR-mediated tumorigenic programs.

To further validate the transcriptional differences in cells expressing WT and the two mutants, and examine their potential participation in tumorigenic programs, we selected a subset of representative genes that are both differentially expressed (**Fig. 3C**) and associated with key tumor-related processes (39–50). This panel included genes involved in epithelial-mesenchymal transition (*SNAIL1*), developmental regulation (*PAX6*, *SOX18*), chromatin remodeling (*PHF19*), DNA damage response (*CHEK2*), cell adhesion (*CADMI*), stress response (*HSPA9*, *HMOX1*), cellular metabolism (*GLUD1*), purinergic signaling (*P2RX4*), and transcriptional regulation (*RIBC1*, *LARP7*) (39–50). We validated the differential gene expression using RT-qPCR (**Fig. 3D**). Remarkably, *SNAIL1*, *PAX6*, *SOX18*, and *PHF19* were uniquely up-regulated only in WT cells upon progesterin-induction (**Fig. 3D**). In contrast, *CHEK2*, *CADMI*, *HSPA9*, and *GLUD1* were only uniquely up-regulated in Δ NTD cells, highlighting a distinct transcriptional response associated with deletion of the PR activation domain (**Fig. 3D**). Additionally, *P2RX4* and *RIBC1* were down-regulated in Δ NTD cells, while both genes were up-regulated in WT cells. Progesterin-induced expression of *LARP7* and *HMOX1* remained unchanged between WT and Δ NTD cells. Importantly, none of these genes showed significant changes in expression in Δ DBD cells. Collectively, these observations highlight the potential of the NTD domain to fine-tune transcription regulation to maintain the condensate landscape for key oncogenic processes.

The NTD and DBD domains are required for PR-driven oncogenic processes

Our DEGs, as well as SPT analysis, indicate that both the NTD and the DBD are required for the PR TCs formation and regulate distinct cellular processes. To functionally validate the results, we examined phenotypes of MCF7 breast cancer cells expressing WT, Δ NTD, or Δ DBD. Because many PR-regulated genes disrupted in the domain-depleted mutants (Δ NTD or Δ DBD) are linked to cell proliferation, we first assessed progesterin-induced proliferation in MCF7 cells expressing WT, Δ NTD, or Δ DBD. Progesterin induction significantly increased the proliferation in WT cells, whereas this effect was significantly reduced in both Δ NTD and Δ DBD cells (**Fig. 4A**). Consistent with the gene expression followed by the Gene Ontology enrichment analysis, we then evaluated progesterin-stimulated cell migration using a transwell migration assay (48). MCF7 cells expressing Δ NTD or Δ DBD showed markedly impaired migration compared to WT (**Fig. 4B, left; Supplementary Fig. 5**) and by acetic acid-based extraction (**Fig. 4B, right**). Together, these results demonstrate that both the NTD and the DBD of PR are required for the establishment of PR-dependent transcriptional programs that promote oncogenic processes such as cell proliferation and migration.

Discussion

Here, we have presented a single-molecule biophysical study dissecting the contribution of distinct PR protein domains to transcriptional condensate formation in cancer cells. SPT combined with trajectory analysis revealed reduced low-mobility populations and a diminished bound fraction in both Δ NTD and Δ DBD mutants, as extracted from MSD and Spot-On analysis, respectively. STEP, a deep-learning SPT analysis pipeline, revealed a two-population-based behavior of PR in the nucleus. Based on the mobility and anomalous exponent clustering by STEP, we identified the anomalous diffusion and the low mobility population as the population belonging to the diffusion within

condensates (29). Importantly, condensate formation was strongly dependent on both the NTD and DBD, indicating that these domains regulate condensate assembly and stability.

We also assessed the contribution of PR protein domains on global gene expression by performing bulk mRNA-sequencing. Remarkably, breast cancer cells expressing Δ NTD or Δ DBD-mutants displayed markedly distinct transcriptional profiles, with only a partial subset of genes shared between the two mutants, underscoring domain-specific regulation of transcription. Notably, the NTD mutant led to a significant effect on the transcription activity of the genes associated with oncogenic processes. In contrast, the DBD mutant resulted in the transcription alteration of the genes associated with metabolic pathways and cellular signaling functions. These transcriptional findings were validated by functional assays of cell migration and proliferation, confirming that disruption of either domain compromises PR-dependent cellular programs.

Based on our SPT and global gene transcription and functional assays, we propose a novel regulatory mechanism of TCs. It has been suggested that the concerted action between DBDs and IDRs might regulate gene transcription(51–53). However, how this regulatory interplay might act in living cells remains to be further studied. In recent years, with the application of live cell imaging, the interaction between IDR domains, via possible biocondensation, and interacting partners is emerging as a complementary mechanism for TF selectivity. Indeed, using single-molecule imaging, a report suggests that the TF specificity is jointly determined by the transient interactions between the unstructured domains of TF with the heterologous DBDs, rather than being driven solely by DNA binding, which has traditionally been considered the main determinant of TF specificity. However, in this study, they did not observe evidences for condensate formation (54). In the same line, a more recent report pointed out that the TFs' binding stability and genomic preferences are mainly governed by unstructured domains of TFs(55). Altogether, these findings support the concept that unstructured domains of TFs transiently maintain the interaction with the structured DBD to dictate the gene regulation.

In the case of PR, our previous study established that the PR forms condensates (29). Here, we compared Spot-On together with STEP and obtained similar populations of low and higher mobility in WT and in both mutants, suggesting that both types of analysis are retrieving the same biophysical mechanism, namely PR condensates. Importantly, when the NTD is depleted, there is still a 60.1% fraction of PR molecules diffusing within a condensate, suggesting that the prominent function of structured DBD is an essential regulator of condensate formation. In the case of the DBD mutant, 51.5% of PR molecules still diffuse within a condensate, showcasing the canonical contribution of IDR in regulating condensate formation. These observations support the notion that the formation of functional PR condensates is driven by transient interactions between unstructured NTD with structured DBD domains, as suggested in the above-mentioned study (54). In this case, and since Δ DBD is only with specific mutations in PR, the ability of DBD to contribute to the condensate formation by maintaining the transcript interaction with IDR is strikingly affected, and also PR-dependent DEGs are significantly compromised. However, we still observed that Δ DBD mutants could not completely abolish the PR-dependent gene expression. The plausible reason for this could be that our Δ DBD mutants still could interact with other TFs and related transcription machinery, as also reported for other human TFs(54). Moreover, we could not ignore that RNA flux at the s transcriptional site could also contribute to the stability and transcriptional activity of TCs (56, 57). Further studies would be required to address the (I) DNA-binding sites

and (II) IDR-IDR interactions promoted by aberrant PR condensates. (III) how RNA-flux contributes to the stability and transcriptional activity of PR TCs. Overall, the differential gene expression profile and the SPT data point to an IDR and DNA-binding dependent mechanism in mediating functional PR transcriptional condensates in breast cancer.

Altogether and to the best of our knowledge, our work provides the first systematic and comprehensive analysis of the distinct roles of PR domains in regulating condensate formation and transcription program. While numerous studies have highlighted the contribution of other nuclear receptor IDRs to TCs, often using artificial gene arrays or focusing solely on IDR manipulation, we establish a tunable condensate model controlled by hormone stimulation in breast cancer cells. By examining both the DBD and the NTD of PR, we uncover a prominent role for the structured DBD in condensate formation and demonstrate how each domain shapes transcriptional output. Based on our observations, we propose that IDRs and structured DBDs jointly affect TCs formation and function, opening a new paradigm in the field of biocondensates.

Materials and Methods

Plasmids. The original plasmid expresses the PR isoform-B under a tetracycline-controllable promoter (TetOff system, Clontech) and a SNAP-tag at the N-terminal domain of the GFP (29). Taking this plasmid as the backbone, we introduced a shRNA-based mutation, thereby knocking down endogenous PR expression. We engineered the wild-type (WT) PR and mutants by either mutating the DBD-PR and deleting the NTD-PR (referred to as Δ DBD or Δ NTD in the manuscript).

Cloning. The original plasmid (pGFP-PRB plasmid) expresses the PR isoform-B under a tetracycline-controllable promoter (TetOff system, Clontech) and a SNAP-tag at the N-terminal domain of the GFP (29). This plasmid was used to insert the shRNA sequence (GGGACAACATAATTATTTGTG) within the exon 3 of the PR-B coding gene, and was confirmed by sequencing. The inhibition activity of the inserted shRNAs sequence was confirmed by the reduced level of PR-B (**Supplementary Fig. 1A**). To generate the catalytically inactive DBD domain (Δ DBD) of the pGFP-PRB plasmid, the potential set of amino acids (G585 to E585, S586 to G586, C587 to A587, V589 to A589, A604 to T604, R606 to W606, R637 to A637, K638 to A638) were mutated (58–60). Whereas, to generate the Δ NTD, the gene fragment (Integrated DNA Technologies) was synthesized by digestion with Bsu36 (R0524S, NEB) and PstI-HF (R3140S, NEB) restriction enzymes, followed by purification and ligation using T4 DNA ligase (M0202S). The final plasmid was confirmed by Sanger sequencing. To generate the Δ NTD pGFP-PRB plasmid, the pGFP-PRB shRNA1 plasmid was cut with BsrGI (R3575S, NEB), Bsu36I (R0524S, NEB), and SrfI (R0629S, NEB) restriction enzymes. After gel purification (QIAquick Gel Extraction Kit), the 5827 and 1044bp fragments were preserved. The following linkers were cut with BsrGI (R3575S, NEB) and Bsu36I (R0524S, NEB) with prior annealing of the 2 oligos (Integrated DNA Technologies) with Buffer 4 from NEB at 95°C for 5 minutes:

Linker:

5'-

CGAGCTGTACAAGTCCGGACTCAGATCTCGAGCGTTGCCTGAGGCCGGAT-3'

5'-

ATCCGGCCTCAGGCAACGCTCGAGATCTGAGTCCGGACTTGTACAGCTCG-3'

Next, a gene fragment without the NTD domain (Integrated DNA Technologies) was digested with BglII (R0144S, NEB) and PstI-HF (R3140S, NEB), ligated with T4 DNA ligase (M0202S). The final plasmid was also confirmed by Sanger sequencing.

Cell culture and electroporation. MCF7 Tet-off cells (Clontech, Cat No. 631154) were grown on Dulbecco Modified Eagle Medium (DMEM) high-glucose media supplemented with 10% Tet-free Fetal Bovine Serum, 2mM L-glutamine, 1 mM sodium pyruvate, and 100 U/ml penicillin-streptomycin. Cells were electroporated simultaneously with the corresponding vectors using the Amaxa Nucleofector Electroporation System. After one week of cell culture, cells were selected under 0.6 μ g/ml Puromycin to enrich for electroporated cells and then sorted in single-cell wells using GFP as a marker, in order to generate a stable cell line.

Hormone stimulation and SNAP labeling for SPT. 16 hours before hormone stimulation (10nM R5020), the cell medium was changed to Phenol Red-free DMEM

media supplemented with 10% charcoal-treated FBS Serum, 2mM L-glutamine, 1 mM sodium pyruvate, and 100 U/ml penicillin-streptomycin; from now on abbreviated as “charcoalized white DMEM”. We changed the medium as standard DMEM medium contains Phenol Red, which has interfering estrogenic activity with our experiments. The Janelia Fluor 549 dye coupled to the SNAP substrate was kindly provided by Luke Lavis (Janelia Farm, Ashburn, Virginia, USA). Cells were incubated with 10 nM for SPT in charcoalized white DMEM for 30 min at 37°C. Subsequently, the cells were washed three times with PBS, and then placed back in the incubator in charcoalized white DMEM for a one-hour washout at 37°C. This step was repeated two times. After the JF549 SNAP labeling, hormone stimulation was done using 10 nM R5020 (Promegestone) solubilized in Ethanol. The R5020 incubation time for SPT was 60 min.

Western Blot. Cell lysates were resolved on SDS-polyacrylamide gels, and the proteins were transferred to Hybond-ECL nitrocellulose membranes (Amersham). Membranes were blocked with TBS-0.1% Tween20 (TBS-T) with 5% of skimmed milk, incubated for 1h at room temperature with a primary antibody (For antibody PR, sc-166169 from Santa Cruz Biotechnology, and antibody for tubulin, T9026 from Sigma, St. Louis, MO, USA, antibody for GFP, A-11122, Invitrogen), diluted in TBS-T with 2.5% skimmed milk. After three washes with TBS-T, membranes were incubated for 1h at room temperature with horseradish peroxidase-conjugated secondary antibodies (GE Healthcare, Chicago, IL, USA). Antibody binding was detected by chemiluminescence on an LAS-3000 image analyzer (Fuji PhotoFilm, Tokyo, Japan).

Experimental setup for SPT. SPT was performed in a Nikon N-STORM 4.0 microscope system for localization-based super-resolution microscopy, equipped with a TIRF 100x, 1.49 NA objective (Nikon, CFI SR HP Apochromat TIRF 100XC Oil). The sample was illuminated by a continuous 561 nm laser line with a power of 30 mW before the objective in HILO configuration. The emission fluorescence was projected into an EM-CCD Andor Ixon Ultra Camera at a framerate of 15 ms. During imaging, the temperature was kept at 37°C by an incubation chamber.

SPT data analysis: SPT trajectories were reconstructed as explained in our previous work (29). Briefly, single fluorophore localizations were detected using Trackmate (61) with sub-pixel localizations. The detected localizations were reconnected using the algorithm Linear Assignment Problem (LAP), and the minimum trajectory length for further analysis was 10 frames. To calculate the time-averaged MSD, we applied the following formula (29):

$$\text{tMSD}(\Delta) = \frac{1}{T-\Delta} \sum_{i=1}^{T-\Delta} ([x(t_i + \Delta) - x(t_i)]^2 + [y(t_i + \Delta) - y(t_i)]^2) \quad [1]$$

where x and y are the positions of the fluorophore, t_i are the time steps, Δ is the time interval, and T the total duration of the trajectory. To extract the instantaneous diffusion coefficient ($D_{2.4}$), we fit the 2nd and 4th point ($\Delta=2-4$) of the MSD curve with the following formula:

$$\text{tMSD} = 4D\Delta + \text{offset} \quad [2]$$

where D is the diffusion coefficient. For the D_{2-4} distributions, extracted D_{2-4} values below $0.0001 \mu\text{m}^2/\text{s}$ were considered as fitting errors and discarded.

For the Spot-On Analysis, the MATLAB software version was downloaded from their dedicated website: <https://github.com/elifesciences-publications/spot-on-matlab>. The trajectories were fitted to a two-population kinetic model with the following parameters: $dZ=0.7 \mu\text{m}$, GapsAllowed=1, TimePoints=8, JumpsToConsider=4, ModelFit=2, NumberOfStates=2, $D_{\text{Free_2State}} = [0.5; 25]$; $D_{\text{Bound_2State}} = [0.0001; 0.05]$.

mRNA-sequencing samples, library preparation, and sequencing. RNA-sequencing experiments were performed as biological triplicate for the cells expressing WT-PR, ΔNTD and ΔDBD mutants before and after the 10nM R5020 treatment for 6 hours. RNA from WT-PR and ΔNTD and ΔDBD cells was extracted using the Quick-RNA kit (ZYMO Research; R1055) according to the manufacturer's instructions. Purified RNA was analyzed on a Bioanalyzer using an RNA Pico assay chip. PolyA plus RNA mRNA libraries were prepared using TruSeq Stranded RNA Library Prep Kit (Illumina) and sequenced using Illumina HiSeq 2500.

mRNA-sequencing data processing and analysis. Bulk mRNA-sequencing was performed on MCF7 cells expressing WT-PR, ΔNTD and ΔDBD mutants before and after R5020 treatment. Each condition was analyzed in triplicate, yielding a total of 18 FASTQ files. Quality control of raw sequencing reads was conducted using FastQC (version 0.72). Reads were aligned to the human reference genome (hg38) using HISAT2 (version 2.1.0) implemented on the Galaxy platform (version 7). Aligned reads were quantified using featureCounts (version 2.0.1), and differential gene expression analysis was performed with DESeq2 (version 2.11.40.6). Genes were considered significantly differentially expressed if they exhibited a Benjamini-Hochberg-adjusted p-value < 0.05 and an absolute $\log_2$ fold change greater than 0.58. For data visualization and downstream analysis, the R packages tidyverse (version 2.0.0), scales (version 1.4.0), and ggplot2 (version 3.5.2) were used.

Gene Ontology (GO) and Regulatory Gene Set Analysis. GO Annotation performed using the Gene Set Enrichment Analysis (GSEA, <http://software.broadinstitute.org/gsea/index.jsp>) database v5 (62, 63). The significant cut-off p-value and FDR q-value < 0.05 . Plots were done with the use of Prism (GraphPad Prism 10.0.3 for MacOS) and R4.0.2.

RNA extraction and RT-qPCR. Purified RNA from WT-PR, ΔNTD and ΔDBD expressing cells was used for reverse transcription of $1\mu\text{g}$ RNA, which was performed using an OneScript® RT Mix for qPCR with gDNAOut (OZYME; OZYA012-100) according to the manufacturer's instructions. Complementary DNA was quantified by qPCR using ONEGreen FAST qPCR Premix (OZYME; OZYA008-1000) at the PCR analyzer Applied Biosystems (7500). For each gene product, relative RNA abundance was calculated using the standard curve method and expressed as relative RNA abundance after normalizing against the human GAPDH gene level as described (64, 65). All gene expression data generated by RT-qPCR represented the average and $\pm$ SEM of at least three biological replicates. Primers used for RT-qPCR are listed in **Table S1**.

Cell Proliferation. Cell proliferation was performed using the BrdU (5'-bromo-2'-deoxyuridine) cell proliferation assay: WT-PR, Δ NTD and Δ DBD expressing cells (0.3×10^3) were seeded in 96-well plates with complete DMEM media for 36 hours, incubated at 37 °C with 5% CO₂. Next, cells were washed with PBS and replaced with “charcoalized white DMEM” for 16 hours before R5020 induction for 24 hours. The cell proliferation ELISA BrdU (5'-bromo-2'-deoxyuridine) colorimetric assay (Roche, 11647229001) was performed as per the manufacturer's instructions. The experiments were performed on at least eight biological replicates.

Cell Migration. Migration method was adopted from Justus et al. 2023 (66). WT-PR Δ NTD and Δ DBD expressing cells were seeded at a density of 2×10^6 cells per P10 dish in complete DMEM medium and incubated at 37 °C with 5% CO₂. After 24 hours, the medium was removed, cells were washed with PBS, and the culture medium was replaced with “charcoalized white DMEM” before 10 nM R5020 induction for 24 hours. Cells were washed with pre-warmed PBS, detached, and washed again before being resuspended in migration buffer (DMEM supplemented with 10 mM HEPES and 0.1% BSA, pH 7.4). Cell viability was assessed, and the suspension was adjusted to a final concentration of 0.7×10^6 viable cells/mL. Migration assays were conducted using 24-well Transwell inserts with 8 μ m pores (Corning-Falcon, ref: 353097). A total of 600 μ L of migration buffer containing 10% FBS as a chemoattractant was added to the lower chamber. Then, 100 μ L of the prepared cell suspension (containing 1×10^5 cells) was carefully added to the upper chamber. Plates were incubated overnight at 37 °C with 5% CO₂. The following day, non-migrated cells were gently removed from the upper surface of the membrane using a cotton swab. Migrated cells on the basal side were fixed in 4% paraformaldehyde and stained with crystal violet. Membranes were imaged under a light microscope at 10 $\times$ and 20 $\times$ magnifications. For quantitative analysis, 30% acetic acid was added to elute the dye, and the absorbance was measured at 565 nm using a CLARIOstar (BMG Labtech) plate reader.

Machine Learning: The pipeline used to characterize the diffusive properties of the various constructs is mainly based on STEP (38). Given an input SPT trajectory, STEP uses state-of-the-art deep learning methods to predict the diffusive properties (diffusion coefficient and anomalous diffusion exponent) for every frame. This allows us to maximize the temporal resolution at which we analyze the trajectories, unveil changes of diffusion properties, and identify different diffusive states (67, 68). We post-process the frame-wise raw diffusive properties prediction with a kernel change detection algorithm to obtain segments of constant properties (69). This allows us to identify the diffusive changepoints as well as generally improve the analysis accuracy (38).

For this work, we have performed a series of adjustments over the original STEP architecture and have trained a new model using the tools provided by the authors (38). More precisely, we train a STEP model only with noiseless fractional Brownian motion trajectories, as opposed to using multiple anomalous diffusion methods with various noise levels. This ensures that the lowest diffusion coefficient predicted in the experiment matches the localization noise and removes all the prediction artifacts in trajectories with low diffusion coefficients. Additionally, we sample a random trajectory length between 20 and 200 for every batch that we use to take a random slice of all the trajectories therein. This makes the model robust to trajectory length variation while avoiding padding and keeping the training loop as efficient as possible. Finally, based on the findings of (67), we improve the original STEP by considering as input the motion displacements instead of using the raw coordinates. To do so, we compute the displacements for every

dimension of an input trajectory and take the logarithm of their absolute value. Then, we compute the angle between consecutive displacements, setting the first one to zero by convention. We feed the STEP model this trajectory representation, which is rotationally invariant and preserves all the physically relevant information. To perform the analysis of the predictions, we first take all the trajectory segments for the three constructs, aggregate them in a dataset, and run a k-means clustering algorithm over it. We choose to do the clustering with all segments to avoid any potential bias some of the constructs may have, and improve the statistical relevance of the analysis with more data. We find that a separation into two clusters gives the best qualitative result.

To reveal changes in the different mutants' distribution, we first constructed joint probability distributions of the diffusion coefficient D and anomalous exponent α for each of them by estimating normalized two-dimensional histograms in the (D, α) plane. These histograms were converted into discrete probability masses by accounting for the bin areas. Pairwise differences between conditions were then quantified using a symmetric, normalized contrast,

$$\Delta P = \frac{P_1 - P_2}{P_1 + P_2},$$

computed pointwise over the (D, α) space. This definition yields a bounded measure $[-1, 1]$ that highlights relative enrichments or depletions between conditions while suppressing regions of low overall probability. The resulting difference maps provide a direct visualization of how the distributions shift across dynamical regimes.

Quantification and statistical analysis. For RT-qPCR experiments, a Two-tailed unpaired Student's t-test was used to determine statistical significance between the groups. For cell proliferation and migration experiments, the significance between groups was calculated by the Wilcoxon-Mann-Whitney test. Plots and indicated statistical analysis were done with the use of Prism (GraphPad Prism 10.0.3 for MacOS) unless otherwise stated. If exact p-values are not shown or indicated in the legend, then p-values are represented in all figures as follows: *, p-value ≤ 0.05 ; **, p-value ≤ 0.01 ; ***, p-value ≤ 0.001 ; °, p-value > 0.05 .

Acknowledgments: We thank Gordon Hager (NIH) for providing the pGFP-PRB plasmid and Luke Lavis (Janelia Research Campus) for kindly providing the SNAP-Janelia Fluor 549 dye. We would like to thank the Advanced Light Microscopy Unit, the Genomics facility, and the Flow Cytometry Unit of the Center for Genomic Regulation (CRG, Barcelona) for their support. The research leading to these results has received the following funding: J.A.T.-P. and C.R. jointly from BIST-Ignite funding (PHASE-CHROM). J.A.T.-P. from the Government of Spain JdC-IJCI-2017-33160 and RYC2023-044870-I. M.G.-P. from the European Commission ERC Adv788546 (NANO-MEMEC), the Government of Spain (Severo Ochoa CEX2019-000910-S and PID2023-147711NB-I00), Fundació CELLEX (Barcelona), Fundació Mir-Puig, and the Generalitat de Catalunya through the CERCA program and AGAUR (Grant No. 2021 SGR 01450). B.R. and M.L. acknowledge support from MCIN/AEI (PGC2018-0910.13039/501100011033, CEX2019-000910-S/10.13039/501100011033, Plan Nacional STAMEENA PID2022-139099NB, project funded MCIN and by the "European Union NextGenerationEU/PRTR" (PRTR-C17.I1), FPI); Ministry for Digital Transformation and of Civil Service of the Spanish Government through the QUANTUM

ENIA project call - Quantum Spain project, and by the European Union through the Recovery, Transformation and Resilience Plan - NextGenerationEU within the framework of the Digital Spain 2026 Agenda; CEX2024-001490-S [MICIU/AEI/10.13039/501100011033]; Fundació Cellex; Fundació Mir-Puig; Generalitat de Catalunya (European Social Fund FEDER and CERCA program); Barcelona Supercomputing Center MareNostrum (FI-2023-3-0024); Funded by the European Union (HORIZON-CL4-2022-QUANTUM-02-SGA, PASQuanS2.1, 101113690, EU Horizon 2020 FET-OPEN OPTologic, Grant No 899794, QU-ATTO, 101168628), EU Horizon Europe Program (No 101080086 NeQSTGrant Agreement 101080086 — NeQST).). C.G.-C. and X.S. were supported by AGAUR (2021 SGR 476), AEI (PID2022-141816OB-I00), and the ERC (CONCERT, contract number 648201). G.M.-G. acknowledges support from the European Union (ERC Advanced Grant, QuantAI, No. 101055129). P.S. received funding from the French National Research Agency (ANR) Young Investigator grant (ANR-21-CE12-0010), Cancer Research Foundation (ARCPJA22021050003683), La Ligue Foundation for Cancer (Haute-Garonne, 285458), Centre national de la recherche scientifique (CNRS), Université de Toulouse, Toulouse, France. The views and opinions expressed in this article are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council - neither the European Union nor the granting authority can be held responsible for them.

Author Contributions: G.M.-G., M.L., M.F.G.-P., J.A.T.-P., P.S. supervised research; C.R.-A., J.A.T.-P., M.F.G.-P., P.S. designed research; L.R., D.S., C.R.-A., J.A.T.-P., L.E., L.I.L.C., S.N., C.G.-A., E.S., N.Y., P.S. performed research; L.R., B.R., G.M.-G., C.G.-A., E.S., J.A.T. N.Y., D.S., R.L.K., P.S. analyzed data; L.R., X.S., C.G.-C., M.L., G.M.-G., M.F.G.-P., P.S. and J.A.T.-P. wrote the paper.

Declaration of Interests

X.S. is founder and scientific advisor of Nuage Therapeutics.

Resource Availability

RNA sequencing data performed and used in this study have been deposited at GEO under the accession code GSE308408 (secure token to access is ojilyckqtrujuz) and will be publicly available on acceptance for publication. Further information and requests for resources and results should be directed to the corresponding authors, Maria Garcia-Parajo (maria.garcia-parajo@icfo.eu), Juan A. Torreno-Pina (ja.torreno@cnb.csic.es), and Priyanka Sharma (priyanka.sharma@ipbs.fr).

Figures

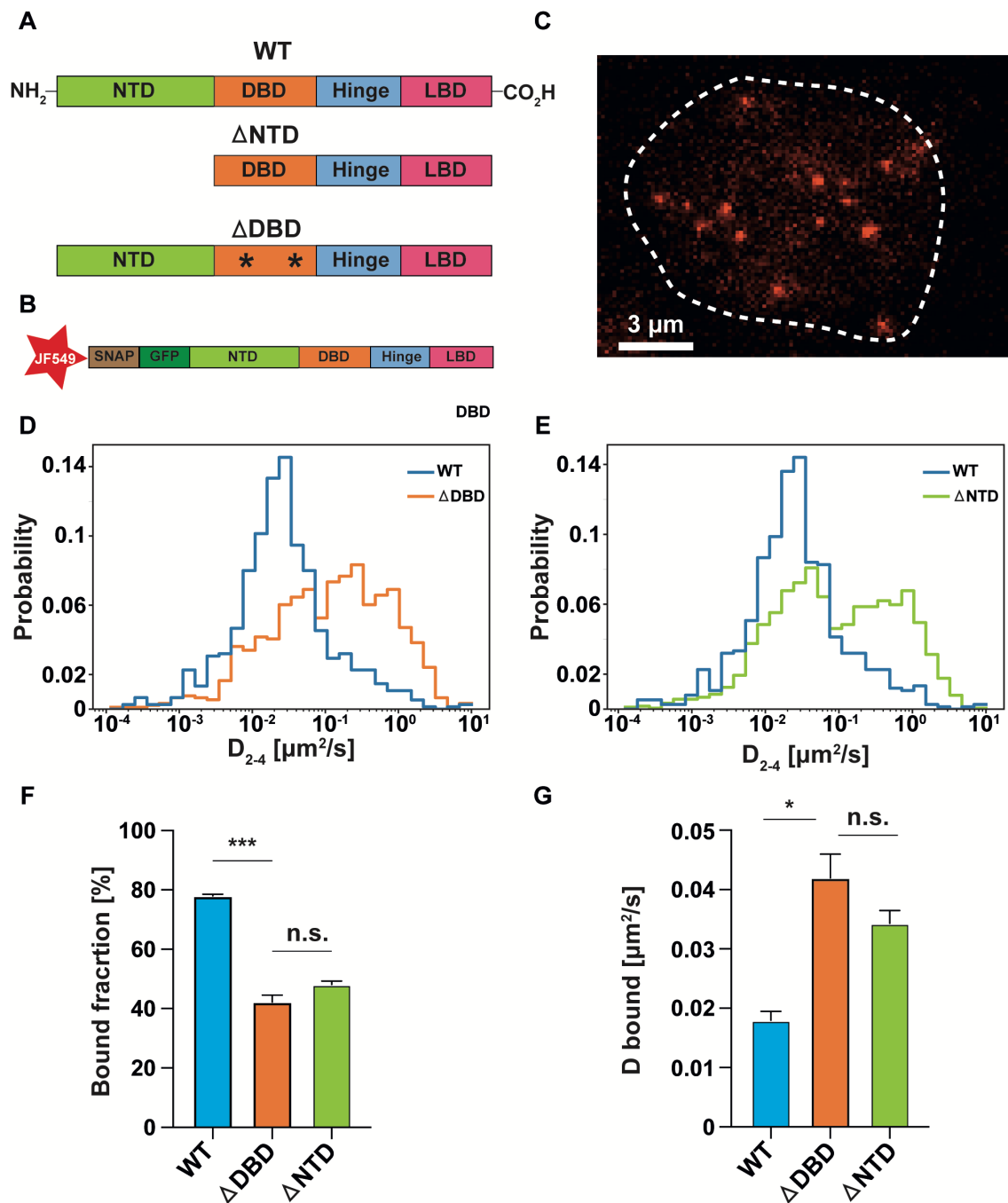


Figure 1: The lateral diffusion of WT and mutants (ΔDBD or ΔNTD) in the nucleus of living breast cancer cells. (A) Schematic representation of the different constructs (WT, ΔDBD, and ΔNTD) used throughout the study. **(B)** Schematic representation of the construct used for SPT in the nucleus. A SNAP-tag was introduced in the N-terminal domain of the GFP-PR. **(C)** Representative frame of the SPT video. Individual PR molecules labelled with Janelia Fluor 549 are visualized using red color coding in the nucleus (white outline) of MCF7 cells. **(D-E)** Normalized histograms of the instantaneous diffusion coefficients (D_{2-4}) for hormone-exposed WT, ΔDBD, and ΔNTD extracted from

the MSD curves for each single SPT trajectory. (F-G) Bound fraction and diffusion coefficients from the three constructs extracted from the Spot-On analysis. The error bars correspond to the standard error of the mean. Results of a one-way analysis of variance (ANOVA) test are shown as: n.s. for not significant, * for p -value < 0.05, and *** for p -value < 0.001. SPT data is extracted from at least 2 independent experiments, 8 cells, and 750 trajectories per condition.

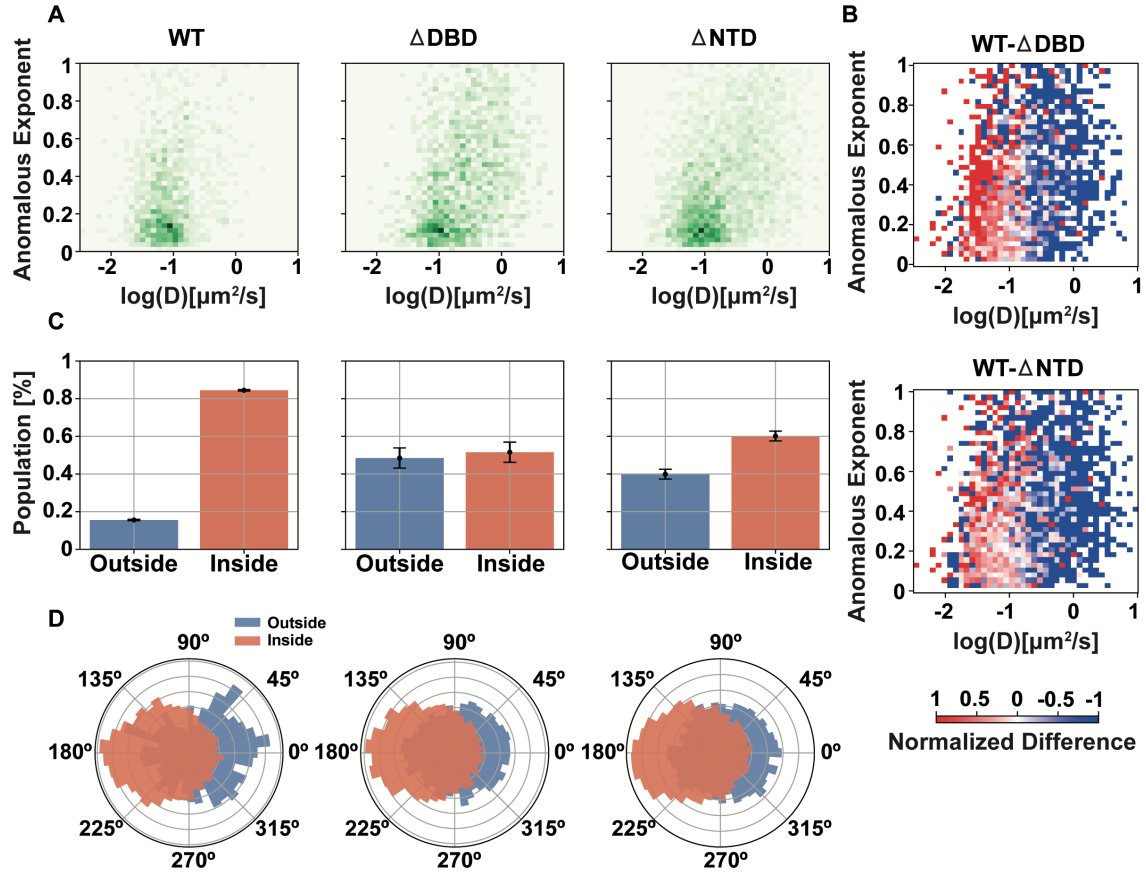


Figure 2. Deep learning analysis (STEP) reveals two sub-populations within TCs. (A) Histograms of the predicted diffusion coefficient and anomalous exponent for every segment of WT, Δ DBD, and Δ NTD found by means of STEP and a kernel changepoint detection algorithm. (B) Difference between the parameter distributions of the different constructs, calculated from the normalized joint histograms as shown in A. (C) Percentage of segments per population as extracted from a k-means clustering algorithm. (D) Angle distribution between consecutive segments for each sub-population and construct. The data used for STEP analysis has been extracted from at least 2 independent experiments, 8 cells, and 903 trajectories per condition.

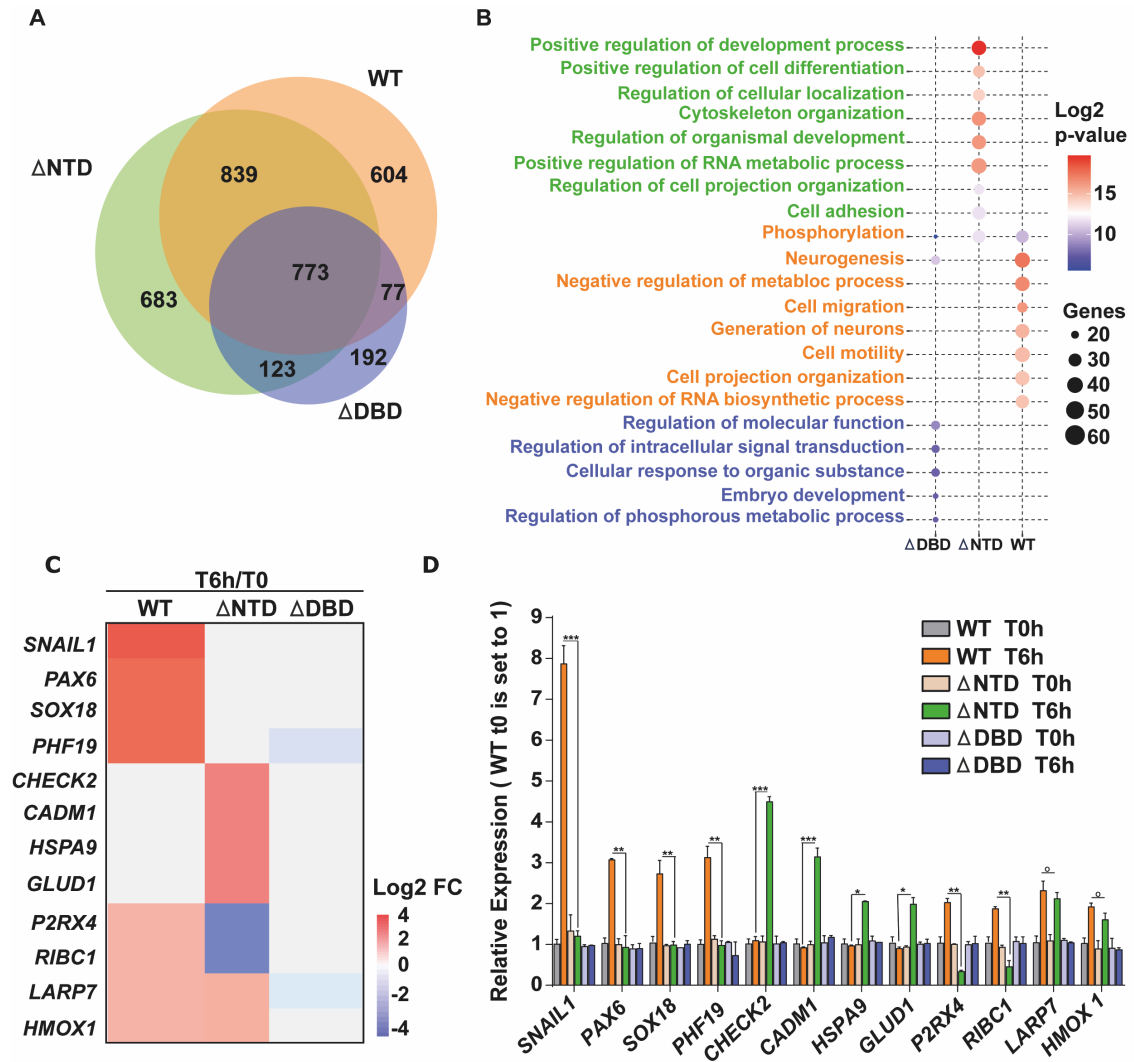


Figure 3: Progestin-induced transcriptional profiling reveals distinct regulatory gene programs for individual domains of the PR. (A). Venn diagram representing the overlap of differentially expressed genes (Fold change > 1.5 or < 1/1.5 and p value < 0.05) modulated by R5020 induction (10nM for 6 hours) in MCF7 cells specifically expressing the WT and mutants (ΔDBD and ΔNTD). (B). Representative plots of key gene sets for R5020-induced differentially expressed genes in WT, ΔDBD, and ΔNTD expressing MCF7 cells. Log2 p-values are indicated in color and size, showing the number of enriched genes. (C). Heatmap representing the log2 fold change in the expression of the representative gene set of “progestin-responsive genes” from the RNA-sequencing analysis performed in WT, ΔNTD, or ΔDBD MCF7 cells. (D) Quantitative RT-PCR validation in MCF7 cells expressing WT or mutants in the absence or presence of R5020. Changes in mRNA levels were normalized to GAPDH mRNA. Data represented as WT, as set to 1, mean ± SEM of four biological experiments.

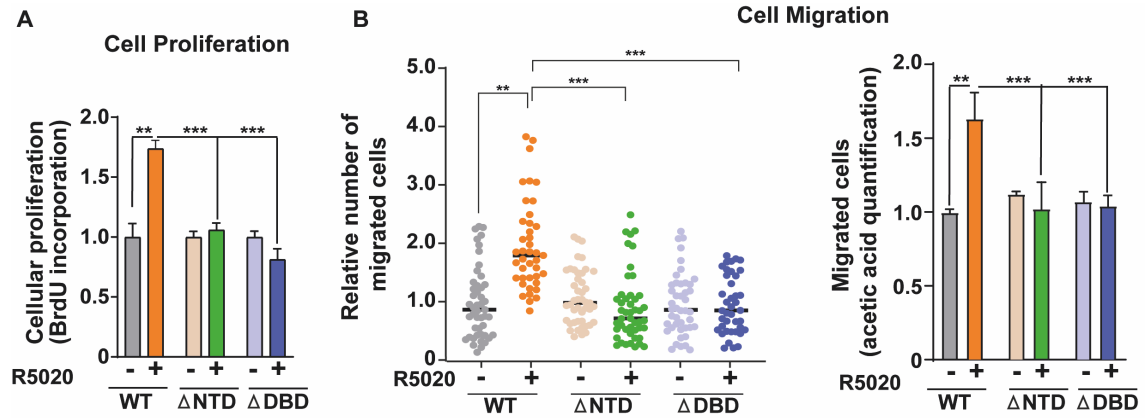


Figure 4: Both protein domains regulate oncogenic processes (A) Cell proliferation of MCF7 cells expressing the WT or the two mutants before and after the induction of R5020. Data represent median $\pm$ SEM of at least eight biological experiments. (B). Cell migration of MCF7 cells expressing WT or mutants (Δ NTD or Δ DBD) before and after the induction of R5020. Data represent median $\pm$ SEM of at least three biological experiments. (Left) Dot plot showing the relative number of migrated cells in the indicated condition. (Right) Bar plot representing the acetic acid quantification of the migrated cells. *p value < 0.05 ; **p value < 0.01 ; ***p value < 0.001 .

References:

1. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285–298 (2017).
2. S. Mehta, J. Zhang, Liquid–liquid phase separation drives cellular function and dysfunction in cancer. *Nat Rev Cancer* **22**, 239–252 (2022).
3. B. R. Sabari, Biomolecular Condensates and Gene Activation in Development and Disease. *Developmental Cell* **55**, 84–96 (2020).
4. B. Wang, *et al.*, Liquid-liquid phase separation in human health and diseases. *Signal Transduct Target Ther* **6**, 290 (2021).
5. Y. Shin, C. P. Brangwynne, Liquid phase condensation in cell physiology and disease. *Science* **357** (2017).
6. M. Stortz, A. Pecci, D. M. Presman, V. Levi, Unraveling the molecular interactions involved in phase separation of glucocorticoid receptor. *BMC Biol* **18**, 59 (2020).
7. B. R. Sabari, *et al.*, Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, 3958 (2018).
8. A. Boija, *et al.*, Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **175**, 1842–1855 e16 (2018).
9. W. K. Cho, *et al.*, Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
10. R. Mann, D. Notani, Transcription factor condensates and signaling driven transcription. *Nucleus* **14**, 2205758 (2023).
11. F. Jonas, Y. Navon, N. Barkai, Intrinsically disordered regions as facilitators of the transcription factor target search. *Nat Rev Genet* **26**, 424–435 (2025).
12. L. Ma, *et al.*, Co-condensation between transcription factor and coactivator p300 modulates transcriptional bursting kinetics. *Mol Cell* **81**, 1682–1697.e7 (2021).
13. R. Lee, *et al.*, CTCF-mediated chromatin looping provides a topological framework for the formation of phase-separated transcriptional condensates. *Nucleic Acids Research* **50**, 207–226 (2022).
14. Y. Shin, *et al.*, Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. *Cell* **175**, 1481–1491 e13 (2018).
15. B. A. Gibson, *et al.*, Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell* **179**, 470–484 e21 (2019).
16. T. Nozaki, *et al.*, Condensed but liquid-like domain organization of active chromatin regions in living human cells. *Sci. Adv.* **9**, eadf1488 (2023).

17. A. R. Strom, *et al.*, Condensate interfacial forces reposition DNA loci and probe chromatin viscoelasticity. *Cell* **187**, 5282-5297.e20 (2024).
18. J. Yang, *et al.*, MYC phase separation selectively modulates the transcriptome. *Nat Struct Mol Biol* **31**, 1567–1579 (2024).
19. J. Yang, *et al.*, Enhancer reprogramming: critical roles in cancer and promising therapeutic strategies. *Cell Death Discov.* **11**, 84 (2025).
20. Y. Shin, *et al.*, Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* **168**, 159-171 e14 (2017).
21. N. Schneider, *et al.*, Liquid-liquid phase separation of light-inducible transcription factors increases transcription activation in mammalian cells and mice. *Sci Adv* **7**, eabd3568 (2021).
22. S. Chong, *et al.*, Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361** (2018).
23. L. Chen, *et al.*, Hormone-induced enhancer assembly requires an optimal level of hormone receptor multivalent interactions. *Mol Cell* **83**, 3438-3456.e12 (2023).
24. Y. J. Kim, *et al.*, Light-activated macromolecular phase separation modulates transcription by reconfiguring chromatin interactions. *Sci Adv* **9**, eadg1123 (2023).
25. K. Shrinivas, *et al.*, Enhancer Features that Drive Formation of Transcriptional Condensates. *Molecular Cell* **75**, 549-561.e7 (2019).
26. M. Beato, R. H. G. Wright, F. L. Dily, 90 YEARS OF PROGESTERONE: Molecular mechanisms of progesterone receptor action on the breast cancer genome. *J Mol Endocrinol* **65**, T65–T79 (2020).
27. E. R. Weikum, X. Liu, E. A. Ortlund, The nuclear receptor superfamily: A structural perspective. *Protein Sci* **27**, 1876–1892 (2018).
28. F. Le Dily, *et al.*, Hormone-control regions mediate steroid receptor-dependent genome organization. *Genome Res* **29**, 29–39 (2019).
29. G. Muñoz-Gil, *et al.*, Stochastic particle unbinding modulates growth dynamics and size of transcription factor condensates in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2200667119 (2022).
30. J. Xie, *et al.*, Targeting androgen receptor phase separation to overcome antiandrogen resistance. *Nat Chem Biol* **18**, 1341–1350 (2022).
31. D. A. Garcia, *et al.*, An intrinsically disordered region-mediated confinement state contributes to the dynamics and function of transcription factors. *Mol Cell* **81**, 1484-1498 e6 (2021).
32. S. Basu, *et al.*, Rational optimization of a transcription factor activation domain inhibitor. *Nat Struct Mol Biol* **30**, 1958–1969 (2023).

33. K. Wagh, D. A. Stavreva, G. L. Hager, Transcription dynamics and genome organization in the mammalian nucleus: Recent advances. *Mol Cell* **85**, 208–224 (2025).
34. K. Wagh, *et al.*, Dynamic switching of transcriptional regulators between two distinct low-mobility chromatin states. *Sci. Adv.* **9**, eade1122 (2023).
35. A. S. Hansen, *et al.*, Robust model-based analysis of single-particle tracking experiments with Spot-On. *Elife* **7** (2018).
36. H. T. Ma, K. F. On, Y. H. Tsang, R. Y. C. Poon, An inducible system for expression and validation of the specificity of short hairpin RNA in mammalian cells. *Nucleic Acids Res* **35**, e22 (2007).
37. A. Gautier, *et al.*, An engineered protein tag for multiprotein labeling in living cells. *Chem Biol* **15**, 128–36 (2008).
38. B. Requena, *et al.*, Inferring pointwise diffusion properties of single trajectories with deep learning. *Biophys J* **122**, 4360–4369 (2023).
39. S. Kaufhold, B. Bonavida, Central role of Snail1 in the regulation of EMT and resistance in cancer: a target for therapeutic intervention. *J Exp Clin Cancer Res* **33**, 62 (2014).
40. Y. Kiselev, *et al.*, Transcription factor PAX6 as a novel prognostic factor and putative tumour suppressor in non-small cell lung cancer. *Sci Rep* **8**, 5059 (2018).
41. C. Dailey, R. B. Oshodi, C. Boull, A. Aggarwal, Expanding the clinical spectrum of SOX18-related Hypotrichosis-lymphedema-telangiectasia-renal defect syndrome. *Eur J Med Genet* **65**, 104607 (2022).
42. G. L. Brien, *et al.*, Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. *Nat Struct Mol Biol* **19**, 1273–1281 (2012).
43. P. Li, J. Sun, Y. Ruan, L. Song, High PHD Finger Protein 19 (PHF19) expression predicts poor prognosis in colorectal cancer: a retrospective study. *PeerJ* **9**, e11551 (2021).
44. I. Rainville, *et al.*, High risk of breast cancer in women with biallelic pathogenic variants in CHEK2. *Breast Cancer Res Treat* **180**, 503–509 (2020).
45. H. Li, J. Gao, S. Zhang, Functional and Clinical Characteristics of Cell Adhesion Molecule CADM1 in Cancer. *Front Cell Dev Biol* **9**, 714298 (2021).
46. B. Royer-Bertrand, *et al.*, Mutations in the heat-shock protein A9 (HSPA9) gene cause the EVEN-PLUS syndrome of congenital malformations and skeletal dysplasia. *Sci Rep* **5**, 17154 (2015).
47. S. Liu, Bioinformatics analysis identifies GLUD1 as a prognostic indicator for clear cell renal cell carcinoma. *Eur J Med Res* **29**, 70 (2024).

48. X. Qiao, C. Wang, J. Ma, The role and function validation of P2RX4 as a novel cancer biomarker in pan-cancer analysis. *Sci Rep* **15**, 11507 (2025).
49. X. Ji, H. Lu, Q. Zhou, K. Luo, LARP7 suppresses P-TEFb activity to inhibit breast cancer progression and metastasis. *Elife* **3**, e02907 (2014).
50. J. Fahrner, S. Wittmann, A.-C. Wolf, T. Kostka, Heme Oxygenase-1 and Its Role in Colorectal Cancer. *Antioxidants (Basel)* **12**, 1989 (2023).
51. G. Pei, H. Lyons, P. Li, B. R. Sabari, Transcription regulation by biomolecular condensates. *Nat Rev Mol Cell Biol* **26**, 213–236 (2025).
52. S. A. Lambert, *et al.*, The Human Transcription Factors. *Cell* **172**, 650–665 (2018).
53. L. F. Soto, *et al.*, Compendium of human transcription factor effector domains. *Molecular Cell* **82**, 514–526 (2022).
54. A. A. Abidi, *et al.*, Unstructured transcription factor interactions enable emergent specificity. *Science* eab6487 (2026). <https://doi.org/10.1126/science.aeb6487>.
55. J. Bugis, *et al.*, Beyond motif recognition: Specificity of human transcription factors in yeast. [Preprint] (2026). Available at: <http://biorxiv.org/lookup/doi/10.64898/2026.04.27.721015> [Accessed 10 May 2026].
56. F. Muzzopappa, F. Erdel, Beyond equilibrium: roles of RNAs in condensate control. *Current Opinion in Genetics & Development* **91**, 102304 (2025).
57. J. Liao, S. Y. Ahn, A. C. Obermeyer, Dynamics of synthetic transcriptional condensates emerge from RNA synthesis and degradation. [Preprint] (2026). Available at: <http://biorxiv.org/lookup/doi/10.64898/2026.05.03.722550> [Accessed 10 May 2026].
58. S. C. Roemer, *et al.*, Structure of the Progesterone Receptor-Deoxyribonucleic Acid Complex: Novel Interactions Required for Binding to Half-Site Response Elements. *Molecular Endocrinology* **20**, 3042–3052 (2006).
59. I. Quiles, *et al.*, Mutational Analysis of Progesterone Receptor Functional Domains in Stable Cell Lines Delineates Sets of Genes Regulated by Different Mechanisms. *Molecular Endocrinology* **23**, 809–826 (2009).
60. G. S. Takimoto, D. M. Tasset, A. C. Eppert, K. B. Horwitz, Hormone-induced progesterone receptor phosphorylation consists of sequential DNA-independent and DNA-dependent stages: analysis with zinc finger mutants and the progesterone antagonist ZK98299. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3050–3054 (1992).
61. J. Y. Tinevez, *et al.*, TrackMate: An open and extensible platform for single-particle tracking. *Methods* **115**, 80–90 (2017).
62. A. Subramanian, *et al.*, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545–15550 (2005).

63. M. J. Daly, *et al.*, PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics* **34**, 267–273 (2003).
64. P. Sharma, *et al.*, Arginine Citrullination at the C-Terminal Domain Controls RNA Polymerase II Transcription. *Mol Cell* **73**, 84-96 e7 (2019).
65. J. L. Villanueva-Cañas, *et al.*, Evolutionary analysis reveals the role of a non-catalytic domain of peptidyl arginine deiminase 2 in transcriptional regulation. *iScience* **27**, 109584 (2024).
66. C. R. Justus, M. A. Marie, E. J. Sanderlin, L. V. Yang, “Transwell In Vitro Cell Migration and Invasion Assays” in *Cell Viability Assays*, Methods in Molecular Biology., O. Friedrich, D. F. Gilbert, Eds. (Springer US, 2023), pp. 349–359.
67. G. Muñoz-Gil, *et al.*, Quantitative evaluation of methods to analyze motion changes in single-particle experiments. *Nat Commun* **16**, 6749 (2025).
68. G. Munoz-Gil, *et al.*, Objective comparison of methods to decode anomalous diffusion. *Nat Commun* **12**, 6253 (2021).
69. S. Arlot, A. Celisse, Z. Harchaoui, A Kernel Multiple Change-point Algorithm via Model Selection. *Journal of Machine Learning Research* **20**, 1–56 (2019).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [RaluyPRcondensatesSupplementaryManuscript.pdf](#)