

Inaugural dissertation
for
obtaining the doctoral degree
of the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of the
Ruprecht - Karls - University
Heidelberg

Presented by
Lukas Frank, M.Sc.
Born in Stuttgart, Germany
Oral examination 14.07.2023

**Perturbing and imaging nuclear compartments
to reveal mechanisms of transcription regulation
and telomere maintenance**

Referees: Prof. Dr. Karsten Rippe
Prof. Dr. Sylvia Erhardt

This work was performed from January 2018 to April 2023 under the supervision of Prof. Dr. Karsten Rippe in the Division of Chromatin Networks at the German Cancer Research Center (DKFZ) and the BioQuant Center in Heidelberg, Germany.

Declaration

I hereby declare that I have written the submitted dissertation: "Perturbing and imaging nuclear compartments to reveal mechanisms of transcription regulation and telomere maintenance" myself and in this process, have used no other sources or materials than those explicitly indicated. I hereby declare that I have not applied to be examined at any other institution, nor have I used the dissertation in this or any other form at any other institution as an examination paper, nor submitted it to any other faculty as dissertation.

Place, Date

Lukas Frank

Table of contents

SUMMARY	III
ZUSAMMENFASSUNG.....	V
LIST OF PUBLICATIONS	VII
ABBREVIATIONS	VIII
I. INTRODUCTION.....	1
1. The cell nucleus is organized into functional compartments.....	1
2. General mechanisms of nuclear compartment formation	3
3. Specialized compartments formed around genomic repeats.....	6
3.1 Mouse pericentric heterochromatin (PCH).....	7
3.2 Transcription clusters	8
3.3 Telomeres	9
4. Scope of the thesis	13
II. RESULTS.....	15
1. Chromatin targeting of effector constructs enables microscopy-based characterization of compartment activities	15
1.1 A toolbox of modular effectors mimics TF binding and activity.....	16
1.2 Directing transcriptional activators to mouse major satellite repeats interrogates the chromocenter compaction mechanism	18
1.3 Assembling TF-like effectors at a reporter gene cluster links readouts of compartmental characteristics to transcription.....	22
1.4 Combinatorial use of TF constructs distinguishes strong and weak ADs based on their functional interaction with co-activators	25
1.5 PHR optodroplets model how phase-separated TF compartments form around chromatin and affect transcription	28
1.6 Different types of light-induced phase-separated TF droplets do not enhance transcription.....	32
2. ALT-FISH – a microscopic assay for the detection of ALT-specific compartments	35
2.1 Single-stranded telomeric repeats of various origin are a unifying feature of ALT	36
2.2 ALT-FISH visualizes compartments of single-stranded telomeric nucleic acids.....	36
2.3 The Telosegment toolkit enables automation of ALT-FISH data analysis.....	38
2.4 Telomeric RNA TERRA and ECTRs contribute to the ALT-FISH signal	40
2.5 SSTR compartments contain telomere repeats in the kilobase range	42
2.6 ALT-FISH foci enrich components of ALT-associated PML bodies (APBs)	43
2.7 ALT-FISH stratifies ALT-positive and ALT-negative cells with high accuracy	46
2.8 The ALT activity profiles of cancer cell lines are predicted by ALT-FISH spot counts	48
3. Mapping the distribution of ALT activity in unperturbed and perturbed cell populations using ALT-FISH	53
3.1 Capturing changes in ALT activity upon telomerase-to-ALT switching	54
3.2 Perturbing ATRX to study genetic ALT activity modulation.....	56
3.3 Integrative comparison of ALT-FISH and C-circle data reveals robust correlation	59
3.4 Adaptation of the ALT-FISH assay to high-content screening formats	60
3.5 Automated analysis and quality control of high-content ALT-FISH data with Telosegment-HT.....	62
3.6 Sensitive and specific detection of ALT activity in hundreds of thousands of cells grown in microplates	64
3.7 High-content ALT-FISH data uncovers links between ALT activity and the cell cycle	66
3.8 Spatial cell isolation correlates with higher ALT activity	68
3.9 Stochasticity of ALT activity revealed by spatial analysis of micro-communities.....	70
3.10 Drug effects on ALT activity and viability are simultaneously tested in microplates	72

4.	Spatial mapping of ALT activity compartments in primary tumors at single cell resolution.....	74
4.1	Considerations for ALT-FISH staining and image analysis on tumor tissue sections.....	75
4.2	The ALT-FISH assay is compatible with primary tumor material.....	75
4.3	Spatial mapping of ALT activity in leiomyosarcoma and neuroblastoma tumors.....	77
4.4	Combining ALT-FISH with a tumor marker to distinguish between the heterogeneity of tumor cells and non-malignant infiltrations.....	81
4.5	Single cell transcriptome information and ALT activity obtained from the same cell in intact neuroblastoma tumor tissue.....	82
III.	DISCUSSION.....	85
1.	Probing nuclear compartment transitions, co-activator interaction and phase separation by perturbation and imaging.....	88
2.	An integrative assay quantifies ALT activity by mapping compartments of single-stranded telomere repeats.....	92
3.	Resolving ALT activity in single cells in their spatial context.....	95
4.	Conclusions and Perspectives.....	99
IV.	MATERIALS AND METHODS.....	100
1.	Materials.....	100
1.1	Plasmids.....	100
1.2	Oligonucleotides and probes.....	100
1.3	Antibodies.....	100
1.4	Relevant fine chemicals.....	101
1.5	Cell lines.....	101
1.6	Tumor samples and tissue sectioning.....	101
1.7	Microscopy instrumentation.....	102
1.8	Software.....	102
2.	Methods.....	103
2.1	Cell culture.....	103
2.2	Plasmid DNA and siRNA transfection.....	104
2.3	Chromocenter perturbation experiments and data analysis.....	104
2.4	Light-induced binding to the reporter and measurement of reporter features.....	105
2.5	Light-induced RNA and BRD4 time course experiments.....	106
2.6	Optodroplet imaging and quantification.....	107
2.7	qRT-PCR.....	107
2.8	C-circle and TRAP assays.....	108
2.9	Western blotting.....	108
2.10	ATRX re-expression.....	109
2.11	ALT-FISH staining on cell lines.....	109
2.12	Combined ALT-FISH and immuno-staining of cell lines.....	110
2.13	Confocal imaging of ALT-FISH stained cell line samples.....	110
2.14	Combined ALT-FISH and other staining of tumor tissue sections.....	111
2.15	Molecular Cartography (MC) and ALT-FISH on neuroblastoma.....	111
2.16	Confocal microscopy of tissue samples.....	112
2.17	Cell line ALT-FISH image analysis with Telosegment.....	112
2.18	High-content microplate image analysis with Telosegment-HT.....	114
2.19	Tissue image analysis.....	115
2.20	Cell cycle inference and spatial analysis of microplate data.....	116
2.21	Additional methods.....	117
	REFERENCES.....	118
	ACKNOWLEDGEMENTS.....	129

Summary

The cell nucleus is organized into functional domains that form around chromatin, which serves as a scaffold composed of DNA, proteins, and associated RNAs. On the 0.1-1 μm mesoscale these domains can form spatially defined compartments with distinct composition and properties that enrich specific genomic activities like transcription, chromatin modification or DNA repair. In addition, extrachromosomal DNA elements and RNAs can separate from the chromatin template and assemble with proteins into nuclear bodies. The resulting accumulations of proteins and nucleic acids in the nucleus modulate chromatin-templated processes and their organization. The assembly of these compartments occurs in a self-organizing manner via direct and indirect binding of proteins to DNA and/or RNA. Recently, it has been proposed that multivalent interactions drive compartmentalization by inducing phase separation with a non-stoichiometric accumulation of factors into biomolecular condensates. Despite the importance of compartments for genome regulation, insights into their structure and material properties and how these affect their function is limited. To address this issue, it is important to devise approaches that can perturb nuclear compartments in a targeted manner, while also measuring changes in genome activities within the same cell. In this thesis, the methodology to reveal the underlying structure-function relationships of nuclear compartments has been advanced and applied to compartments involved in activation and silencing of chromatin, and telomere maintenance in cancer cells.

I first established a toolbox of chromatin effector constructs to probe and perturb properties of nuclear compartments in living cells that comprised different combinations of DNA binding, transcription activation and light-dependent interaction domains. In addition, I developed workflows to quantitatively assess relevant compartment features by fluorescence microscopy. These methods were employed to study the compaction mechanism of mouse pericentric heterochromatin (PCH) foci and to investigate the interplay between transcriptional co-activators, phase separation and transcription at an inducible reporter gene cluster. It revealed determinants of PCH compaction and identified differential co-activator usage and multivalent interactions as contributors to transcription factor (TF) strength. The results furthermore challenged the model of TF phase separation as a general positive driver of gene transcription. In the second part, I focused on exploiting the detection of compartments for measuring activity of the alternative lengthening of telomeres (ALT) pathway used by cancer cells to extend their telomeres in absence of telomerase. I developed ALT-FISH, a scalable and quantitative imaging assay that detects ALT pathway-specific compartments containing large amounts of single-stranded telomeric nucleic acids. I applied the method to cell line models from different cancer entities and to tumor tissue from leiomyosarcoma and neuroblastoma patients. By devising automated ALT-FISH data acquisition and analysis

workflows, I implemented an approach, which enabled ALT activity measurements in hundreds of thousands of single cells. These technological advancements provided a quantitative description of ALT activity at single cell resolution and were used to characterize the spatial distribution of ALT activity in relation to other biological features and in response to perturbations. Finally, a novel approach for studying the regulation of ALT in tumors could be established by integrating the method with the spatially resolved detection of single cell transcriptomes.

In summary, this thesis introduced and utilized several methods to establish connections between nuclear compartment organization, chromatin features, transcription regulation, and telomere maintenance. These perturbation and imaging techniques are versatile and may be applied to dissect nuclear activities related to other compartments and biological model systems. Furthermore, the detection of ALT activity has demonstrated that compartments can offer valuable biological insights into how phenotypic cellular heterogeneity is encoded and linked to diseases such as cancer.

Zusammenfassung

Der Zellkern ist in räumliche Einheiten, sogenannte Kernkompartimente, gegliedert, die sich aus einer bestimmten Chromatin-, Protein- und RNA-Zusammensetzung auf der Größenskala von 0.1-1 μm ergeben und spezifische Funktionen erfüllen. Dazu zählen beispielsweise die Regulation der Genexpression, die kontrollierte posttranslationale Modifikation von Histonproteinen und die Reparatur von lokalen DNA-Schäden. Darüber hinaus können sich extrachromosomale DNA-Elemente und RNAs vom Chromatin lösen und sich mit Proteinen zu Kernkompartimenten zusammenschließen. Diese freien Komplexe aus Proteinen und Nukleinsäuren im Zellkern können wiederum weitere Chromatin-vermittelte Prozesse und deren Organisation modulieren. Die Bildung von Kernkompartimenten erfolgt auf selbstorganisierende Weise durch direkte und indirekte Bindung von Proteinen an DNA und/oder RNA. Ein neues Entstehungsmodell sieht vor, dass die lokale Anreicherung für bestimmte Kompartimente durch den Prozess der Phasentrennung zustande kommen kann. Die hierbei gebildeten "biomolekularen Kondensate" können Faktoren nicht-stöchiometrisch anreichern und haben weitere besondere biophysikalische Eigenschaften. Jedoch sind die wechselseitigen Beziehungen zwischen Kompartimentierung, Materialeigenschaften und biologischen Funktionen des Zellkerns wenig verstanden. Es bedarf neuartiger methodischer Ansätze, um Kernkompartimente gezielt zu manipulieren und deren Eigenschaften sowie Aktivität in einzelnen Zellen zu messen. In dieser Studie wurden experimentelle Methoden entwickelt, mit denen gezielt Kernkompartimente untersucht und ihre Aktivität gemessen werden kann. Die hier untersuchten Kompartimente sind an der Regulierung der Genexpression sowie der Telomerverlängerung von Krebszellen beteiligt.

Im ersten Teil der Arbeit habe ich eine Toolbox aus Chromatineffektoren etabliert und Methoden zur quantitativen fluoreszenzmikroskopischen Messung von Kompartiment-eigenschaften entwickelt. Mithilfe eines aktivierbaren Reporter-Gen-Clusters unter Anwendung dieser Methoden auf perizentrisches Heterochromatin konnte ich Faktoren identifizieren, die die Kompaktierung von Heterochromatin regulieren und die Aktivität von Transkriptionsfaktoren verstärken. Die Ergebnisse zeigten außerdem, dass die Phasentrennung von Transkriptionsfaktoren auf Chromatin nicht immer mit einer verstärkten Transkription einhergeht.

Im zweiten Teil meiner Arbeit lag der Fokus auf der Identifizierung von Kompartimenten zur Messung der Aktivität des alternativen Telomerverlängerungsmechanismus (ALT), welcher von Tumorzellen genutzt wird, um ihre Telomere in Abwesenheit des Enzyms Telomerase zu verlängern. Dazu entwickelte ich ALT-FISH, ein stark skalierbares und quantitatives Mikroskopieverfahren, das spezifische Kompartimente für den ALT-Signalweg nachweist. Die

Methode habe ich anschließend an Zelllinienmodellen verschiedener Krebsarten als auch an Tumorgewebe von Patienten mit Leiomyosarkom und Neuroblastom validiert. Durch die Entwicklung von automatisierten Workflows zur ALT-FISH Datenerfassung und -analyse konnte ich Messungen der ALT-Aktivität in Hunderttausenden von Einzelzellen durchführen. Dank dieser technologischen Fortschritte konnte die ALT-Aktivität erstmals quantitativ und mit Einzellauflösung beschrieben werden. Außerdem wurde untersucht, wie sich ihre räumliche Verteilung im Verhältnis zu anderen biologischen Merkmalen verhält und wie sie auf bestimmte Einflüsse reagiert. Schlussendlich wurde ein innovativer Ansatz entwickelt, um die Regulation der ALT-Aktivität in Tumoren zu untersuchen, der auf der Kombination von ALT-FISH mit der räumlich aufgelösten Erfassung von Einzelzelltranskriptomen basiert.

In dieser Arbeit wurden mehrere Methoden entwickelt und eingesetzt, um mechanistische Verbindungen zwischen der Organisation von Kernkompartimenten, deren Chromatinstatus und biologischer Aktivität zu analysieren. Die hier präsentierten experimentellen Ansätze sind vielseitig einsetzbar und können daher auch für die Manipulation und Untersuchung anderer Kernkompartimente angepasst werden. Darüber hinaus hat der Nachweis der ALT-Aktivität gezeigt, dass Kernkompartimente wertvolle biologische Erkenntnisse darüber liefern können, wie zelluläre Heterogenität kodiert ist und mit Krankheiten wie Krebs im Zusammenhang steht.

List of publications

In the course of this thesis I contributed to the following publications:

Frank, L., Rademacher, A., Mucke, N., Tirier, S.M., Koeleman, E., Knotz, C., Schumacher, S., Stainczyk, S.A., Westermann, F., Frohling, S., Chudasama, P., and Rippe, K. (2022). ALT-FISH quantifies alternative lengthening of telomeres activity by imaging of single-stranded repeats. *Nucleic Acids Res* 50, e61. 10.1093/nar/gkac113.

Trojanowski, J.*, Frank, L.*, Rademacher, A., Mucke, N., Grigaitis, P., and Rippe, K. (2022). Transcription activation is enhanced by multivalent interactions independent of phase separation. *Mol Cell* 82, 1878-1893 e1810. 10.1016/j.molcel.2022.04.017.

* equal contribution

Frank, L., and Rippe, K. (2020). Repetitive RNAs as Regulators of Chromatin-Associated Subcompartment Formation by Phase Separation. *J Mol Biol* 432, 4270-4286. 10.1016/j.jmb.2020.04.015.

Frank, L., Weinmann, R., Erdel, F., Trojanowski, J., and Rippe, K. (2021). Transcriptional Activation of Heterochromatin by Recruitment of dCas9 Activators. *Methods Mol Biol* 2351, 307-320. 10.1007/978-1-0716-1597-3_17.

Erdel, F., Rademacher, A., Vlijm, R., Tunnermann, J., Frank, L., Weinmann, R., Schweigert, E., Yserentant, K., Hummert, J., Bauer, C., Schumacher, S., Al Alwash, A., Normand, C., Herten, D.P., Engelhardt, J., and Rippe, K. (2020). Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of HP1-Driven Liquid-Liquid Phase Separation. *Mol Cell* 78, 236-249 e237. 10.1016/j.molcel.2020.02.005.

Poos, A.M., Schroeder, C., Jaishankar, N., Roll, D., Oswald, M., Meiners, J., Braun, D.M., Knotz, C., Frank, L., Gunkel, M., Spilger, R., Wollmann, T., Polonski, A., Makrypidi-Fraune, G., Fraune, C., Graefen, M., et al. (2022). PITX1 Is a Regulator of TERT Expression in Prostate Cancer with Prognostic Power. *Cancers (Basel)* 14, 1267. 10.3390/cancers14051267.

Brito, D.V.C., Gulmez Karaca, K., Kupke, J., Frank, L., and Oliveira, A.M.M. (2020). MeCP2 gates spatial learning-induced alternative splicing events in the mouse hippocampus. *Mol Brain* 13, 156. 10.1186/s13041-020-00695-1.

Abbreviations

(E)GFP	(enhanced) green fluorescent protein
a.u.	arbitrary units
AD	activation domain
ALT	alternative lengthening of telomeres
ANOVA	analysis of variance testing
APB	ALT-associated PML body
ASF1	Anti-silencing function protein 1
ATM	ataxia telangiectasia mutated
ATR	ataxia telangiectasia and Rad3-related
ATRX	alpha-thalassemia/mental retardation, X-linked
BH	Benjamini-Hochberg correction
BIR	break-induced replication
BLinCR	blue light Induced chromatin recruitment
BRD4	bromodomain containing protein 4
CC	C-circle
CI	confidence interval
CIBN	N-terminal region of the CIB1 protein
CMV	cytomegalovirus
CRISPR	Clustered Regularly Interspersed Short Palindromic Repeats
CTD	carboxy-terminal domain of RNA polymerase II
ctrl	control
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA-binding domain
dCas9	dead/catalytically inactive CRISPR associated protein 9
DDR	DNA damage response
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
dn	double null
DSB	DNA double-strand break
ECTR	extrachromosomal telomeric repeat
EU	5-ethynyl-uridine
FBS	fetal bovine serum
FCS	fluorescence correlation spectroscopy
FISH	fluorescence in situ hybridization
FPR	false positive rate
FRAP	fluorescence recovery after photobleaching
FUSN	N-terminal IDR of the fused in sarcoma protein
GBP	GFP-binding protein
H3K27ac	histone H3 acetylation at lysine 27
H3K9me3	histone H3 tri-methylation at lysine 9
HAT	histone acetyl-transferase
HP1	heterochromatin protein 1
HR	homologous recombination
HRP	horse radish peroxidase
IDR	intrinsically disordered region

IF	immunofluorescence
iMEF	immortalized mouse embryonic fibroblasts
kd	knock-down
ko	knock-out
Lacl	lactose repressor protein
<i>lacO</i>	<i>lac</i> operon
LMS	leiomyosarcoma
loop	construct with sgRNA-PP7/tdPCP interaction modules
LT	long telomeres
mCherry / mCh	monomeric Cherry red fluorescent protein
MRN	Mre11-Rad50-NBS1
MSR	major satellite repeats
n.s. / ns	not significant
NB	neuroblastoma
NHEJ	non-homologous end joining
NLS	nuclear localization signal
opto	construct with CIBN/PHR dimerization modules
optolooop	construct with sgRNA-PP7/tdPCP-CIBN/PHR interaction modules
PBS	phosphate buffered saline
PCH	pericentric heterochromatin
pedGBM	pediatric glioblastoma
PHR	photolyase homology region of the CRY2 protein
PML	promyelocytic leukemia
PML-NB	promyelocytic leukemia nuclear body
Pol II	RNA polymerase II
qRT-PCR	quantitative real-time polymerase chain reaction
rTetR	reverse tetracycline repressor protein
s.d.	standard deviation
s.e.m	standard error of the mean
sgRNA	single guide RNA
ssDNA	single-stranded DNA
SSTR	single-stranded telomeric repeats
ST	short telomeres
SUMO	small ubiquitin-related modifier
t-loop	telomeric loop
tdMCP	tandem dimer MS2 coat protein
tdPCP	tandem dimer PP7 coat protein
tdTomato	tandem dimer Tomato red fluorescent protein
TERRA	telomere repeat-containing RNA
TERT	protein subunit of the telomerase enzyme complex
<i>tetO</i>	<i>tet</i> operon
TF	transcription factor
TMM	telomere maintenance mechanism
TRF2	telomere repeat binding factor 2
VPR	VP64-p65-Rta tripartite activator
wt	wild type

I. Introduction

1. The cell nucleus is organized into functional compartments

The cell nucleus contains the genetic information in the form of DNA wrapped around histone proteins to form chromatin. A pool of diverse macromolecules regulates how and when the genome is transcribed, compacted, replicated or repaired by spatially confining its active components in a coordinated manner within the nucleus. Many nuclear components show punctate staining patterns on the mesoscale of 0.1-1 μm when analyzed by immuno-staining, suggesting that they are compartmentalized into local accumulations with sometimes steep concentration boundaries [1]. In these compartments, biological processes can be aggregated to accelerate enzymatic and binding reactions, to regulate only specific gene loci or to reduce crosstalk between opposing processes. Nuclear compartments may range from the nanometer to micrometer scale [2] and can be classified by functional, compositional and structural criteria. One main distinction is made based on their chromatin content and association with gene transcription [3]. Compartments that have a high concentration of chromatin are usually linked to densely packed chromatin regions that are often transcriptionally silent. On the other hand, compartments found in actively transcribed regions or in the interchromatin space typically have a higher RNA and protein content and are sometimes referred to as nuclear bodies [4]. **Figure 1** provides an overview of specific compartments found in the mammalian cell nucleus.

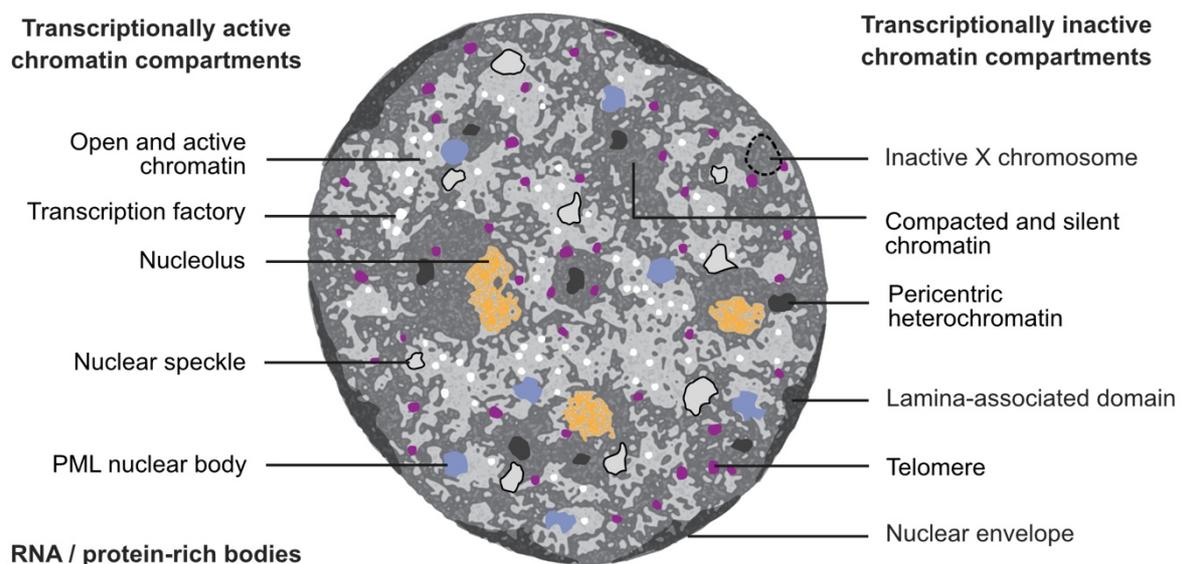


Figure 1. Specific examples of functional compartments formed in the cell nucleus.

Examples of compartments organizing active chromatin include nucleoli, which transcribe rDNA and produce ribosomal subunits [5], and transcription factories, which are bodies of enriched active RNA polymerase II and transcription factors (TFs) that regulate transcription throughout the genome [6, 7]. Nuclear speckles are assemblies of RNAs, splicing factors and other proteins involved in post-transcriptional RNA processing, storage and transport [8]. They mostly form in the interchromatin space, where they dynamically exchange factors with the nucleoplasm and the surrounding chromatin. PML nuclear bodies (PML-NBs) are formed by spherical shells of PML protein filled with a core of numerous other proteins, of which many are SUMOylated [9]. Although their exact function remains enigmatic, PML-NBs have been linked to the regulation of cellular stress responses [10]. Similar to nuclear speckles, they are mostly devoid of chromatin [11]. However, they can associate with specific genomic loci in situations that include viral infection, DNA damage and telomere maintenance [12].

Several regions of the genome that are transcriptionally silent also form unique structural or functional barriers that distinguish them from their surrounding environment. Chromosomal regions adjacent to the nuclear envelope (lamina-associated domains) are mostly silenced, but also regulated by the unique interactions with the lamin proteins [13, 14]. The inactive X chromosome forms a highly compacted and transcriptionally repressed "Barr body" that locates to the nuclear periphery and selectively accumulates specific chromatin marks [15, 16]. Other examples include pericentric heterochromatin in mouse cells [17] and telomeres at chromosome ends [18]. Both form compacted structures around repetitive DNA sequences and are crucial for genome stability and chromosome segregation [19, 20].

Ongoing research aims to understand how the structural and functional barriers of various compartments in the nucleus are established and maintained. This knowledge is essential to comprehend how cells process and integrate biological information at different size and time scales within the nucleus. In the following sections, some general organizational principles and formation mechanisms of nuclear compartments will be introduced. In addition, their biological function will be illustrated with specific examples.

2. General mechanisms of nuclear compartment formation

The basis for nuclear compartmentalization is provided by self-organizing interaction circuits that confine soluble factors to a specific location or a set of locations that share molecular cues for initiating compartment assembly [21]. The specific assembly mechanism depends on the type of compartment and can involve both cross- and self-interactions between RNA, proteins and binding sites on the chromatin scaffold. Hereby, the role of the nucleosome chain can vary [22] and assembly by stoichiometric binding is distinguished from assembly by phase separation (Figure 2).

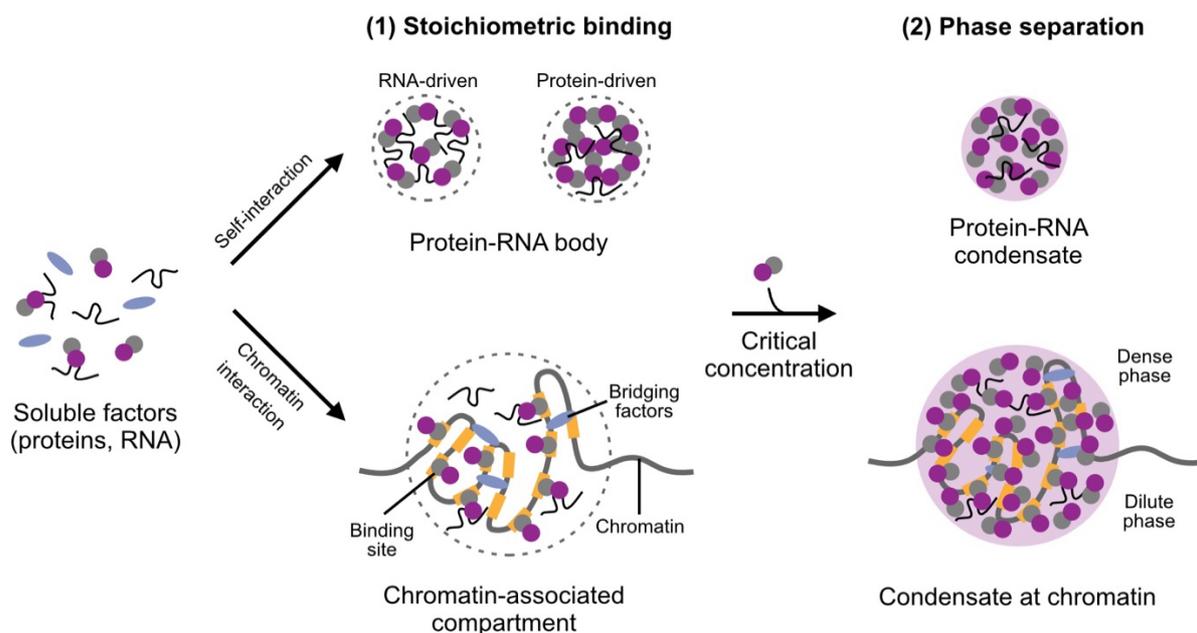


Figure 2. Mechanisms of nuclear compartmentalization. Two main mechanisms of compartment formation are distinguished: (1) *Stoichiometric binding*: soluble factors form complexes of defined composition with or without the involvement of chromatin. The assembly is achieved by multivalent self-interacting proteins or RNAs and/or factors which recognize specific binding sites on chromatin. Chromatin-bound bridging factors may additionally introduce structural changes on the level of the chromatin fiber. (2) *Phase separation*: soluble factors separate into a dense (condensate) and dilute phase above a so called critical concentration. Inside condensates, factors are concentrated non-stoichiometrically. Condensates may or may not involve chromatin binding sites. Factors with the ability to form multivalent self-interactions typically have a higher propensity to undergo a phase transition.

In one model, chromatin acts as the core scaffold by providing clusters of sites to which effectors bind in a DNA sequence or chromatin mark-specific manner. This is for example the case for mouse pericentric heterochromatin [23, 24] and telomeres, which both contain significant amounts of repetitive binding sites that recruit specific proteins. Compartments can form along a continuous region on the chromatin fiber or comprise multiple chromatin loci that come together in 3D [25]. Folding into a higher order structure may involve cross-linking interactions mediated by bivalent bridging factors. In alternative, chromatin is not directly or

only transiently involved as scaffold for the formation of compartments. Instead, protein and RNA complexes assemble in the interchromatin space, wherein RNA, protein or both can act multivalent linkers to drive self-assembly [26, 27]. This mode of assembly is likely common to both PML-NBs and nuclear speckles, which carry out chromatin-associated functions, but do not necessarily, or only under certain circumstances, contain chromatin at their core.

Nuclear components rapidly diffuse through the nucleus and many DNA binding proteins have short residence times on chromatin [28]. This raises the question of how compartments are stabilized without investing cellular energy, even though diffusion works to balance these steep concentration gradients. In general terms, proteins and nucleic acids can form macromolecular assemblies in two ways that are not mutually exclusive. The first is through high-affinity stoichiometric interactions between well-structured domains, which may include DNA binding sites on the nucleosome chain. The second is through variable-affinity interactions between intrinsically disordered regions (IDRs) of proteins [22, 29]. IDRs are prevalent in nucleic acid binding proteins across species [30] and exhibit high conformational variability, which results in many possible interaction modes. These range from highly specific lock-and-key type binding [31] to more promiscuous "sticky" interactions [32]. In many instances, structured domains work together with IDRs, for example, to accomplish efficient targeting of proteins to DNA [33]. Most notably, many IDRs can form multivalent links due to their repetitive sequence nature [34], making them prime candidates for building large protein networks that are driven by self-interaction. In recent years, this property of IDRs has been used to rationalize the assembly of IDR-containing proteins into bodies that form by phase separation, which are referred to as biomolecular condensates or liquid droplet [35]. This mode of formation is fundamentally different from stoichiometric assembly into higher order complexes.

In broad terms, phase separation is the unmixing of molecules into a concentrated and less concentrated phase at a critical saturating concentration where self-interaction is energetically more favorable than interaction with the solvent (**Figure 2**, right). In addition, factors such as temperature, pH, and ions can affect the critical concentration by modulating the self-interaction propensity. Phase separation is an intriguing concept to explain compartmentalization, since phase-separated systems have key characteristics that could potentially be exploited for biological regulation [26]. Phases maintain a sharp concentration boundary, while exchange of molecules between them is still possible. This can locally enrich large amounts of molecules and also create reactive surfaces for biological processes. In equilibrium, the molecule concentration within a phase-separated compartment is constant, and concentration changes lead to expansion or shrinkage of the dense phase volume. This concentration buffering ability could protect enzymatic reactions against changes of the

environment. Finally, molecules can be excluded or included in the dense or dilute phase based on their chemical properties (e.g. hydrophobicity, charge, etc.). Such selective solubility can be utilized to confine functionally related molecules within a compartment and segregate them from unrelated processes or inhibiting reaction products.

While originally described for P-granules in the cytoplasm [36], phase transitions have recently been invoked to explain the formation of numerous nuclear compartments. These include nucleoli [37], mouse chromocenters [38], DNA damage foci [39], PML bodies [40] and transcription factories or TF clusters [41, 42]. However, the classification as genuine phase-separated systems remains controversial and is a topic of ongoing research. This also applies to defining the exact nature of the molecular contacts that maintain these bodies [43] and how they relate to biological activity. It is often argued that *in vitro* evidence for condensate formation is flawed by the non-physiological conditions used in these experiments. The question on how to best distinguish compartments that form by phase separation from those that use other mechanisms has been addressed in a number of studies [44-46]. Here, the focus is on how biological function inside the cell can be probed and related to the type of compartmentalization mechanism. To address this question, it is necessary to introduce perturbations of the relevant molecular properties (e.g., binding to chromatin, self-interaction, etc.) of the compartments in question and readout their relevant biological activities.

3. Specialized compartments formed around genomic repeats

The present and following sections will illustrate the organizational principles and functions of three different prototypical compartments that form around genomic repeats: (i) Silenced mouse pericentric heterochromatin (PCH), which is associated with compaction and large scale chromosomal interactions; (ii) Transcription clusters, groups of genomic loci that co-localize for efficient transcription or regulation; and (iii) Telomeres, which are specialized chromosome end compartments with relevance for genome integrity and cancer. Although these compartments differ in terms of their function, location, size, activity, and chromatin interactome, they are governed by common organizational principles. Here, the common principles of compartment formation around these genomic repeats are highlighted in an integrative manner (**Figure 3**) and later on discussed in more detail in the respective sections.

	Mouse PCH	Transcription clusters	Telomeres
			
DNA repeat	Major satellite	TFBS, REs	TTAGGG
Marker proteins	HP1, SUV39H	TFs, Pol II, co-activators	Shelterins
Chromatin	H3K9me3, compacted	Acetylated, open	Variable
Activity	Silencing, chromosome segregation	Gene expression regulation	DDR inhibition, repeat extension (TERT, ALT)
Higher order organization	Chromocenters	Multi-way contact hubs	T-loop, APBs (ALT)

Figure 3. Concept of how mouse PCH, transcription clusters and telomeres are organized as compartments around genomic repeats. Repetitive DNA binding sites on the chromatin scaffold enrich specific marker proteins, which determine biological activity and link to a characteristic chromatin state. Furthermore, organization into higher order structures may occur to fulfil compartment-specific functions. Details are described in the main text.

Mammalian genomes contain large amounts of repetitive DNA [47, 48], which can be either dispersed or organized into regular arrays that span kilobase to megabase regions on chromosomes. Densely arrayed repetition is common for non-protein coding DNA elements such as satellite and simple repeats found in (peri-) centromeres and telomeres, respectively. The 234 bp AT-rich major satellite repeat (MSR) constitutes a large fraction of mouse PCH [49], while telomeres consist of regular head to tail arrays of (TTAGGG)_n repeats [50]. In addition, transcribed protein or RNA coding genes and their associated regulatory elements

(RE) and TF binding sites (TFBS) can occur in large repetitive clusters, for example, as observed for the histone [51] and the rDNA gene loci [52] found on multiple chromosomes.

On the DNA level, repeat arrays can bind large amounts of similar effector proteins, which in turn establish a spatially segregated functional state on the level of chromatin. In PCH, heterochromatin proteins are bound to compact chromatin and silence transcription. Transcriptional clusters jointly accumulate TFs, RNA Pol II and co-activators, to maintain an open, hyper-acetylated and transcriptionally active chromatin environment. Telomeres are bound by shelterins to inhibit DNA damage response (DDR) pathways and regulate telomere extension. Epigenetic patterns can also repeat themselves independent of the underlying DNA sequence and forming self-sustaining feedback loops [53, 54]. Furthermore, the chromatin complex of repeats, bound proteins and RNAs may form higher order structures that are crucial to mediate the respective specialized functions. In the following sections, the interconnections between these levels of organizations will be described and related to the corresponding biological function. PCH, transcription clusters will be discussed with emphasis on chromatin feedback circuits and the role of phase separation in transcription. Telomeres will be primarily discussed with respect to their state and function in cancer cells using the alternative lengthening of telomeres (ALT) pathway.

3.1 Mouse pericentric heterochromatin (PCH)

Mouse PCH is one example where sequence and chromatin state repetition work hand in hand to form large transcriptionally inactive domains. MSR chromatin is enriched in histone H3 tri-methylated at lysine 9 (H3K9me3), the H3K9 methyltransferase SUV39H and heterochromatin protein 1 (HP1) [20, 55]. PCH regions from several chromosome furthermore associate *in trans* into densely compacted DNA foci, called chromocenters. Both SUV39H and HP1 can bind to H3K9me3 marks and to each other, thereby forming a feedback circuit that maintains these marks at and around MSR sequences. The circuit is additionally backed up by a direct interaction of MSR transcripts and the SUV39H enzymes [24] and possibly by spreading of the repressed state to neighboring regions [56]. The resulting constant enrichment of repressive factors on MSR arrays protects them against intruding deposition of activating marks.

How exactly transcription is prevented at PCH and to which extent these self-sustained circuits can buffer perturbations from activators, remains an open question. While loss of the H3K9me3-HP1-SUV39H axis impacts PCH integrity and thereby destabilizes chromosome segregation, *Suv39h*-deficient cell models remain viable [55]. And remarkably, its lack does not affect the higher order organization into chromocenters [55, 57]. Consequently, multiple redundant pathways seem to be at play to control repression, compaction and coalescence

into higher order structures. Given the ability of PCH-associated proteins like HP1 and MeCP2 to form droplets *in vitro*, phase transitions have been proposed as mechanism that drives PCH formation [38, 58]. Although the repetitive binding sites could certainly nucleate large amounts of proteins, there is ongoing debate about the precise role of phase transitions and their specific role in this heterochromatin compartment [57].

3.2 Transcription clusters

The concept of self-sustaining circuits introduced for PCH can also rationalize how some transcriptionally active regions are maintained and regulated. Transition from a silenced to a transcribing state can be induced by sequence-specific binding of transcription factors (TF) to regulatory elements, leading to the recruitment of RNA polymerase and transcriptional co-activators in a multi-step process [59]. Feedback loops between writing and reading of histone modifications, such as H3K27ac and H3K4me3 [60, 61], can contribute to preserving marks of activation, even after transcription has stopped. Histone acetylation can directly facilitate transcription by reducing histone-DNA interactions and relaxing chromatin compaction [62]. It also entails recognition of acetylated histones by co-activator proteins, such as bromodomain containing protein 4 (BRD4), which in turn directly stimulates RNA polymerase activity [63].

Positive feedback loops can occur at the level of individual transcription units, but also in space between multiple activated chromatin segments. Physical proximity or topological constraints [64] can enhance the exchange of regulatory signals and factors and reinforce the active state or increase the responsiveness to transcription initiation [65, 66]. The underlying structure-function relationships of these stabilization mechanisms are not fully understood and different models have been proposed [67]. Transcription is known to occur in stochastic bursts, whose frequency is modulated by the TF concentration [68, 69]. Therefore, mechanisms that stabilize high concentrations of activators and co-activators around a transcription cluster are interesting models that could explain some of the observed synergistic effects.

One possible way to enrich additional activator molecules is by multivalent self-interaction. In addition, multivalency could drive a phase transition that embeds the active chromatin regions in the dense (condensate) phase, where high concentrations of activator are maintained. Such a mechanism is probable for transcription clusters or hubs of regulatory elements like super-enhancers [70, 71], which both offer high binding valency and are located in less dense regions of the nucleus. Many transcriptional regulators contain IDRs and show the ability to phase-separate in isolation or with partner proteins *in vitro* [41, 72]. Provided the unique characteristics discussed previously, a phase separation *in vivo* could have several functional implications also for transcriptional regulation. For instance, aggregation into droplets on chromatin has been proposed to reduce target search time on chromatin [73] and cause

specific enrichment or exclusion of co-activators like BRD4 or MED1 between the different phases [42, 74, 75]. Furthermore, phase separation may serve as positive feedback mechanism, for example, by bringing additional TF molecules to regions with a high level of actively engaged RNA polymerase II (Pol II) through interactions with its C-terminal domain (CTD) [76]. In addition, the disordered CTD of Pol II has been proposed to drive condensates formed by the polymerase itself [77].

Beyond the initial description of condensate material properties as liquid-like, i.e. exhibiting high internal mobility of molecules, it is now well recognized that this is not generally the case. Instead, mobility can range by orders of magnitude, from fast (liquid-like) to intermediate (gel-like) to slow (solid-like) [78, 79], although unifying definitions are lacking in the field. Together with other properties, like electrostatic charge or hydrophobicity, this specific condensate material state may directly regulate biological function. This was recently proposed for the phosphorylation of the Pol II CTD in transcriptional condensates, which could regulate switching from transcription initiation to RNA processing [80]. However, the persistent challenge of studying transcriptional condensates in living cells [45] in combination with functional readouts, has left many questions open on how their properties impact transcription-related processes. Direct comparisons of TF activity in living cells, above versus below the critical concentration for phase separation, are mostly missing. Moreover, many recent findings point at functions of TF multivalent interactions without the involvement of a phase transition [74, 81]. These functional contributions need to be investigated in more detail and isolation from phase separation, which requires novel *in vivo* methodologies. In conclusion, having a quantitative view of how condensates form and how they can regulate nuclear processes has implications for the understanding of certain disease phenotypes in cancer and prion-like diseases [79, 82]. The unique material properties can potentially be exploited for more selective targeting of drugs to transcription factors hubs or other compartments that form by proteins which lack targetable domains or high-affinity binding pockets [83].

3.3 Telomeres

Telomeres form protective nucleoprotein compartments at the end of linear chromosomes [84]. They are composed of TTAGGG sequence repeats that are bound by specialized proteins, called shelterins [85]. Telomere ends are uniquely processed into a G-rich single stranded overhang [86], which folds back, invading part of the double-stranded region to form the so called t-loop structure [87, 88]. The t-loop prevents them from being recognized as sites of DNA damage. The shelterins, in particular POT1 and TRF2, contribute to the suppression of DDR signaling [89, 90]. POT1 binds to the single-stranded portion of telomeres, competing for ssDNA sensors, that would trigger ATR signaling. The protein TRF2 directly binds to double-

stranded telomeric DNA together with TRF1 and supports maintenance of the t-loop [91], to avoid activation of the MRN complex to initiate ATM signaling [92].

To protect the chromosome ends from being processed as DNA breaks, telomeres rely on integrity of the shelterin complex and a certain minimum length [85, 93]. Due to intrinsic features of the DNA replication process, linear chromosome ends are incompletely replicated, causing telomeres to shorten after each cell division [19, 94, 95]. This limits the proliferative lifespan. When telomeres reach a critical length, their DNA damage suppressive function is lost, resulting in an unstable state, termed "crisis". The moment of crisis is characterized by end-to-end fusions and breakage-fusion-breakage (BFB) cycles between chromosomes and rampant genomic instability [96]. Under physiological conditions, crisis typically leads to senescence and/or apoptosis [97]. However, loss of central DNA damage checkpoints, such as p53, and accumulation of other genetic alterations, can rescue these adverse effects and enable escape from crisis [98]. Both the escape from crisis and the resulting gain of unlimited proliferation potential are hallmarks of cancer cells [99].

However, in order to sustain proliferation indefinitely, cancer cells need to activate telomere maintenance mechanisms (TMMs) that extend telomere repeats to counteract their net loss. Two TMMs are active in cancer: reactivation of the reverse transcriptase telomerase (TERT) [100, 101] and the alternative lengthening of telomeres (ALT) pathway [102, 103]. While TERT is active the majority of tumors, the overall prevalence of ALT is around 10-15% or lower [104]. However, it is overrepresented and linked to poor disease outcome [105, 106] in many cancer entities, which include leiomyosarcoma (78%) [107], pancreatic neuroendocrine tumors (61%) [108] and relapsed neuroblastoma (47%) [109]. Most notably, no ALT-targeted tumor therapies exist and an integrated mechanistic understanding about the factors that control ALT activation and suppression is lacking. Various molecular and cytological assays are employed to detect ALT activity in cancer cells by looking at changes in telomere repeat metabolism in these cells. These alterations manifest as changes in the positioning of telomere repeat sequences within the nucleus, their abundance, transcription, and their interactions with particular proteins.

It is now established that ALT exploits break-induced replication (BIR)-like processes for telomere extension [110, 111]. Herein, net telomeric DNA is gained at short telomeres by homologous recombination (HR) with telomere sequences from other chromosomes or sister chromatids. In contrast to TMM by telomerase, where addition of telomere repeats is controlled by a single enzyme, the stochastic recombination-based extension could explain why highly heterogeneous telomere length are observed ALT-positive cells [102, 112, 113]. The level of recombination at telomeres furthermore needs to be balanced to not compromise cell fitness [114]. One danger of this process is the release of immunogenic extrachromosomal telomeric

DNA repeats (ECTRs) [115-117], likely from the resolution of HR intermediates, internal damage sites or the trimming of excessively long telomeres [118, 119]. ECTRs are circular or branched and contain double-stranded and/or single-stranded regions. One abundant species are DNA C-circles (CCs). CC abundance can be quantified from bulk genomic DNA by a rolling-circle amplification reaction called the CC assay. This assay is widely used and considered the gold standard for ALT detection in research applications and the clinic [120].

The clustering of telomeres to facilitate recombination is another feature that needs to be strictly regulated in ALT. This involves the establishment of interactions between two or more telomeres and the dissolution of such. In particular, recombination intermediates need to be resolved and coordinated with DNA replication, transcription and mitosis, in order not to endanger chromosomal integrity. Telomere clustering is believed to predominantly occur in specialized nuclear compartments, termed ALT-associated PML bodies (APBs) [121], and may involve phase separation of PML protein and other APB components [122, 123]. APBs, which form by the interaction of PML proteins and numerous DNA damage factors with some but not all telomeres, are rarely found in somatic or telomerase-positive cells. Their microscopic detection by a combination of PML immuno-staining and telomere DNA FISH serves as reliable cytological marker of ALT [124].

Similar to the reorganization of telomeres into APBs, the chromatin state and transcription of telomeres undergoes significant changes in cells that employ the ALT pathway. The majority of ALT-positive tumors show loss-of-function mutations in the *ATRX* and *DAXX* genes [108], two factors with numerous roles in chromatin regulation at repetitive elements [125]. Loss of *ATRX* is not generally sufficient to induce ALT [126, 127]. Rather, it acts as an ALT suppressor and its loss facilitates ALT emergence by mechanisms that are not fully understood and are likely cell type specific [128]. In most commonly studied ALT-negative human cell lines, telomeres exhibit euchromatic features [129], but are only transcribed to very low levels. Telomeres are transcribed from the sub-telomeres towards the chromosome ends, producing a heterogeneous pool of G-rich telomere repeat containing RNA transcripts, named TERRA [130]. Oddly, telomeres of ALT-positive cancer cells exhibit heterochromatic features, like enrichment of H3K9me3 [129, 131], but produce significantly more TERRA [130, 132-134]. Furthermore, HP1 is enriched in APBs and HP1-bound telomeric chromatin can nucleate APB formation [135, 136]. TERRA can form RNA:DNA hybrids at telomeres in *cis* and in *trans* [137] and these hybrids are thought promote ALT telomere extension by increasing replication stress and initiation of recombination (reviewed in ref. [138]).

Altogether, these observations point at a complex network of genetic and epigenetic alterations that coordinate recombination, clustering and transcription in ALT. Despite ongoing research, several aspects of ALT are not understood, and tools for the systematic dissection

of this mechanism are lacking. For instance, the means by which ALT-positive cells compartmentalize and maintain the different species of telomeric nucleic acids in their nucleus is still unclear, and how these interact with nuclear bodies like APBs and other proteins to modulate telomere lengthening by HR. Additionally, it is unclear whether ALT is uniformly active in all cells of a tumor to sustain bulk telomere length, or whether there are temporally or spatially defined patterns in ALT activity, which may also link to other cancer cell features, such as genomic instability and cell cycle. Lastly, large scale ALT-targeted drug susceptibility and genetic screens are missing as resources for the efficient exploration and identification of factors involved in ALT. Numerous studies have reported the co-occurrence of ALT and telomerase signatures in the same tumor [139-142], suggesting that both mechanisms may co-exist in the same cell population to drive tumor adaptability [143]. However, the absence of scalable universal assays that can quantify ALT activity from the subcellular to the patient level represents a gap that needs to be addressed to draw an integrative mechanistic and clinical picture of the ALT pathway in cancer. These methods are also needed to understand when and how single telomeres engage in ALT-mediated extension and how the extrachromosomal by-products of ALT are partitioned inside cells and linked to nuclear organization.

4. Scope of the thesis

The molecular properties of chromatin-associated nuclear compartments are crucial in targeting and regulating genome activities in the cell as described in the preceding sections. Recently, the concept of phase separation introduced the classification of some compartments as biomolecular condensates. However, it is only partly understood if and how phase separation is relevant for their assembly and biological function. For example, it remains unclear whether transcriptional condensate formation occurs under endogenous conditions and how these condensates would modulate transcriptional output [74, 81, 144]. Apart from phase separation, there are several other structure-function relationships on the mesoscale that are not resolved, including the maintenance of large silenced and active chromatin domains, and how specific nuclear organization changes are linked to disease. To address these issues, methods and model systems are needed that enable the examination of nuclear compartment function, reveal the mechanism of their formation, and provide measurements of their activities within the cell at both high temporal and spatial resolution. To achieve this goal, it is essential to develop approaches for tagging, building, imaging and manipulating compartments at their chromatin nucleation sites. In addition, strategies to modify compartment composition and location must be devised. My thesis advances this subject in several areas with specific applications to three genome activities, namely heterochromatin formation, transcription regulation and telomere maintenance in cancer cells using the ALT pathway.

Different targeted approaches to perturb, artificially reproduce and probe compartment features in living cells were developed. They were applied to trigger chromatin state transitions of mouse pericentric heterochromatin foci involved in repeat silencing as well as to dissect the role of phase separation in transcription regulation at a reporter gene cluster. Furthermore, the microscopic detection of nuclear compartments was exploited for *in situ* phenotyping of cancer cells that employ the ALT pathway for telomere maintenance. Universal and quantitative methods to reliably assess ALT activity in single cancer cells are lacking, but a spatially resolved view on ALT activity is needed to comprehensively describe tumor heterogeneity and identify novel links between cellular features and activity of this pathway. Here, a microscopic assay (ALT-FISH) was developed to detect ALT-specific compartments containing large amounts of single-stranded telomeric nucleic acids. The assay was applied to dissect heterogeneity of cancer cell lines from different tumor entities and primary tumor tissue from patients. Automated image analysis workflows were established to enable compartment analysis by microscopy on two main levels: (i) To quantitate functional features of compartment activity (chromatin state, transcription, etc.) with high spatial and temporal resolution and (ii)

To measure compartment abundance in hundreds of thousands of single cells and relate heterogeneity to cell phenotypes in their spatial context.

The thesis integrates novel perturbation and imaging approaches to advance our understanding of the factors that link compartment organization to their genome-associated activities. Furthermore, I demonstrate with the ALT-FISH method that quantitative nuclear compartment analysis can be exploited as an activity readout for high-throughput cellular phenotyping. By mapping ALT activity from cell lines to cancer tissues, screening assays for potential ALT-linked cancer vulnerabilities can be established. The resulting novel layer of information obtained on tumor heterogeneity of the ALT pathway is expected to refine patient stratification and clinical decision making in tumor entities like neuroblastoma and sarcoma where ALT is active.

II. Results

1. Chromatin targeting of effector constructs enables microscopy-based characterization of compartment activities

The assembly mechanism, maintenance, and biological activity of nuclear compartments can be dissected by perturbing their composition and material properties. To implement such an approach, techniques have been developed here that mimic and modify compartment properties *in vivo* and are integrated with functional readouts. Genomic regions containing DNA repeats make an appealing model case for research on nuclear compartments. Their repetitive binding sites can concentrate effectors along a large chromatin region to create a compartment that is physically and/or functionally isolated from its surroundings. This same property can also be exploited to tether large amounts of fluorescently tagged effectors or sensors to a region of interest for its characterization by fluorescence microscopy.

In this first chapter of the thesis, I focused on the functional characterization of compartments formed around clustered DNA binding sites. First, a toolbox of synthetic TF-like constructs with fluorescent tags, adjustable binding and effector characteristics was generated to allow for flexible perturbation and imaging of chromatin-associated compartments. In a second step, these constructs were used to assess factors that control the compartmentalization of mouse chromocenters as a prototypical heterochromatin compartment. Third, approaches based on the toolbox constructs were devised that specifically enabled testing of structure-function relationships of phase-separated TF droplets in transcriptional control using a repetitive and inducible reporter gene cluster.

The section on chromocenter perturbation is part of two recent publications [57, 145]. The characterization and application of the effector toolbox to investigate the role of phase separation in transcription is published under shared first authorship with Jorge Trojanowski [146]. Contribution of co-authors to data acquisition or analysis are stated at the corresponding parts of the *Results* and the *Materials and Methods* sections.

1.1 A toolbox of modular effectors mimics TF binding and activity

Local enrichment of effectors on chromatin that carry a given biological activity can be harnessed to induce chromatin states or nucleate defined (ribo-) nucleoprotein assemblies on the chromatin scaffold. For chromatin regions that contain clustered DNA sequences, this can be achieved by fusing effectors to a single sequence-specific DNA-binding domain (DBD), in analogy to prototypical TFs consisting of a transcription activation domain (AD) linked to a DBD. Beyond DNA-protein binding, this modular design principle is extensible by RNA-protein and protein-protein interaction modules with specified structural or functional arrangements. Such a setup is ideally suited to image and manipulate endogenous nuclear compartments or to nucleate synthetic compartments at DNA binding site clusters of interest.

The Blue-light Induced Chromatin Recruitment (BLinCR) system from previous work of our group [65] had already pioneered tools to trigger effector recruitment to a reporter gene cluster by light-dependent PHR-CIBN dimerization [147]. Likewise, effectors based on catalytically dead Cas9 (dCas9) have been created to perturb and image gene activity [148, 149], also in conjunction with PHR-CIBN [150]. In this thesis, I aimed at expanding the combinations of DBDs, ADs/effectors and linker modules that had previously been described elsewhere [65, 149-151] and specifically tailored them to the task of probing and perturbing features of chromatin-associated compartments. The feature space covered by my construct toolbox focused on modulating the following key properties: (i) turnover on chromatin, (ii) transcriptional activation strength and (iii) self-interaction and phase separation propensity.

Constructs were based on two types of DBDs: the *lacO/tetO* binding Lac repressor (LacI) and reverse Tet repressor (rTetR) proteins that were already established in the BLinCR system [65] and dCas9 to target sequences by programmable single-guide RNAs (sgRNAs). For the effectors, I focused exclusively on activation domains (ADs) from TFs, as transcription is linked to chromatin state changes and is well quantifiable through a range of established assays. I selected VP16, VPR, STAT2, p65 and Rta, which have varying activation strengths and are either of viral, human or mixed origin [151-154]. VPR is a synthetic activator consisting of a fusion of VP64 (four copies of VP16), p65 and Rta. Four different designs were chosen to link a given AD effector to the above-mentioned DBDs: (i) direct DBD-AD fusion (ii) fusion of the AD to tandem PP7 coat protein (tdPCP), bound to PP7 loops engineered into the sgRNA [149], and ultimately light-dependent heterodimerization between PHR-AD and (iii) DBD-CIBN or (iv) tdPCP-CIBN fusion proteins. The linkage designs were termed direct, loop, opto and optolooop, respectively and are additionally distinguished by the type of DBD used (**Figure 4**). The properties of the DBD and linker modules covered a large range of binding modes at the target site.

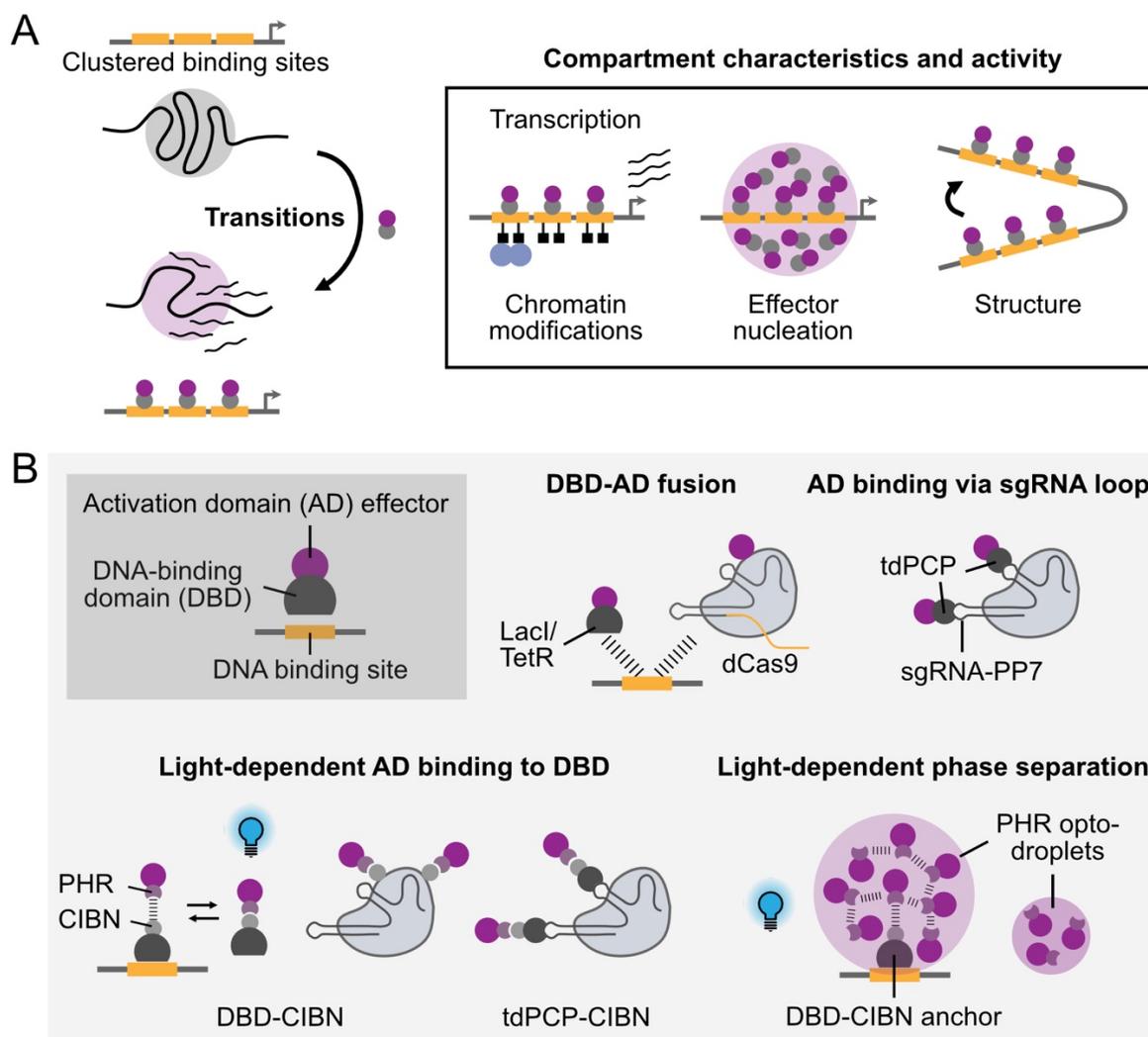


Figure 4. A toolbox of modular effectors to observe and perturb transitions, structure and activity of nuclear compartments. (A) Recruitment of effectors to a genomic region with clustered DNA binding sites can be used to evoke functional compartment transitions and at the same time study compartment characteristics by microscopy, including activity (transcription, chromatin state), the amount and binding mode of nucleated protein and structure. (B) Toolbox of modular effector constructs that follow the AD-DBD design principle of endogenous TFs. ADs are linked to the DBD (Lacl, TetR or dCas9) by direct fusion, recruited via PP7 loops in the sgRNA that targets dCas9 (loop) or in a light-dependent fashion through CIBN-PHR dimerization modules. These can be attached to the DBD (opto) or via sgRNA loops (optoloop). Self-interaction of the PHR domain to form optodroplets also enables light-dependent nucleation on chromatin via phase separation.

The use of PHR-CIBN domains enabled fast control over recruitment [65] and a way to induce AD compartmentalization by a light-induced phase separation mechanism [78] that is based on self-interaction of the PHR moiety (see *Results section 1.5*). In all constructs, the ADs were tagged with fluorescent protein domains in order to follow their subcellular localization in living cells. In conclusion, the constructs aimed at manipulating chromatin by means of triggering transcription and assembling large protein effector complexes with specified composition or binding mode at genomic regions of interest.

1.2 Directing transcriptional activators to mouse major satellite repeats interrogates the chromocenter compaction mechanism

Chromocenters arise from the interaction of PCH regions of multiple chromosomes [155] but the exact nature of their repressive state and regulation of their compaction remain poorly understood. Genetic perturbations that affect their compact state have been identified to some extent [156], but targeted tools to induce full de-compaction for mechanistic dissection are lacking. HP1 has been hypothesized to cause PCH compartmentalization into chromocenters by a phase separation mechanism [38, 157]. However, recent findings by our group called this model into question [57], and the links between the H3K9me3-SUV39H-HP1 and compaction require further mechanistic dissection. In addition, a role of MSR transcripts in maintaining PCH repressed through feedback interactions with SUV39H enzymes has been proposed [24]. But since MSR transcripts are expressed at very low levels under physiological conditions, studying their biological role remains challenging.

I set out to identify determinants of chromocenter compaction by manipulating the heterochromatic state of PCH using activator constructs from the here established toolbox. I furthermore devised microscopy-based readouts to obtain quantitative information on the structural and functional changes that occur during the induced heterochromatin re-activation. Due to their elevated DNA density and higher A/T content, chromocenters are intensely stained by the DNA dye DAPI. This signal can be used to read out their compaction. Further, immuno-staining can be performed in addition, to detect chromatin marks that are linked to the induced perturbation.

The arrays of MSR sequences contained within PCH can be targeted by dCas9 constructs using a single sgRNA targeting sequence [158]. By this, effectors can be specifically enriched in chromocenters. Since PCH is protected by a redundant network of silencing factors, strong and universal activation triggers are needed to override its repressive state. The synthetic activator VPR has proven as potent, context-independent activator [151]. In addition, direct interactions of the histone acetyl-transferases CREB-binding protein (CBP) and p300 have been reported for the VPR subcomponent p65 [159]. I therefore hypothesized that recruiting VPR to MSRs could suffice to evoke PCH activation and trigger de-compaction, potentially by a mechanism that would involve the deposition of activating chromatin marks, such as histone H3 acetylation at lysine 27 (H3K27ac), which also influence chromatin compaction.

I established targeting of GFP-tagged dCas9 constructs with or without VPR activator to chromocenters in immortalized mouse fibroblasts (iMEFs) using the previously described MSR sgRNA [158]. Detection of targeted dCas9 was combined with H3K27ac immuno-staining and DNA staining to assess active chromatin and compaction, respectively.

Furthermore, wild-type (wt) iMEF cells were compared to an isogenic cell line lacking the H3K9 methyltransferases SUV39H1 and SUV39H2 (*Suv39h* double-null, short dn). iMEF dn cells lack H3K9me3 and HP1-alpha enrichment at chromocenters [55, 57]. This comparison was included to study the impact of the H3K9me3-SUV39H-HP1 axis on chromocenter activation and de-compaction. Transfected GFP-tagged dCas9 constructs were successfully targeted to MSRs in chromocenters, as evidenced by their enrichment in the characteristic DAPI-dense nuclear foci (Figure 5 A).

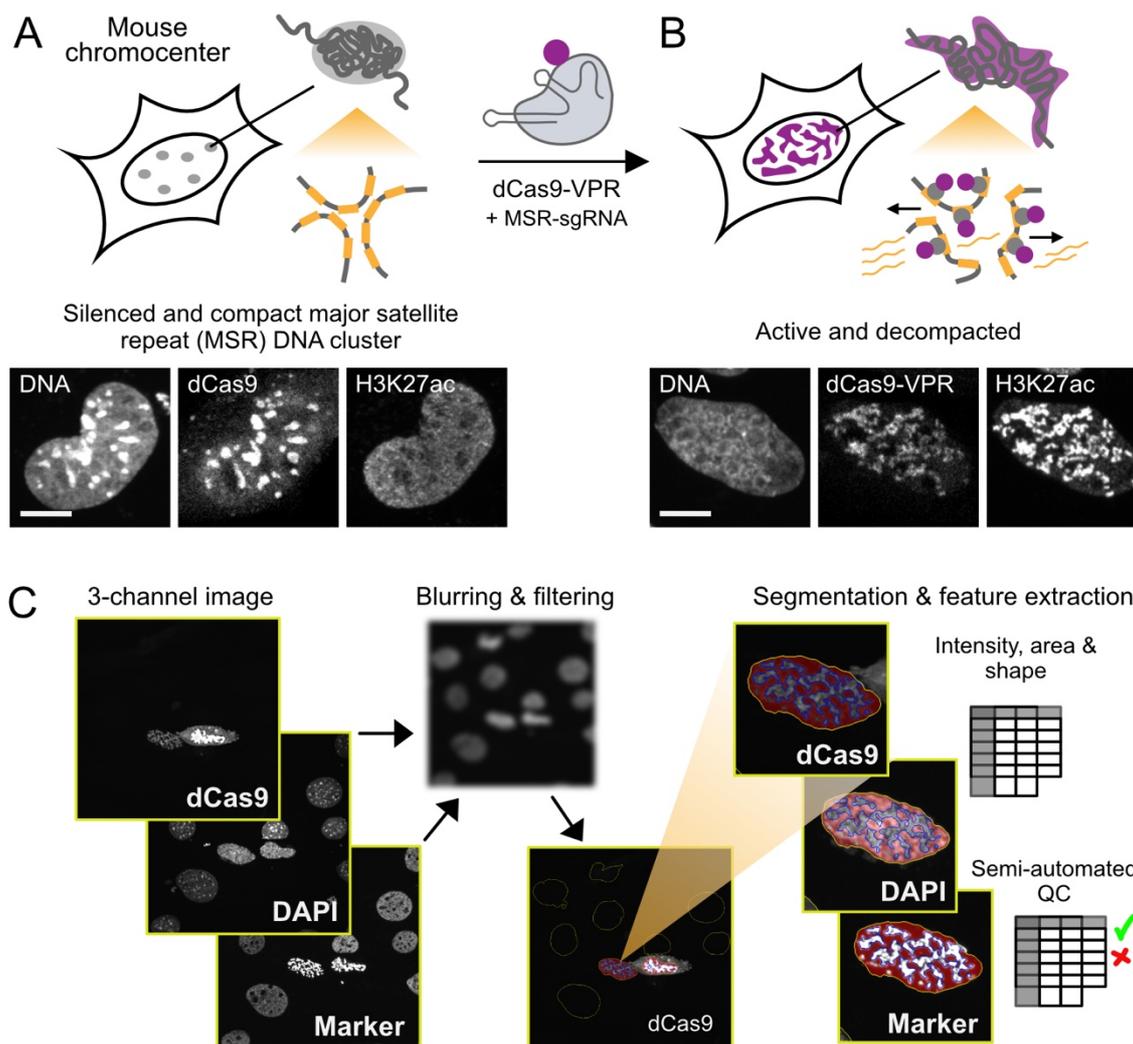


Figure 5. Experimental setup to quantify resulting changes in chromocenter organization by imaging upon activation of mouse pericentric heterochromatin. (A) Recruiting dCas9-GFP as mock effector to MSRs in iMEF cells preserves the organization of PCH into DNA-dense chromocenters (left) (B) Recruitment of a dCas9-GFP-VPR activator fusion evokes PCH de-compaction and area expansion and causes localized H3K27ac deposition as detected by immuno-staining. Scale bars, 10 μ m. (C) Automated workflow for the imaging-based analysis of chromocenter changes upon perturbation. Analysis of DNA/DAPI, dCas9 and an additional channel (marker) by segmentation of nuclei (DAPI) and chromocenters (dCas9) is followed by intensity, area and shape feature extraction and quality check (QC) of the final results. Figure modified from ref. [145].

dCas9-VPR also accumulated in micrometer-sized areas of the nucleus. However, these areas were widespread and appeared fibrous in their structure (**Figure 5 B**), suggesting that chromocenters had de-compacted upon binding of VPR. To quantitate even small changes in compaction and H3K27ac for single cells, I developed an R programming language [160] based image analysis workflow applicable to fluorescence microscopy images of chromocenter-perturbed cells (**Figure 5 C**) [145]. In brief, the method uses intensity-based segmentation of the DAPI (DNA) and GFP (dCas9) signals in each cell to produce masks for the cell nucleus, chromocenters, and nucleoplasm. Quantifications of area, shape, and intensity are carried out in these masks, followed by a semi-automated workflow to inspect the segmentation results for each individual cell. Curated results in a universal tabular format are then generated for downstream analysis. The automated approach produced a considerably more scalable and unbiased alternative to manually assessing the microscopy images.

Quantitative analysis of the H3K27ac and DNA enrichment in dCas9-bound areas (**Figure 6 A, E**) versus their surrounding nucleoplasm revealed a marked increase in acetylation and de-compaction with increasing amount of bound dCas9-VPR (low, medium, high) (**Figure 6 B-D, F-H**). In contrast, DNA density was higher and acetylation levels were much lower in chromocenters targeted by dCas9 only fused to GFP (mock). The strong structural changes upon VPR recruitment were comparable between iMEF wt and iMEF dn cells lacking HP1 and H3K9me3 at chromocenters (**Figure 6**). Chromocenter area correlated well with the level of bound VPR (**Figure 6 I**) and already low H3K27ac levels as a result of VPR recruitment, were associated with a strong de-compaction phenotype (**Figure 6 J**).

The results from the perturbation experiments indicated that already low levels of bound VPR were sufficient to strongly acetylate and decondense chromocenters and that, strikingly, the large-scale transition to a de-compacted state did neither depend on HP1 nor H3K9me3. To the best of my knowledge this is the first account of an activator that can de-compact mouse chromocenters to the degree demonstrated here. MSR transcription induction upon VPR recruitment was not measured here. However, additional experiments found positive immunostaining for the CTD serine 5-phosphorylated (elongating) form of RNA polymerase II at VPR-perturbed chromocenters [57] as indicator of transcription. Thus, the use of the VPR will likely also prove useful to induce expression of the lowly-abundant MSR transcripts arising from PCH regions and study their function. In conclusion, mechanisms that underly PCH compartmentalization could be studied using effectors from the toolbox established here and combined with quantitative image analysis. By exchanging the activators or additionally perturbing histone modification pathways, the present experimental setup can be easily extended to dissect other structure-function relationships in PCH or in other targetable chromatin-associated compartments.

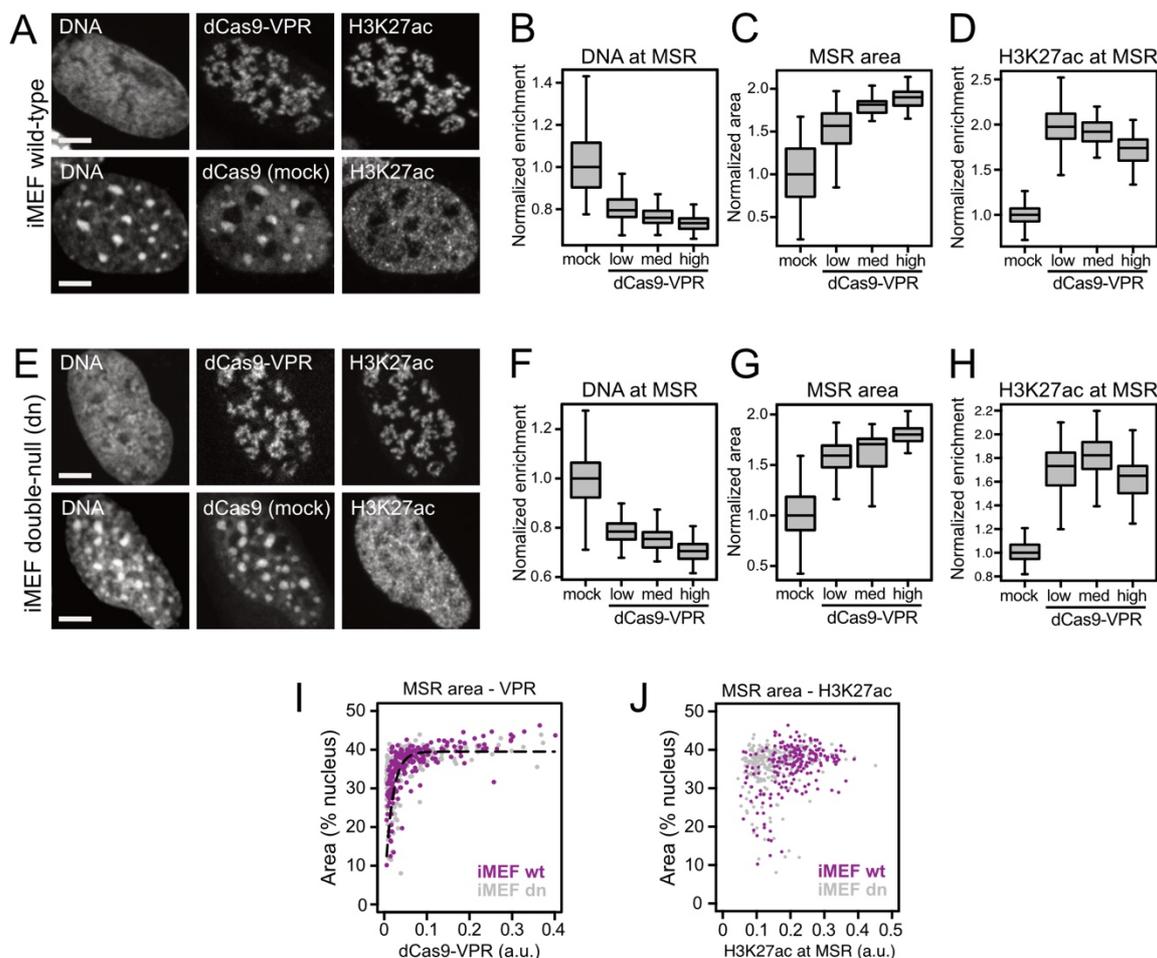


Figure 6. Quantitative analysis of chromocenter changes upon VPR targeting to MSRs in *Suv39h* wild-type (wt) and double-null (dn) iMEF cells. iMEF cells were targeted with GFP-tagged dCas9-VPR or dCas9 (mock) constructs. 30 hours post-transfection, cells were stained by DAPI (DNA) and H3K27ac immuno-staining. (A) Representative images of iMEF wt cells. (B-D) Normalized DNA enrichment, MSR area and H3K27ac levels as a function of dCas9 enrichment (low, medium, high) at MSRs. Mock: 155 cells. dCas9-VPR low, medium and high: 83, 42 and 83 cells, respectively. Data was normalized to the respective mean value of the mock condition. (E) Representative confocal images of iMEF dn cells. (F-H) Same as B-D, but for iMEF dn cells. Mock: 148 cells. dCas9-VPR low, medium and high: 95, 48 and 95 cells, respectively. (I) Relationship of the relative nuclear area covered by MSRs and the amount of MSR-bound dCas9-VPR in all analyzed wt and dn cells. The line represents an exponential function as a guide to the eye. (J) Relationship between relative MSR area and the amount of H3K27ac at MSRs. All scale bars, 5 μ m. Figure panels adapted from ref. [57].

1.3 Assembling TF-like effectors at a reporter gene cluster links readouts of compartmental characteristics to transcription

To investigate how compartmental characteristics related to actively transcribed gene clusters are established and regulated, I took advantage of the U2OS 2-6-3 reporter cell line [65, 161], which contains an artificial inducible reporter gene cluster. In contrast to the transcription of single genes, whose time-resolved measurement is technically challenging in single cells, the here used reporter cell line contains tandemly repeated copies of an MS2 loop reporter RNA transgene unit (**Figure 7 A**). Each transgene unit is controlled by a CMV minimal promoter flanked by 256 *lacO* and 96 *tetO* sites. These clustered binding sites can be targeted by dCas9 with *lacO/tetO*-specific sgRNAs or directly bound by LacI and (r)TetR proteins to deliver effectors. By default, the reporter is in a silenced state. However, transcription can be induced to high levels by recruitment of ADs like VP16 [65, 161].

Due to its high binding site valency and repetitiveness, it represents a model for studying condensate nucleation on chromatin and other regulatory mechanisms that involve the sharing of regulatory factors between genes in the same cluster. Its size (~1 μm) and position at a single genomic locus allows direct fluorescence microscopy-based analysis of its features in both in living and fixed cells without signal amplification. Immuno-staining can be used to assess its chromatin state, in a similar fashion to the mouse chromocenters introduced previously. Fluorescent tandem MS2 coat protein (tdMCP) allows for monitoring of nascent RNA synthesis at the reporter locus in real time (**Figure 7 A**, right panel). Other fluorescently-tagged constructs (e.g., co-activators) can be introduced to test for interactions with the reporter locus or endogenous proteins.

In a first set of experiments, I used the reporter cell line to validate the binding of the here established toolbox components (**Figure 7**) and explore their activation properties and effects on the chromatin state (**Figure 8**). To test the binding, VP16 was recruited to the *lacO* repeats in all possible construct configurations (direct, loop, opto, optoloop), visualized either by a GFP or an mCherry tag. TetR constructs were used to locate the reporter. Following transfection, all constructs were successfully enriched at the target site, as seen by co-localization with TetR, which labeled the *tetO* repeats (**Figure 7 B**). I could also confirm that the light-dependent components CIBN and PHR were functional (with opto and optoloop with dCas9 or LacI as DBD) (**Figure 7 C**) and behaved as in the BLinCR setup [65]. I next tested the constructs' ability to activate, focusing on VP16 and VPR as comparison between two ADs with reportedly different strength (**Figure 8**).

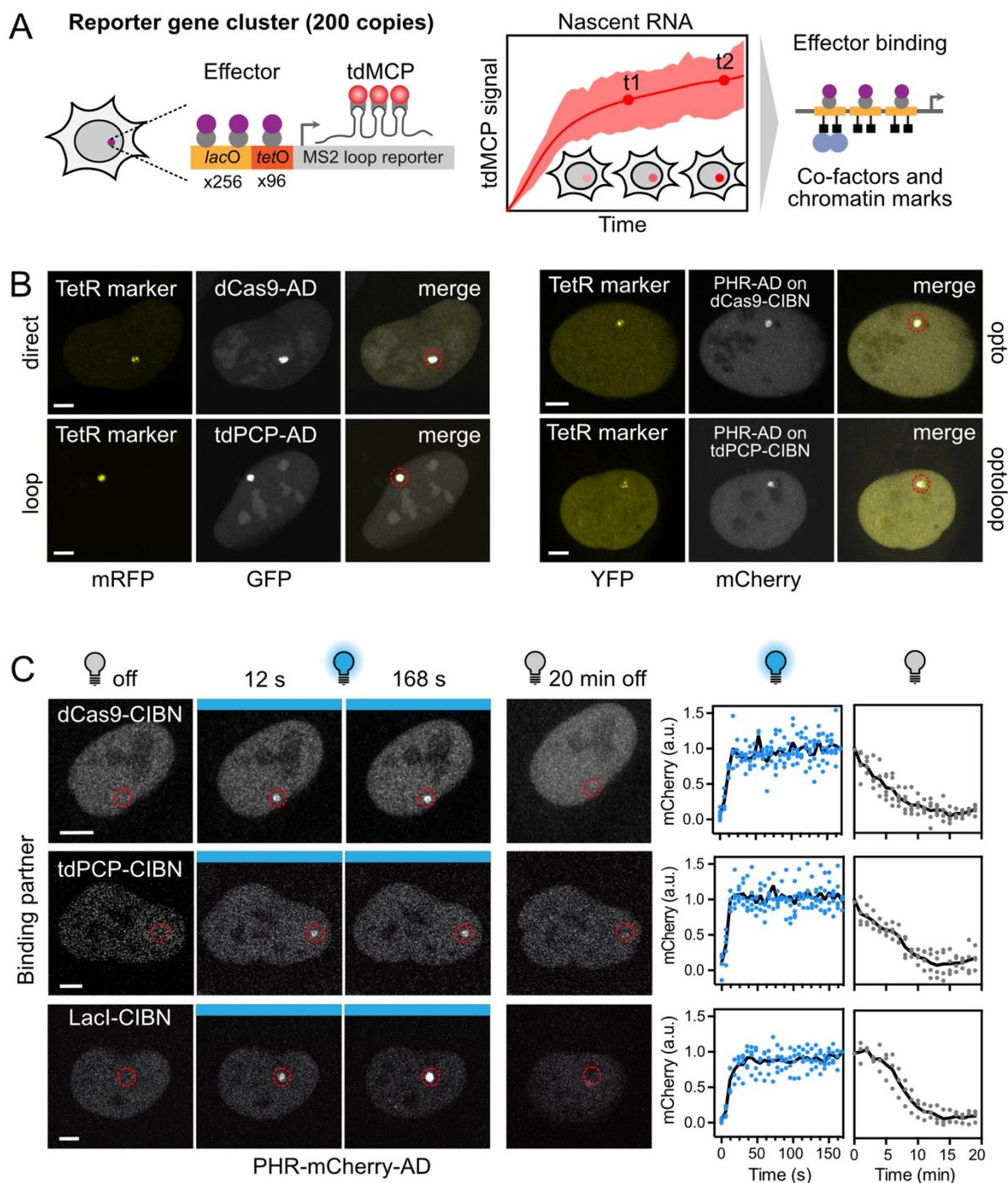


Figure 7. Binding of toolbox effectors at a reporter gene cluster. (A) The U2OS 2-6-3 reporter gene cluster [161] contains *lacO* and *tetO* binding sites for effectors that can induce transcription of an MS2 loop RNA detected by fluorescent tdMCP. Time-resolved transcription can be recorded alongside co-factor enrichment and chromatin marks. (B) Representative images of U2OS 2-6-3 cells expressing dCas9 fusion, dCas9-loop, dCas9-opto and dCas9-optoloop constructs that bring VP16 as activation domain to the *lacO* sites. TetR was co-transfected as marker. (C) Representative images and light-dependent binding kinetics of 3-5 cells expressing the opto/optoloop constructs. Intensities were normalized to the first (association) or last (dissociation) timepoint. Solid lines depict averaged intensity. All scale bars, 5 μ m. Figure panels adapted from ref. [146]

To this end, the constructs were recruited to the *tetO* repeats and qRT-PCR measurements of total reporter RNA levels after 24 hours revealed robust activation for most constructs (**Figure 8 A**). Overall, VPR lead to a ~200 to 500-fold induction of reporter RNA, while VP16 induction ranged between ~5 to 35-fold. To quantify chromatin changes possibly linked to the level of activation and the different construct topologies, enrichment of BRD4 and H3K27ac were assessed by microscopy as two prototypical measures of transcriptionally active chromatin (**Figure 8 B, C**). I observed that high levels of total reporter RNA were generally correlated to elevated BRD4 and H3K27ac profiles. Also, nascent MS2 transcripts could still be detected by tdMCP labeling, indicating that the reporter gene cluster stayed in a transcriptionally active state, even 24 hours after induction.

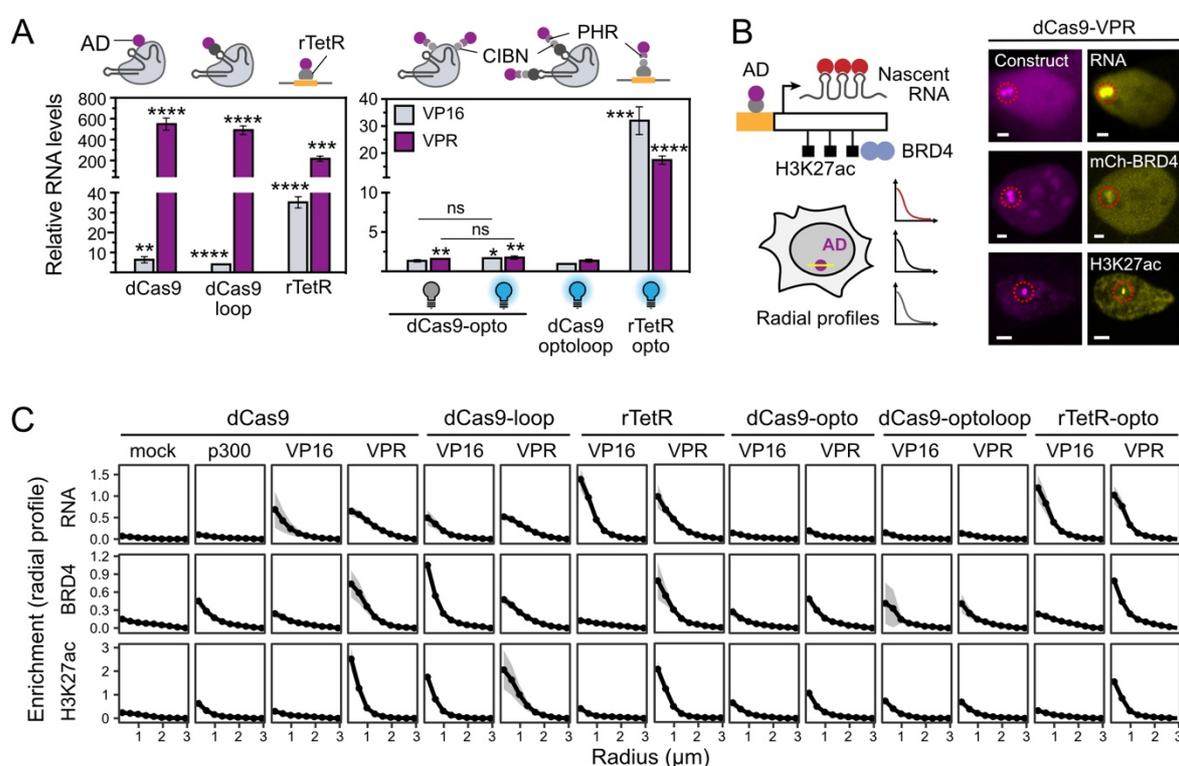


Figure 8. Transcription and chromatin changes induced by toolbox effectors targeted to the reporter gene cluster. The ADs VP16 and VPR were compared for all construct topologies. (A) Reporter RNA levels normalized to β -actin mRNA and relative to mock-transfected samples were quantified by qRT-PCR 24 hours post-induction. Mean and s.d. of $n = 3$. rTetR and opto(loop) construct activation involved the addition of doxycycline and/or constant illumination. Two-sided unpaired Student's t test, with not significant (ns), * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $P < 0.0001$ (B) Representative images and scheme for radial enrichment analysis of steady-state nascent RNA, co-transfected mCherry-BRD4 and H3K27ac immuno-staining signals at the studied reporter. (C) Radial enrichment profiles of RNA, BRD4 and H3K27ac for all construct topologies examined with either VP16 or VPR effectors recruited to the *tetO* sites ($n = 16$ -250 cells per condition). dCas9 without AD (mock) and dCas9-p300 served as controls. Radial profile acquisition and analysis was done jointly with Jorge Trojanowski. All scale bars, 5 μ m. Figure adapted from ref. [146].

Overall, VPR was a more potent activator than VP16, and in the majority of cases, showed a markedly stronger enrichment of BRD4 and H3K27ac than VP16. No nascent transcripts and no H3K27ac and BRD4 enrichment were detectable when recruiting dCas9 alone as non-activating control. As an additional assay to validate the specificity of the BRD4 and H3K27ac signals, dCas9 was also coupled to the histone acetyltransferase p300 core domain (p300), which produced no RNA, but as expected, increased H3K27ac and BRD4 at the reporter.

The dCas9-opto and dCas9-optolooop constructs did not activate the reporter. They nevertheless produced a BRD4/H3K27ac pattern that was comparable to that of dCas9-p300. This finding was relatable to the context-dependent activity previously reported for optogenetic dCas9 constructs [150], but showed that transcription and co-activating marks can be uncoupled under certain conditions. In summary, the reporter gene cluster assays revealed that the designed effectors were able to induce transcription and chromatin changes to an active state and that these can be quantified by fluorescence microscopy. Furthermore, the observed construct topology-dependent range of activities, offered ways to alter and explore activation features in isolation.

1.4 Combinatorial use of TF constructs distinguishes strong and weak ADs based on their functional interaction with co-activators

The co-activator protein BRD4 binds to acetylated histones via its bromo-domains [162]. Together with other factors, it accumulates on active regulatory elements and promoters, thereby modulating transcription at both the initiation and elongation step [63, 163]. RNA Pol II and other factors have been proposed to form phase-separated transcriptional condensates with BRD4 to regulate transcription [42]. The protein's large C-terminal IDR is thought to be responsible for this co-compartmentalization behavior. However, it remains unclear to which extent these bromodomain-independent interactions also take place with TFs directly and how they would contribute to the activation process. The co-occurrence of several events during the transcription activation process makes it difficult to accurately deconstruct such mechanistic relationships. However, the effector toolbox validated in *section 1.3* provides a range of effectors with different properties and can be applied to the reporter gene cluster as a model transcription compartment. The previous characterization had revealed several constructs with distinct activities. Notably, acetylation and BRD4 levels at the reporter locus could be raised by the dCas9-p300 and dCas9-optolooop constructs without inducing transcription and comparison of the ADs VPR and VP16 identified VPR as the stronger activator, also causing more BRD4 accumulation and H3K27ac deposition. I next exploited these unique construct properties to look into how both BRD4 interaction and histone acetylation affect the activity of VP16 and VPR at the reporter gene cluster.

In a first step, assays were devised to not only determine the steady state interaction of BRD4 with VP16 and VPR, but also capture its dynamics in absence of transcription as well as its functional impact in the presence of transcription (**Figure 9**). BRD4 localization in living cells was followed by transfecting mCherry-BRD4. For highest temporal control over AD recruitment and examining the BRD4 without triggering transcription, BRD4 intensity was measured at the reporter after recruiting VP16 or VPR by the inactive dCas9-optolooop construct (**Figure 9 A, B**). Compared to VP16, VPR caused BRD4 to accumulate to higher levels more quickly. For VPR but not VP16, there was an initial sharp surge in BRD4 recruitment over the first 10 min, which was followed by a slower period (**Figure 9 C**, control). BRD4 accumulation for VP16 was entirely eliminated by pretreatment with the bromodomain binding inhibitor JQ1 (**Figure 9 C**, JQ1). VPR's first phase was unaltered, however the second phase of BRD4 enrichment was found to be JQ1-dependent. These findings suggested that both bromodomain-dependent and independent BRD4 interactions occur with these ADs, even in absence of transcription. However, bromodomain-independent interaction was predominantly seen for VPR and not for VP16, possibly pointing at a so far undescribed VPR-specific BRD4 interaction.

Next, to examine how BRD4 modulated transcription, the JQ1 experiment was repeated using the transcriptionally active rTetR-opto-VP16 and -VPR constructs under comparable conditions (**Figure 9 D**). Upon JQ1 pretreatment, total reporter RNA levels as measured by qRT-PCR were largely unchanged for VP16 (1.1-fold reduction), but reduced for VPR (1.8-fold reduction). Thus, also bromodomain-dependent BRD4 binding can improve activation by VPR. To determine how histone acetylation was related to the differences between the two ADs, the transcriptional output of VP16 and VPR in a high versus low acetylation chromatin context was tested (**Figure 9 E**). dCas9-p300 was employed to acetylate the reporter's *lacO* sites before inducing transcription with rTetR-opto-VP16 and -VPR recruited to the *tetO* sites. In the presence of dCas9-p300, VP16 produced more reporter RNA than the control (dCas9-GFP), while VPR did not (**Figure 9 F**). Thus, VP16 may be less active than VPR due to a reduced ability to bind BRD4 and attract HATs, making it more reliant on pre-existing histone acetylation. In conclusion, the combinatorial use of the toolbox components dissected behavior of the BRD4 co-activation pathway to explain differences in the activity of two ADs. The here devised approaches can be further applied to investigate the interaction of other TF-derived domains or quantify the involvement of other co-activators like e.g., MED1 and TAF15 in the activation process.

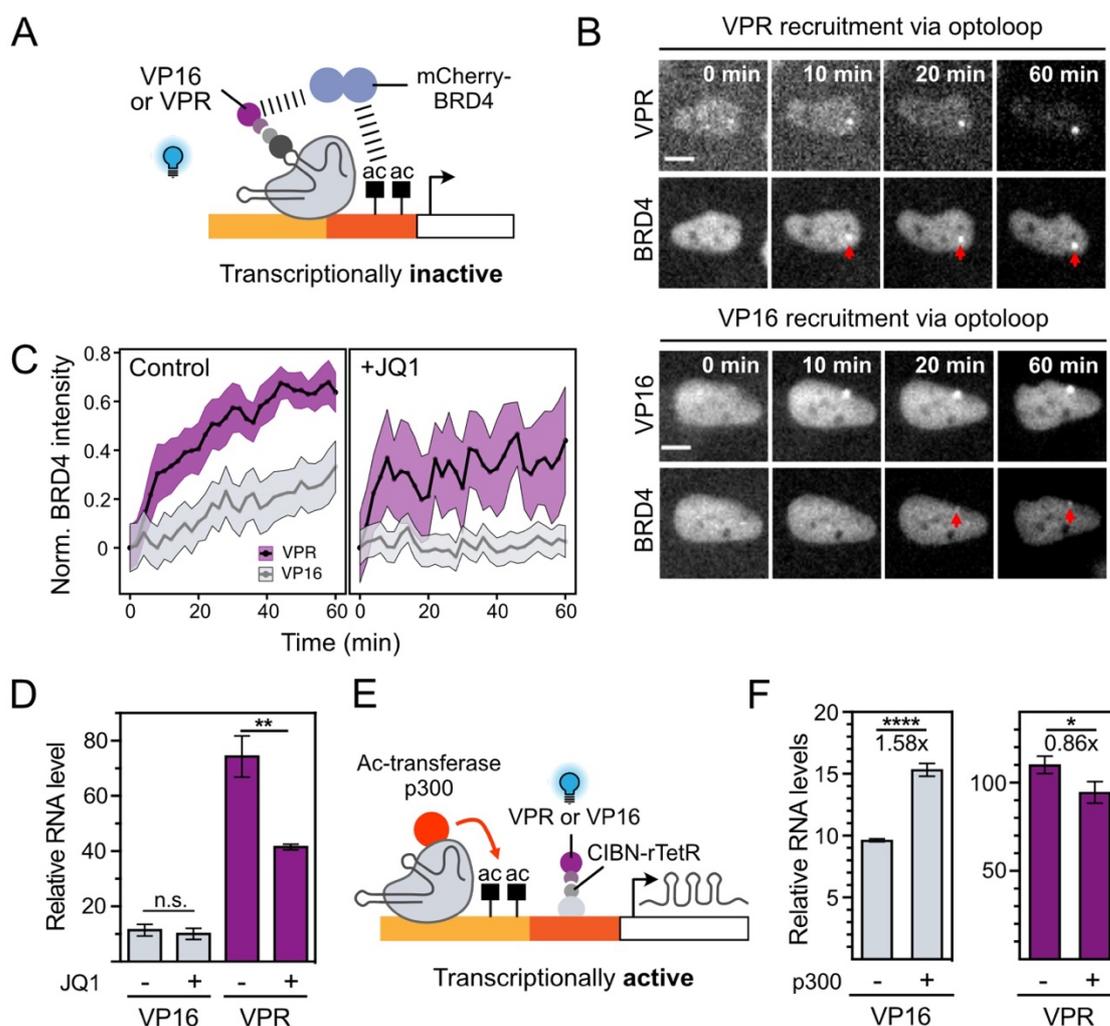


Figure 9. Combinatorial use of effector constructs from the toolbox dissects interrelations of transcription, BRD4 binding and histone acetylation. (A) Setup to monitor differential BRD4 co-enrichment at the reporter for VP16 and VPR in absence of transcription. The inactive dCas9-optolooop construct was used for recruitment to both *tetO* and *lacO* to improve detection of the mCherry-BRD4. (B) Representative time series of BRD4 enrichment at the reporter (arrows). Scale bars, 10 μ m. (C) Temporal dynamics of BRD4 accumulation at the reporter after light-induced AD binding (at 0 min), both with and without JQ1 pre-treatment (n = 10-85 cells per condition). Mean values of normalized intensity with 95% CI as ribbons. (D) Relative reporter RNA levels (qRT-PCR) measured 90 min after recruiting transcriptionally active rTetR-opto-VP16 or VPR, with and without JQ1 pre-treatment as done in C. (E) Experimental setup to test how pre-deposition of acetylation by dCas9-p300 (*lacO*) affects transcriptional output of the rTetR-opto-VP16/VPR (*tetO*) constructs. (F) Reporter RNA quantification by qRT-PCR for E, 90 min post-recruitment. Binding of dCas9 alone (-) to *lacO* was used as control. All qRT-PCR data was normalized to β -actin mRNA and to mock-transfected samples. Bar plots depict mean and s.d. of n = 3 replicates. Statistical analysis: Two-sided unpaired Student's t test, with not significant (ns), *P < 0.05, **P < 0.01, and ****P < 0.0001. Figure panels adapted from ref. [146]. BRD4 time courses were acquired and analyzed by Jorge Trojanowski.

1.5 PHR optodroplets model how phase-separated TF compartments form around chromatin and affect transcription

Phase separation has been proposed to compartmentalize transcription-related proteins on chromatin and coined as a process that enhances TF activity [76]. One caveat is provided by the lack of assays that employ functional readouts at high temporal and spatial resolution, but at the same time grant control over condensate formation and composition in living cells. Depending on the specific biological context, endogenous transcriptional condensates may not be amenable to mechanistic dissection. In contrast, the modular reporter cell line and toolbox framework developed here, is broadly applicable to interrogate these features in a bottom-up approach. Individual components (DBDs, ADs, readouts, etc.) can be flexibly exchanged, to provide control over molecular properties, such as DNA binding, self- and co-activator interaction. In this section, the light-dependent properties of the toolbox were harnessed to artificially create phase-separated compartments around the reporter gene cluster and test their functional impact on transcription (**Figure 10** and **Figure 11**). For this, I used the second light induced activity of the PHR domain, which is its oligomerization into so called PHR optodroplets [78] (**Figure 10 A**). These condensates can be controlled by light and concentration, and anchored to chromatin with CIBN-localizer constructs. Three lines of experiments were conducted: (i) The conditions and critical concentration (C_{crit}) for optodroplet formation were established for fluorescently-tagged VP16, and VPR, as well as three additional ADs (STAT2, p65 and Rta) (ii) The state of AD assemblies at the reporter was characterized below and above C_{crit} . (iii) Stratification of cells according to the phase separation propensity and to C_{crit} was carried out to compare the transcriptional response in presence and absence of the phase-separated compartments.

To quantify the phase separation propensity of the five ADs, GFP-labeled PHR fusions were recruited to the reporter via rTetR-opto and imaged after a fixed period of illumination. The cumulative area of all droplets in relation to the nuclear area was then measured and related to the nuclear fluorescence intensity as a proxy for concentration (**Figure 10 B**). C_{crit} was determined from the intersection of a logistic fit of the data and threshold value set at 1% relative droplet area (**Figure 10 C**). The visual classification of cells containing droplets and this value matched. The assay revealed that Rta, p65, and VPR were able to form droplets at much lower concentrations than VP16 and STAT2. The ADs examined here were thus divided into two groups: those with low (VP16, STAT2) and high (Rta, p65, VPR) phase separation propensity. Next, it was investigated whether PHR-AD concentrations above C_{crit} also resulted in droplet nucleation at the reporter gene cluster under these conditions. Confocal microscopy revealed that PHR-AD construct accumulation was considerably stronger and more granular above C_{crit} (**Figure 10 D**).

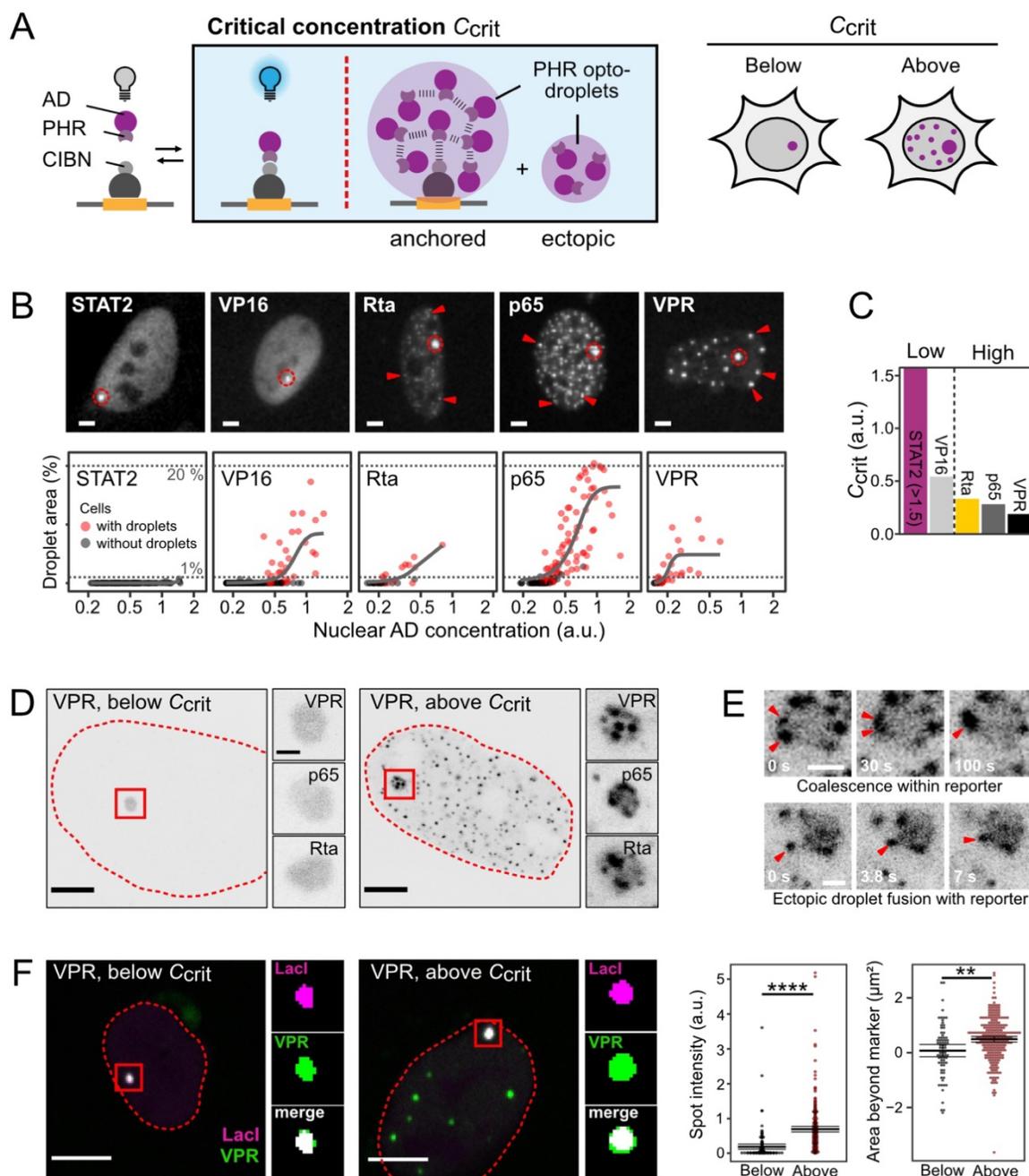


Figure 10. Building phase-separated activator compartments on chromatin by light and concentration dependent PHR oligomerization. (A) PHR-AD fusions guided to DNA by dimerization with CIBN anchors form PHR-PHR optodroplets above a critical concentration C_{crit} . If not stated otherwise, PHR-GFP-AD recruitment to rTetR-opto was triggered by illumination for 10 min. (B) Top: U2OS 2-6-3 cells recruiting AD-PHR fusions to the reporter gene cluster (dashed circle) and forming optodroplets (arrows). Scale bars, 5 μ m. Bottom: C_{crit} was determined from the intersection of a logistic fit on the data and the 1% nuclear area threshold, with additional visual droplet positive (red) and negative (gray) classification. (C) C_{crit} values for the different ADs, stratified into low (STAT2, VP16) and high (Rta, p65, VPR) propensity to form droplets. (D) Live cell imaging of VPR, p65 and Rta recruitment in cells below and above C_{crit} . Cell for VPR (scale bar, 5 μ m) next to zoom-ins (scale bar, 1 μ m) for all ADs. (E) Droplet-like coalescence of Rta assemblies at the reporter. Scale bars, 1 μ m. (F) Imaging of fixed cells expressing VPR and LacI marker. Zoom-ins show segmentation of signals to quantify intensity and area below ($n = 121$ cells) and above ($n = 255$ cells) C_{crit} (right). Scale bars, 10 μ m. Two-sided unpaired Welch's t test: ** $P < 0.01$, and **** $P < 0.0001$. Fitting in B-C and quantification in F were conducted by Jorge Trojanowski. Figure panels adapted from ref. [146].

As shown for PHR-VPR, fluorescence signal was 3.6-fold higher at the reporter and covered a larger area beyond the LacI marker in cells above C_{crit} (**Figure 10 F**). This indicated the enrichment of additional molecules beyond the level observed at below C_{crit} . The reporter-localized assemblies resembled diffusible optodroplets in their appearance, and they were able to coalesce on the reporter cluster as well as with droplets outside of it (**Figure 10 E**). These results confirmed that above C_{crit} , PHR-AD constructs indeed formed assemblies at the reporter gene cluster with model properties of phase-separated compartments.

To test how the additional accumulation by phase separation affected the activation process, the ADs with low (VP16, STAT2) and high (Rta, p65, VPR) phase separation propensity were compared. Nascent reporter RNA production was measured over 90 minutes after binding to rTetR-opto (**Figure 11 A, B**). The duration until half-maximal activation, the RNA plateau level at 90 minutes and the percentage of responsive cells (RNA visible or not) were evaluated. When compared to VP16 and STAT2, the ADs Rta, p65 and VPR showed higher RNA plateau values and were characterized by shorter half-maximal activation times (26-28 min versus 38-42 min). Between 67-92% of the cells responded to Rta, p65 and VPR, while the percentage was lower for VP16 and STAT2 (42 and 67%) (**Figure 11 C, D**). Overall, these results indicated faster and stronger activation of the ADs with high phase separation propensity and were furthermore consistent with the earlier discovered differences in activation strength between VP16 and VPR. The phase-separated state of reporter-bound TFs might be directly responsible for stronger activation. However, since grouping of the five ADs was done based on an arbitrary value for C_{crit} , there could be other factors involved the observed differences. To examine the direct effect of droplet formation on transcription and control for AD-intrinsic properties, the nascent RNA time traces were further divided into cells with and without visible droplets and analyzed separately (**Figure 11 E**).

Strikingly, neither the RNA plateau values and nor the half-activation time showed any statistically significant difference between the two groups ($P > 0.05$, two-way ANOVA accounting for AD and droplet presence) (**Figure 11 F, G**). The obtained results revealed that ADs with a higher tendency for phase separation were stronger activators, but that TF constructs in a droplet state at the reporter gene cluster were not generally associated with an increased transcriptional output.

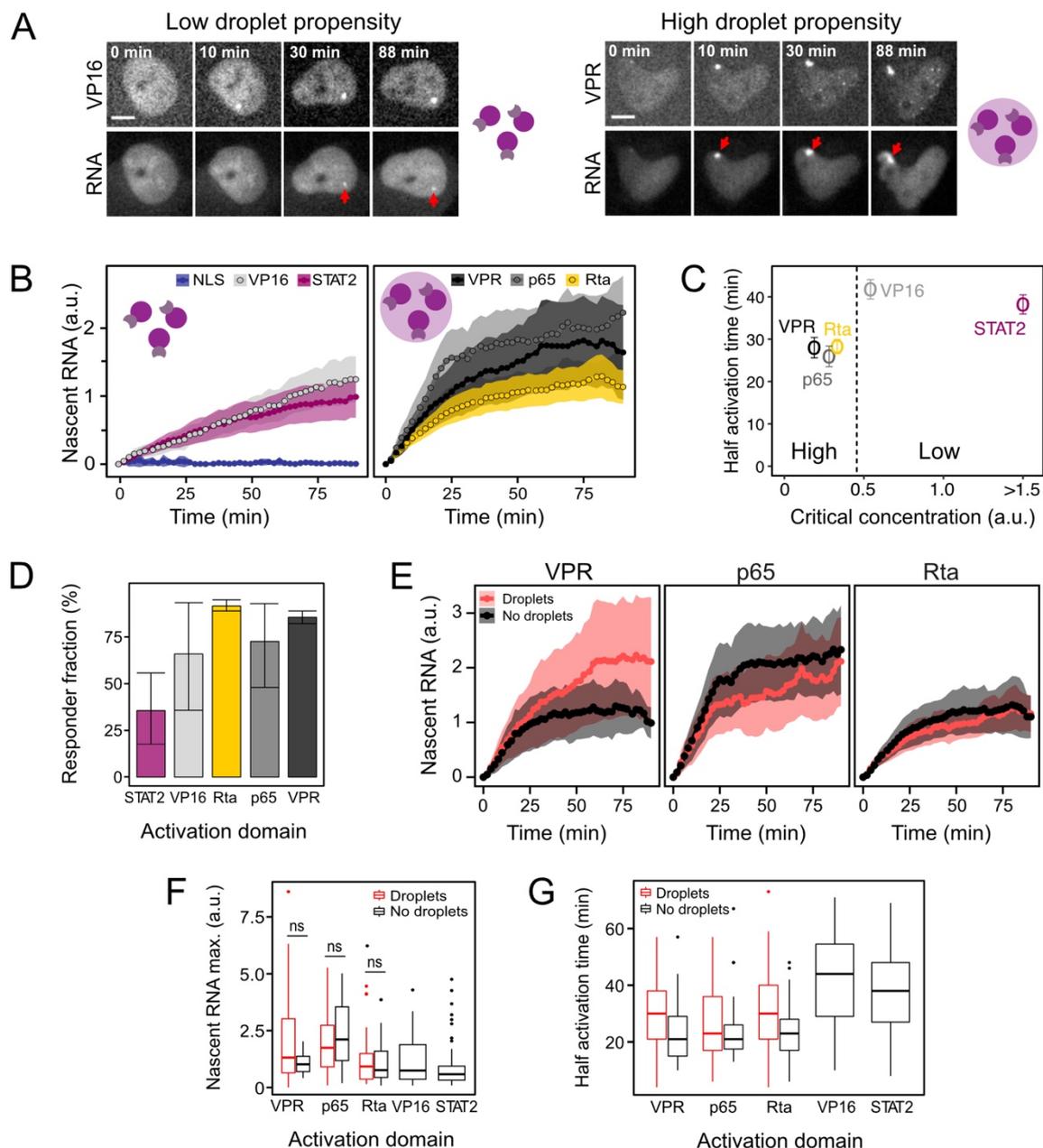


Figure 11. Effects of phase-separated activator compartments on transcription. (A) Representative time series showing PHR-VP16 and PHR-VPR binding to rTetR-opto and droplet formation alongside nascent RNA production (arrows). Light-dependent recruitment started at 0 min. Scale bars, 10 μ m. (B) Averaged nascent RNA production trajectories (ribbon: 95% CI) quantified for all studied ADs split by low and high droplet propensity ($n = 31-71$ cells per condition). NLS does not activate and was used as control. (C) Previously determined C_{crit} values from Figure 10 B and C plotted against time to half-maximal activation determined from the time courses in B. Error bars: s.e.m. (D) Fraction of responding cells (visible nascent RNA) for each AD. Mean with min. and max. values from 2-3 replicates. (E) Averaged nascent RNA trajectories visually classified by droplet presence (ribbon: 95% CI) ($n = 13-18$ cells per condition). (F) Distribution of nascent RNA trajectory plateau values for all cells of each AD and condition ($n = 13-55$ cells per condition). Two-sided unpaired Welch's t test: $P > 0.05$, not significant (ns). (G) Half maximal activation time plotted for the data in F. Activation was slower or similarly fast in droplet-positive and droplet-negative cells. Two-way ANOVA, $p = 0.09$ (n.s.). Nascent RNA trajectories were acquired and analyzed by Jorge Trojanowski. Figure panels adapted from ref. [146].

1.6 Different types of light-induced phase-separated TF droplets do not enhance transcription

The previous experiments established tools for inducing and probing the function of transcriptional condensates in living cells. However, they did not address the issue on how to control the formation of droplets and modulate their composition. I therefore developed constructs that alter the composition of PHR-AD droplets and the concentration at which they form. The PHR-VP16 construct was chosen due to its inherently low phase separation propensity. VP16 was recruited using rTetR-opto. I figured that introducing multivalent linkers would alter droplet formation of VP16. To test for different compositions, three different strategies were applied in parallel (**Figure 12 A**, 1-3): (1) Co-transfection of CIBN-LacI, which forms dimers [164] and therefore bridges two PHR molecules. (2) Adding PHR coupled to GFP-binding protein (GBP), which binds GFP with high affinity and thereby attaches a second PHR domain to PHR-GFP-VP16. (3) Fusion of VP16 to the N-terminal IDR of the fused in sarcoma protein (FUSN), which forms droplets both *in vitro* and *in vivo* [79].

All three approaches lead to a substantial increase in the fraction of cells with visible droplets (**Figure 12 B**). With almost 100% of the cells having droplets, the GBP strategy was the most effective. CIBN-LacI and FUSN-VP16 both increased the fraction to around 50%. So did the PHR-FUSN control without AD. The unperturbed condition showed no droplet formation (0%). This value was lower than previously, but since conditions were assayed in parallel in the same setup, relative comparisons remained valid. After having confirmed that these strategies could produce VP16 droplets with different composition and at lower concentrations, reporter activation was assessed as functional readout. CIBN-LacI or PHR-GBP strongly impaired activation (**Figure 12 C**). Total reporter RNA levels measured by qRT-PCR after 90 minutes dropped by 2.3-fold (CIBN-LacI) or 3.7-fold (PHR-GBP) compared to the controls. Nascent RNA production was also reduced by CIBN-LacI and almost undetectable in presence of PHR-GBP. These changes were also reflected in the proportion of responding cells, which dropped from 70% (unperturbed) to 24% (CIBN-LacI) or 4% (PHR-GBP). Co-transfection of a non-bridging GFP-LacI control led to some inhibition, but to a lower extent, likely due to an independent effect of LacI binding to the reporter. Global changes in transcription signals above/below C_{crit} were also characterized to rule out any side effects of droplet formation. PHR-VPR was included as positive control for droplet formation and a neutral effect of droplets on transcription. In cells with reinforced VP16 or native VPR droplets, ethynyl-uridine (EU) pulse labeling showed no appreciable decline in global RNA levels (**Figure 12 F**). However, when RNA Pol II was inhibited with actinomycin D as a control, global EU incorporation was abolished. Additional control experiments were conducted to exclude mis-localization of the transcriptional machinery or effects of GBP binding on transcription and are shown in ref. [146].

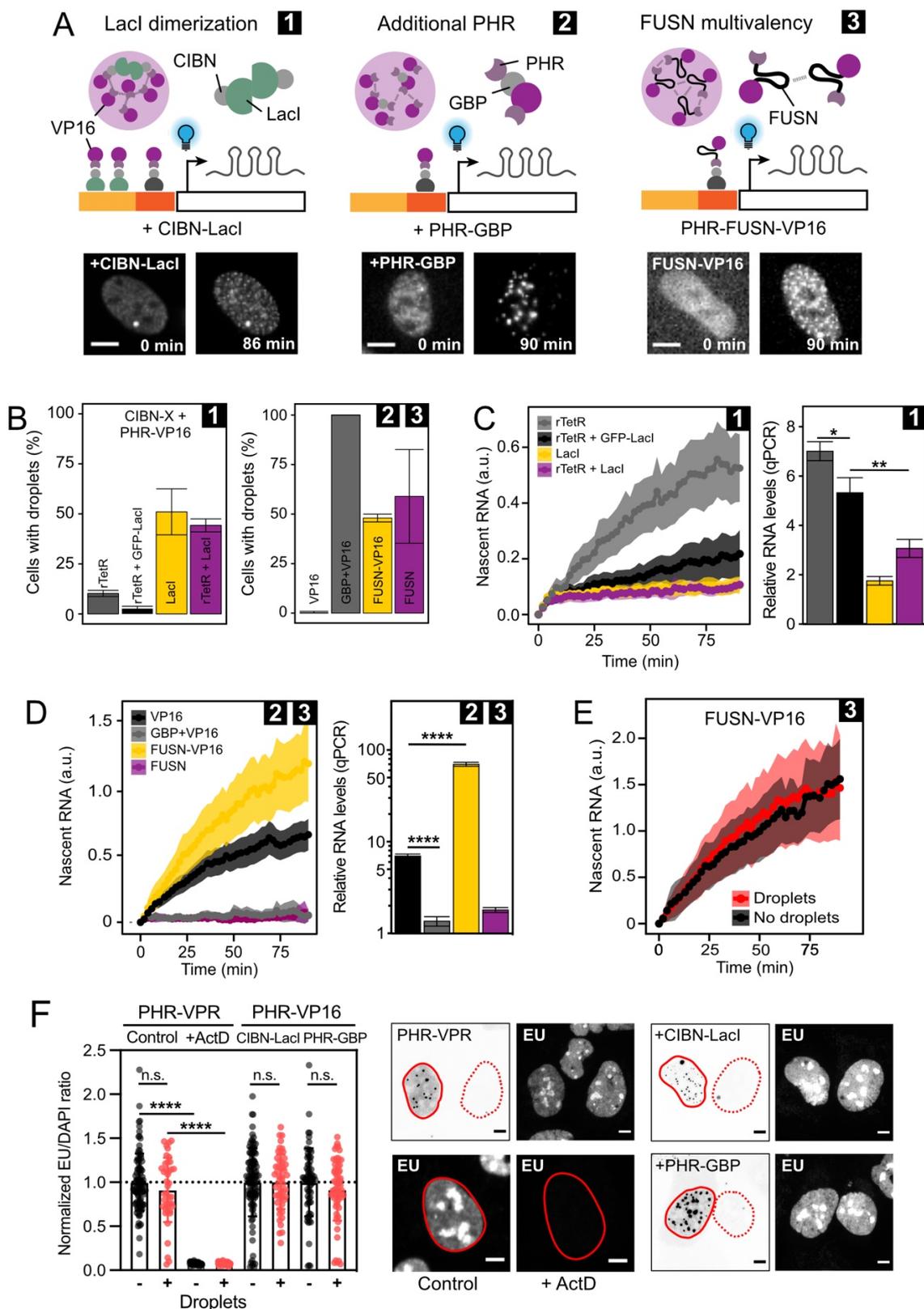


Figure 12. Composition of phase-separated compartments and transcriptional activity. (A) Three different approaches to increase the droplet formation propensity of VP16 and alter droplet composition. (1) Unbound CIBN-Lacl dimers bridge PHR molecules. (2) GFP-binding protein (GBP) coupled to PHR attaches a second PHR domain to PHR-GFP-VP16. (3) Fusion of FUSN to VP16 increases its self-interaction. *Legend continued on next page.*

Figure 12. (legend continued) (B) Fraction of cells with droplets formed by the different approaches and controls (90 min illumination). Left: GFP-LacI is the only construct without CIBN. Right: All constructs are fused to PHR and were recruited via rTetR-opto. PHR-FUSN contains no activation domain. Mean, max. and min. (error bars) of two replicates. (C) Nascent RNA time courses ($n = 74$ -126 cells per condition) and endpoint measurements of total reporter RNA (90 min) for the LacI approach. (D) Nascent RNA time courses ($n = 24$ -154 cells per condition) and endpoint measurements of total reporter RNA (90 min) for the GBP and FUSN approaches. (E) Nascent RNA time courses ($n = 47$ cells per condition) for cells activated by FUSN-VP16 and grouped by the presence of droplets. (F) Representative images and quantification of ethynyl-uridine (EU) labeled global nascent RNA 90 min after activation with the indicated constructs and grouped by the presence (solid nucleus outline) or absence (dashed nucleus outline) of droplets. Actinomycin D (ActD) treatment served as positive control for full inhibition of transcription. Bars depict the mean and dots correspond the single values of $n = 41$ -103 cells per condition. Scale bars, 5 μm . Two-sided unpaired Student's t test: $P > 0.05$, not significant (ns) and **** $P < 0.0001$. All qRT-PCR data was normalized to β -actin mRNA and to mock-transfected samples. Bar plots depict mean and s.d. of $n = 3$ replicates. Two-sided unpaired Student's t test, with not significant (ns), * $P < 0.05$, ** $P < 0.01$, and **** $P < 0.0001$. Nascent RNA trajectories were acquired and analyzed by Jorge Trojanowski. Figure panels adapted from ref. [146].

These controls corroborated the hypothesis that these two types of VP16 droplets could in fact inhibit transcription at the reporter gene cluster. In contrast to the LacI and the GBP approach, droplet reinforcement by FUSN was found to greatly increase transcription activation (**Figure 12 D**). In comparison to VP16 alone, the total reporter RNA levels of FUSN-VP16 were ten times higher after 90 minutes of illumination. Throughout the whole activation period, this was also reflected in increased levels of nascent RNA. As previously done for the other ADs (compare **Figure 11**), nascent RNA trajectories were once more classified into cells with and without discernible droplets to determine if the elevated activity of FUSN-VP16 was connected to droplet formation (**Figure 12 E**). There were no differences in the quantities of RNA between the two groups, in line with the earlier findings that droplets had a neutral effect on transcription. The results of the droplet perturbances with PHR-VP16 lead to the following conclusions: The additional enrichment of TFs at the target site by phase separation is not generally beneficial for transcription. The specific composition of phase-separated compartments strongly impacts their activity. While not explored in detail here, the effect of composition could be linked to differential internal mobility or association with co-activators. Ultimately, it still needs to be determined which properties precisely mediate strong TF activity. The FUSN experiment demonstrated unequivocally that boosting multivalent contacts can change a weak activator like VP16 into a strong activator without the need of a phase transition. Additional data addressing the role of multivalent interactions for TF activity is provided in ref. [146] and further discussed at the end of this thesis. In conclusion, strategies were presented here on how one can modulate droplet formation using the optogenetic toolbox components and multivalent linkers. This type of functional assays can be applied to study other chromatin effectors and will yield further insights on the role of compartmentalization by phase separation for regulating chromatin-templated processes.

2. ALT-FISH – a microscopic assay for the detection of ALT-specific compartments

As demonstrated in the preceding chapter, the relationship between the assembly mechanism, composition and function of chromatin-associated compartments can be probed with tunable effector constructs in different model systems. Another aspect of functional nuclear compartmentalization is exploiting it as a marker for the activity of specific biological pathways that are associated with spatially confining the corresponding biological processes. In particular, disease-associated processes produce large quantities of detrimental by-products or promote cell survival through sustained regulatory activity at specific genomic target loci, which can lead to the formation of distinct nuclear compartments. These considerations were followed up in the present thesis with respect to telomeres. They assemble into a specific nucleo-protein complex, the integrity of which is essential for controlling cell proliferation and genomic stability. Structure and activity of telomeres are drastically altered in cancer cells that use the alternative lengthening of telomeres (ALT) pathway. In these cells, recombinogenic activities are funneled towards telomeric repeat sequences to prevent their net loss. Although it is crucial to understand the ALT mechanism in order to improve cancer treatment, the absence of scalable assays has made it difficult to measure the activity of this telomere lengthening pathway in intact cells.

In the following chapters, I presented a novel assay that exploits compartment detection for the universal quantification of ALT activity across a broad range of sample types. I integrated multiple documented ALT characteristics to design the microscopic ALT-FISH assay that labels compartmentalized single-stranded telomeric nucleic acids. These indicators of ALT activity allowed for the accurate distinction of individual ALT-positive and ALT-negative cells and the resolution of ALT heterogeneity in cancer cell populations. Along with the assay, automated image analysis workflows were developed to enable unbiased quantification and comprehensive biological characterization of ALT-FISH signals.

Development and validation of the ALT-FISH assay has resulted in a separate first author publication, see ref. [165]. Selected data from this study are presented and discussed in the following three chapters of the thesis. The contribution of study co-authors to the data collection or analysis is stated where applicable.

2.1 Single-stranded telomeric repeats of various origin are a unifying feature of ALT

While it remains largely unknown what series of events ultimately leads to productive telomere extension in ALT, various characteristic telomere traits have been documented for this pathway (**Figure 13 A**, scheme). They include (i) heterogeneous telomere length and recombination intermediates [102, 112, 113], (ii) excessively long single-stranded (ss) telomeric DNA overhangs [113] and damage-induced internal telomere loops (i-loops) [118] (iii) elevated replication stress burden at telomeres [111, 166], (iv) accumulation of extrachromosomal telomeric DNA repeats (ECTRs) [115, 116] and, (v) increased levels of TERRA [130, 132-134].

It is technically infeasible to quantify all these traits for a single cancer cell with their respective specialized assays. However, a strategy that incorporates shared molecular aspects of these attributes would provide an integrated measure of ALT activity. I reasoned that most of the aforementioned ALT traits share one common theme: the exposure of C- or G-rich single-stranded telomeric repeats (SSTRs) in the form of either RNA or DNA (**Figure 13 A**, yellow marks). SSTRs are expected to be rare in telomerase-positive cancer cells and somatic cells, where their only sources are the (considerably shorter) overhangs of telomeres and the low quantities of G-rich TERRA transcripts. On the basis of this assumption, I concluded that SSTR detection might be a method to specifically capture a range of ALT traits at once.

2.2 ALT-FISH visualizes compartments of single-stranded telomeric nucleic acids

A fluorescence microscopy-based approach was devised to detect SSTRs at the highest possible sensitivity while maintaining spatial information for assignment of ALT activity to single cells. The detection method was named ALT-FISH and followed the principle of non-denaturing fluorescence *in situ* hybridization (FISH) with two fluorescently-labeled DNA probes that target either the C-rich strand (TelG probe) or the G-rich strand (TelC probe) (**Figure 13 B**). The information provided by the two probes is complementary. While the TelG probe maps C-rich single-stranded regions, such as those seen in C-circle ECTRs, the TelC probe specifically targets TERRA and G-rich telomeric ssDNA repeats. In addition, TERRA and telomeric DNA signals can be distinguished using an RNase treatment. To ensure that the assay is quantitative and scalable, the ALT-FISH staining procedure was designed in such a way that it only needs a single isothermal hybridization step, no signal amplification, and a minimum of liquid handling steps with standard buffers. The probe concentration was titrated for highest signal-to-noise ratio and reproducibility.

Confocal imaging of the ALT-positive osteosarcoma cell line U2OS stained for either C-rich or G-rich SSTRs using ALT-FISH revealed intense sub-micrometer sized nuclear and cytoplasmic SSTR foci (**Figure 13 C**). In contrast, and despite similar nuclear background

levels, these intense accumulations of SSTRs were not seen in ALT-negative (telomerase-positive) HeLa cells (**Figure 13 D**). This result raised the possibility that higher levels and focal accumulation of SSTRs into compartments is an ALT-specific trait that ALT-FISH can visualize.

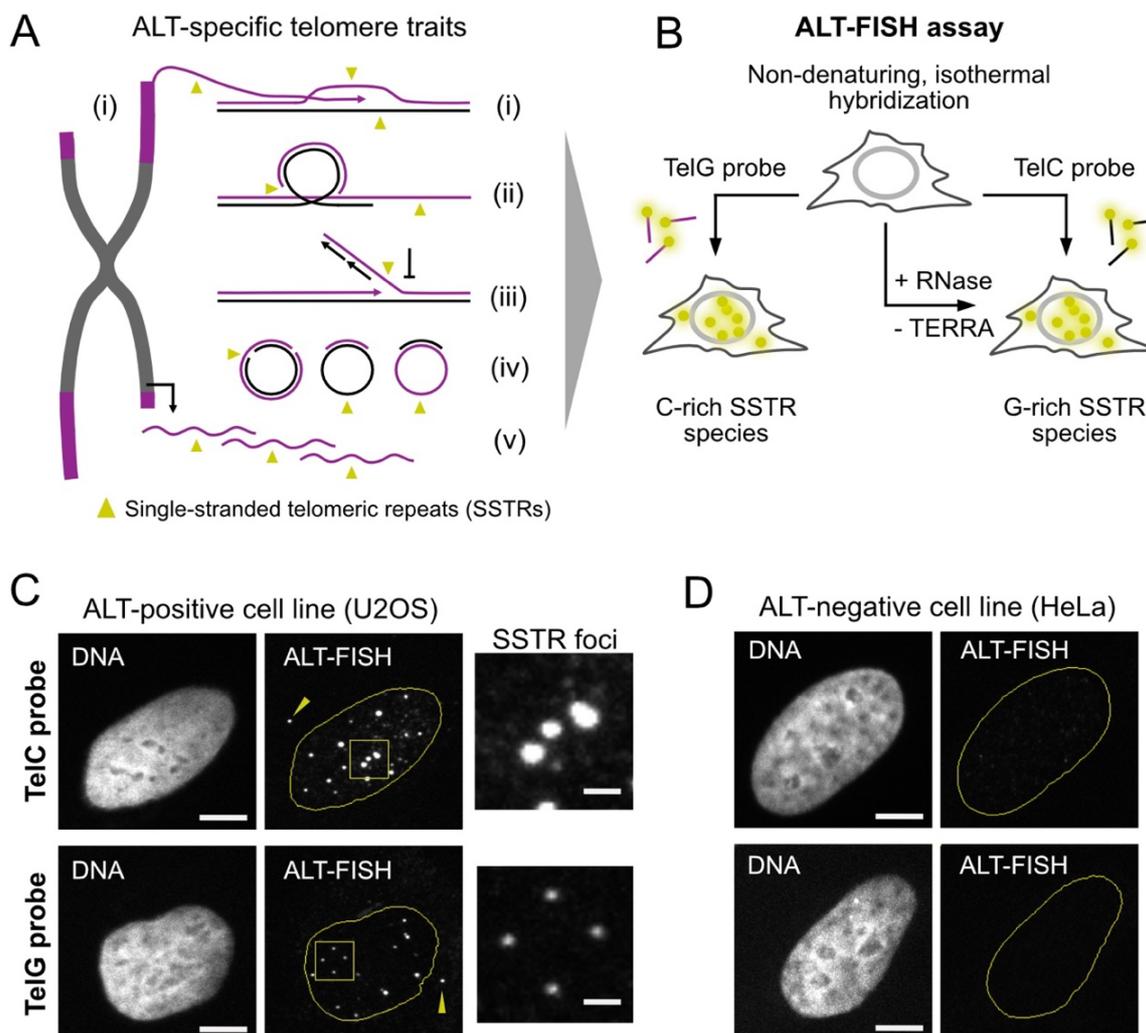


Figure 13. The ALT-FISH assay detects compartments containing single-stranded telomere repeats (SSTRs). (A) ALT-specific telomere traits that potentially expose SSTRs: (i) heterogenous telomeres and recombination intermediates, (ii) damage-induced internal loops and long telomeric overhangs, (iii) replication stress intermediates, i.e. stalled replication forks, (iv) extrachromosomal telomeric DNA (C-circles, G-circles, t-circles, etc.), (v) telomere repeat containing RNA TERRA. (B) Principle of the ALT-FISH assay to detect G-rich or C-rich SSTRs. TERRA can be probed by an RNase pre-treatment. (C) ALT-FISH staining reveals mostly nuclear and some cytoplasmic (arrows) SSTR foci in ALT-positive U2OS cells. (D) ALT-negative (telomerase-positive) HeLa cells show no apparent SSTR foci. Yellow outlines represent nuclei segmented on the DAPI channel (DNA). Scale bars, 5 μm and 1 μm (zoom-in). Some images of panels C and D were adapted from ref. [165].

2.3 The *Telosegment* toolkit enables automation of ALT-FISH data analysis

Unbiased and automated feature extraction from microscopy images is key to make microscopy-based assays like ALT-FISH quantitative and scalable. I thus developed an analysis software toolkit called *Telosegment* for the automated quantification of ALT-FISH image data of hundreds of cells (**Figure 14**). The output of the workflow was to mainly provide ALT-FISH signal features (spot number, location, intensity, etc.) for single cells, with the option to integrate up to three additional image channels for other readouts. The toolkit comprised fully automated 2D cell segmentation, spot detection and intensity feature extraction. In the standard setup, a nucleus marker channel (DAPI) and the ALT-FISH channel are needed for analysis. Co-localization analysis of ALT-FISH signals with up to two more fluorescence channels is possible. With a cytoplasm or cytoskeleton marker channel, *Telosegment* also offers a separate cytoplasm segmentation module. *Telosegment* performs annotation and filtering of data based on quality criteria (e.g. nucleus size and shape). These metrics can later be assessed to refine downstream data analysis. The *Telosegment* pipeline was implemented in the R programming language [160] and builds on functions of the EBImage R package [167]. It is computationally inexpensive to run on a local machine. With the data format used here, it can typically process 800 cells per hour (about 60 image stacks, ~25 GB in total) on a machine with 16GB RAM and 4 CPU cores.

The major analysis steps and functions of *Telosegment* are outlined in the following. Specific steps and functions are detailed in the *Materials and Methods* section. The analysis starts with importing the raw confocal image z-stacks (positions), where one position typically corresponds to a single tile of larger tile scan (**Figure 14 A**). Each image position may have several cells and up to four image channels in the setup utilized here. First, the function *findBestSlices* evaluates the stacks, identifies positions with no cells or poor focus and flags them for further processing (**Figure 14 B**). It also estimates the optimal z-range for subsequent projection. Second, the image stacks are reduced to 2D maximum intensity projections and objects are segmented on the respective channels (**Figure 14 C**). Cell nuclei are segmented on the DAPI image using the *makeNucMask* function. If available, the cytoplasm area is segmented based on the cytoplasm marker. Partial area masks located at the image borders or cytoplasm masks without a matching nucleus are identified and excluded. Spots on the ALT-FISH images are segmented using the *makeSpotMask* function on the corresponding mask areas. A local per-cell intensity threshold is used for segmentation. For quantitative comparisons, the spot segmentation stringency is initially calibrated by the user and then maintained constant across all experimental conditions. Once the nucleus and spot masks have been computed, position shape and pixel intensity features are quantified in all masks and across all available channels (**Figure 14 D**).

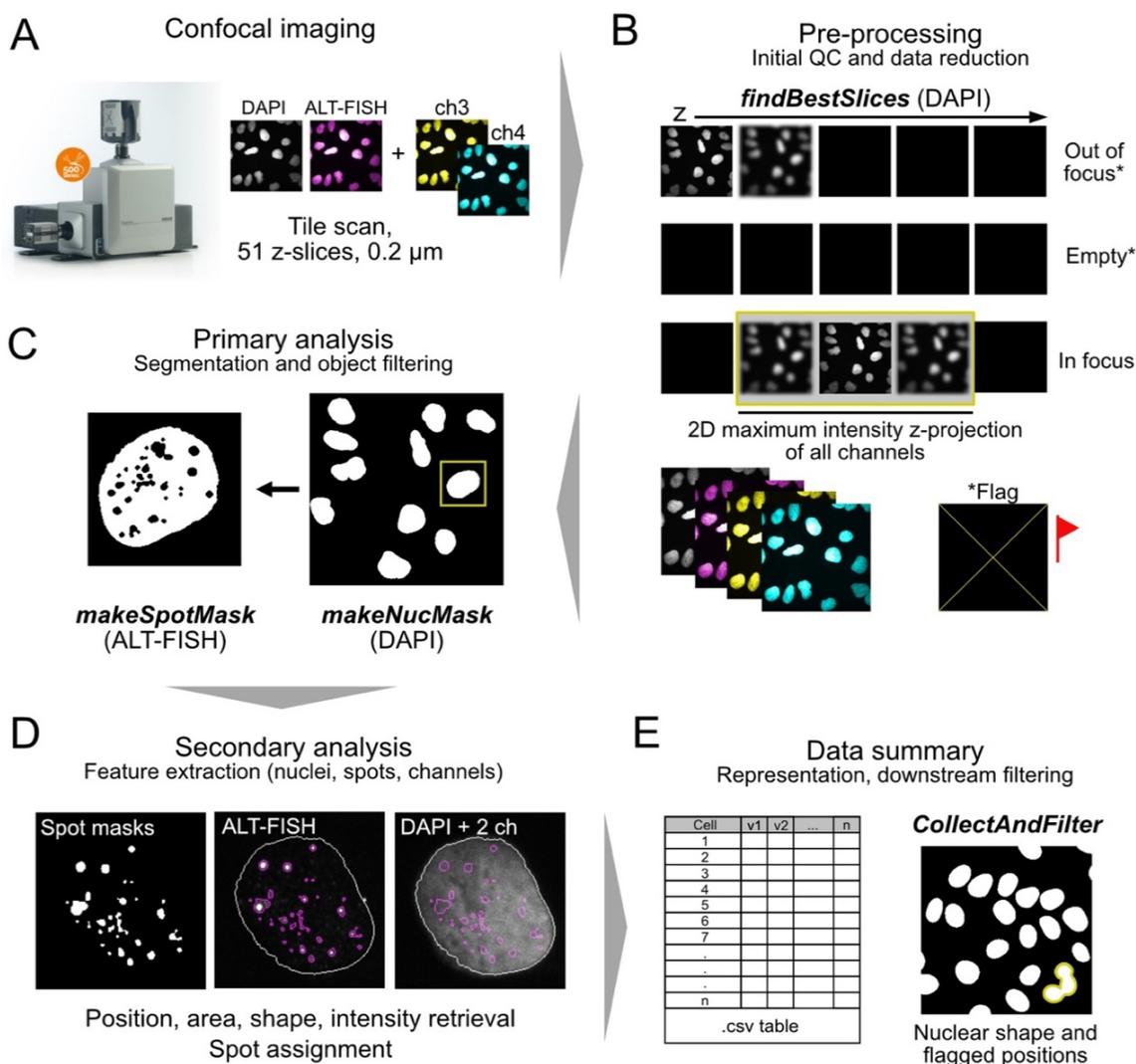


Figure 14. Automated ALT-FISH image analysis workflow using the *Telosegment* toolkit. (A) Confocal z-stacks of DAPI, ALT-FISH and optional channels are acquired, with each image position containing multiple cells. (B) The *findBestSlices* function evaluates DAPI intensity distribution and contrast features for each position, flagging out of focus and empty positions and generating range-optimized 2D maximum intensity z-projections for the in-focus positions. (C) The *makeNucMask* and *makeSpotMask* functions sequentially segment nuclei and spots in the whole image (DAPI channel) and single nuclei regions (ALT-FISH channel), respectively. Small and incomplete border nuclei are filtered out. A cytoplasm marker channel can be included for spot quantification in the cytoplasm. (D) Features, including xy-positions, areas, shape metrics, and intensity values across available channels, are retrieved from the mask objects. (E) The final quantification and segmentation results are presented in a universal tabular format and a set of overlay images for inspection (compare D). The *CollectAndFilter* script comprises customizable routines for removing flagged positions and excluding low quality data points based on shape criteria of the nuclei or other user-specified criteria.

In a final step, the quantification results are summarized into tabular file format compatible for downstream analysis (**Figure 14 E**). A supplementary R script *CollectAndFilter* is used for downstream data filtering, such as removal of flagged image positions, nuclear segmentation artifacts or aberrantly shaped nuclei. Detection of artifacts is done via a custom nuclear shape

quality score (NSQS), which is described in the *Materials and Methods* section. In addition to the final results, *Telosegment* exports a set of overlay images for position-wise visual inspection or optional re-import of image masks. In conclusion, the automated image analysis workflow developed here can be applied to quantitatively compare subcellular ALT-FISH signals and other features between cells and between experimental conditions.

2.4 Telomeric RNA TERRA and ECTRs contribute to the ALT-FISH signal

In the previous sections, I presented initial evidence that SSTRs accumulated in ALT-positive U2OS cells, as detected by ALT-FISH and proposed that they originated from distinct sources. To corroborate this proposition, the contribution of TERRA and extrachromosomal SSTR species to the ALT-FISH signal was tested in U2OS cells (**Figure 15**). The G-rich RNA TERRA can be degraded with RNase. And while it is technically not trivial to differentiate DNA SSTRs in telomeric overhangs, internal loops and ECTRs in the nucleus, cytoplasmic ALT-FISH spots are attributable to extrachromosomal SSTR species. The median number of nuclear spots for untreated U2OS cells was 17 (TelC probe) and 6 (TelG probe), with a range of 0 to >60 spots and 0 to 25, respectively (**Figure 15 A, B**). RNase pre-treatment abolished approximately half of the TelC foci but virtually none of the TelG foci (**Figure 15 B**). Similar losses were observed for RNase A or a combination of RNase A+H, while RNase H pre-treatment alone preserved slightly more spots. These results established that TERRA significantly contributed to the G-rich SSTR pool detected by the ALT-FISH assay. The findings were furthermore consistent with earlier studies that have detected TERRA in the form of ssRNA and RNA:DNA hybrids [168] and studies that have observed a lack of C-rich telomeric RNA in human cells [130].

To confirm that ALT-FISH also detected single-stranded regions in ECTRs, I next quantified the amount of cytoplasmic signal for both probes in U2OS cells. Fluorescent phalloidin labeling of the actin cytoskeleton was used to segment the cytoplasm area (**Figure 15 C**). ALT-FISH spots were substantially less frequent in the cytoplasm than in the nucleus. The median spot number was 3 (TelC) and 1 (TelG), ranging from 0 to 22 and 0 to 16 (TelG) (**Figure 15 D**). Signals residing in cytoplasmic spots made up less than 10% of the cumulative spot intensity per cell (**Figure 15 E**). This comparably low abundance was in accordance with previous reports on the abundance of cytoplasmic ECTRs studied by denaturing FISH [117, 169]. The no probe control staining showed median spot numbers of 0 (nucleus and cytoplasm) and ranging from 0 to 1 (nucleus) and 0 to 2 (cytoplasm), demonstrating that autofluorescence had little impact on overall spot detection. In conclusion, this second set of experiments showed the ability of ALT-FISH to detect ECTRs. However, to identify whether the ECTR species detected by ALT-FISH predominantly relate to circular (e.g. C-circles, G-circles) or other types of intermediates will need further investigation.

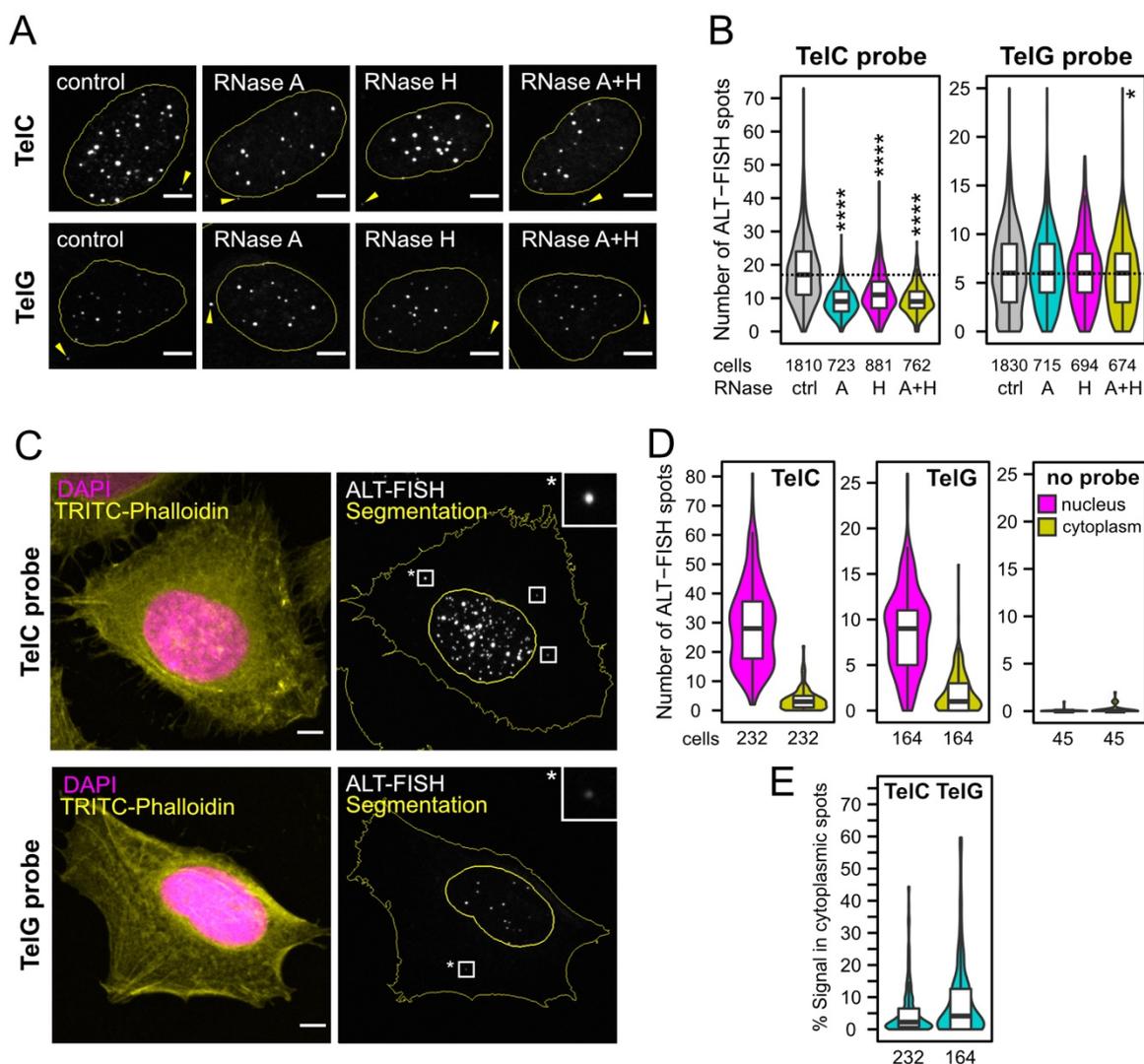


Figure 15. ALT-FISH detects several sources of SSTRs. (A) Images of ALT-FISH stained ALT-positive U2OS cells without treatment (control) or pre-treated with RNase A, H or both. Outlines show nucleus segmentation and arrows non-nuclear foci. (B) Corresponding quantification of nuclear spot numbers for both probes. Dotted lines mark the median of the control. Data was pooled from 2-4 replicates. Adjusted P values from a pair-wise Wilcoxon rank sum test with BH correction compared to control groups were >0.05 , $*0.024$ or $****< 2e-16$. (C) Images of U2OS cells co-stained by DAPI, ALT-FISH and TRITC-Phalloidin to stain the cytoplasm. Insets show exemplary cytoplasmic spots. (D) Quantification of nuclear and cytoplasmic spot numbers for C, including a no probe control. Images of the no probe control are shown in ref. [165]. (E) Per cell cumulative ALT-FISH signal in cytoplasmic spots as fraction of total spot signal per cell. All boxplots depict median, first and third quartiles of the data with the number of analyzed cells indicated below each condition. Scale bars, 5 μm . Figure adapted from ref. [165].

2.5 SSTR compartments contain telomere repeats in the kilobase range

Somatic cells and telomerase-positive cancer cells have telomere lengths between 3 and 12 kb, but telomeres of ALT-positive cells can reach 50 kb in some cases [102]. With appropriate standards, it is possible to determine the absolute telomere length and repeat copy number of single cells using denaturing telomere FISH [170]. Repeat copy number is essential to derive quantitative mechanistic models, e.g., on the number of telomere-repeat binding proteins or size of a single telomere. In the present section, the proof-of-concept of a technique was developed to estimate SSTR copy number in individual ALT-FISH foci of intact U2OS cells that was based on a combination of fluorescence correlation spectroscopy (FCS) and confocal microscopy (**Figure 16**).

Using FCS, a fluorescent target molecule's brightness can be determined by the photon count rate per molecule when that molecule passes through a small sample excitation volume [171]. Here, this technique was exploited to measure the brightness of single TelC ALT-FISH probe molecules (labeled with Atto633 dye) in solution and relate it to the brightness of single 0.1 μm poly-fluorescent beads with a comparable fluorophore (**Figure 16 A**, left). FCS data acquisition and analysis were performed by Norbert Mücke, a research engineer in our group. The FCS results estimated that a single fluorescent bead was about 114 times brighter than a single TelC probe molecule.

To derive SSTR copy numbers using this information, I prepared a sample of TelC ALT-FISH stained U2OS cells with fluorescent beads added after the staining (**Figure 16 A**, right). The sample was imaged using the previous confocal setup. Because the beads carried an additional Alexa488 dye, they could be distinguished from ALT-FISH spots using a separate channel (**Figure 16 B**). In the microscopy images, spot and bead intensities were quantified in the Atto633 channel. Knowing the exact probe length (30 nt) and the brightness ratio from FCS, intensities could be converted into repeat copy number for 200 randomly selected ALT-FISH spots from six U2OS cells. The number of exposed TTAGGG repeats per spot ranged from 50 (0.3 kb) to 6,800 (41 kb) (**Figure 16 C**). This range was consistent with the size ranges of 0.1 to 9 kb for repeats in TERRA transcripts and the total length of 0.8 to >50 kb for ALT-specific telomeric circles that have been reported [116, 130].

Despite the limited dataset examined here and the fact that the quantification presupposes full hybridization of the probes to their targets, the feasibility of the approach was demonstrated. I found that hundreds to thousands of bases of exposed SSTRs were accumulated in a prototypical ALT-FISH compartment of U2OS cells. Further experimental advancements of this method will derive how the single-stranded stretches are distributed among individual molecules of the various chromosomal and extrachromosomal SSTR species.

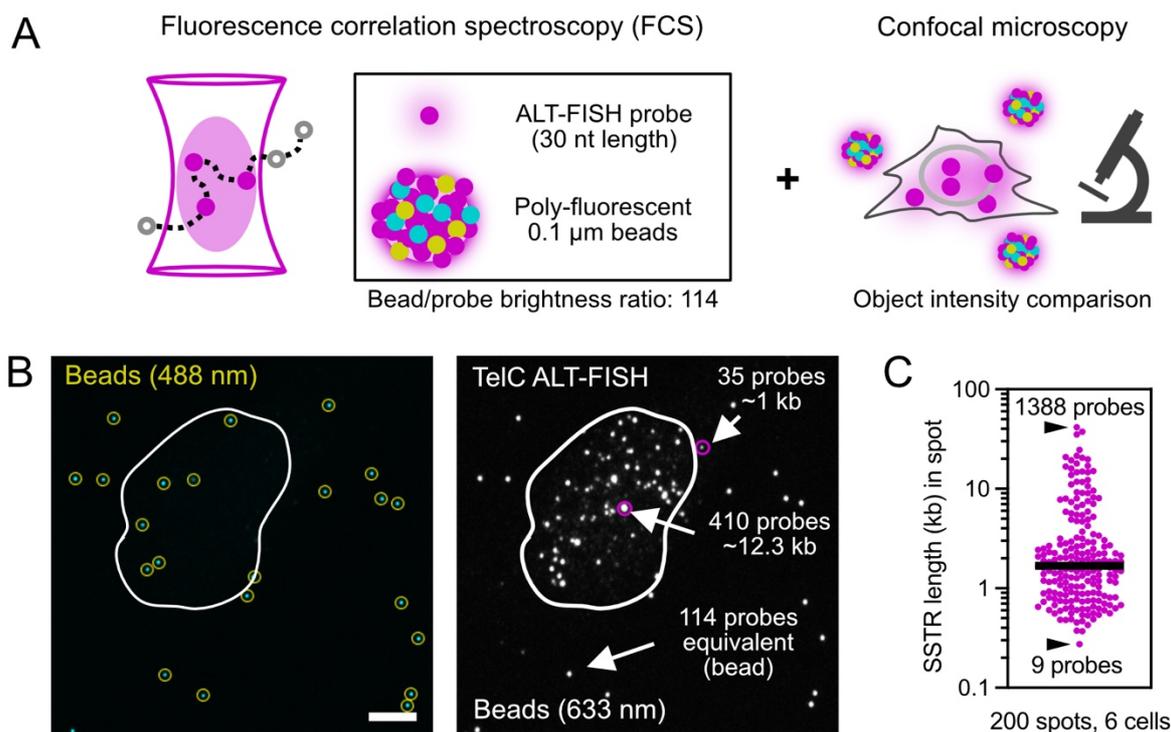


Figure 16. Estimation of SSTR copy number in ALT-FISH compartments. (A) Overview of using FCS to determine the brightness ratio of a single ALT-FISH probe molecule to a poly-fluorescent bead, which can be used to estimate the number of probe molecules in single foci by comparing their absolute intensities using confocal microscopy. (B) Representative TelC ALT-FISH stained U2OS cell with beads added to the mounting medium. Beads are detectable in both the 488 nm and ALT-FISH probe (633 nm) channel. Exemplary spots and selected probe number/SSTR length estimates are highlighted. (C) SSTR length distribution of 200 randomly selected SSTR foci across 6 U2OS cells. Maximum and minimum probe equivalents (arrows) and median (bar) are indicated. FCS data to determine the bead/probe ratio was acquired and analyzed by Norbert Mücke. Figure adapted from ref. [165].

2.6 ALT-FISH foci enrich components of ALT-associated PML bodies (APBs)

In addition to the SSTR-containing ALT compartments identified here, previous studies had already reported ALT-specific nuclear compartments formed by the co-localization of PML proteins with telomeres, termed ALT-associated PML bodies (APBs) [121]. APBs are an established ALT marker and accumulate large amounts of telomeric DNA. Alongside chromosomal telomeric DNA ends and TERRA molecules [172], some of the telomeric DNA enriched in APBs is single-stranded [173] and possibly belongs to both circular and linear ECTRs species [174]. In light of these previous findings, it was important to further characterize the nuclear compartments that the ALT-FISH assay detects and determine their overlap with APBs. To this end, I carried out a combination of ALT-FISH and immunofluorescence (IF) staining for four different candidate proteins (PML, SUMO2/3, TRF2 and RPA) that had been previously linked to APBs, in order to map their localization relative to ALT-FISH foci (**Figure 17**).

PML is the main structural component of APBs and PML nuclear bodies (PML-NB). The shelterin protein TRF2 binds to double-stranded telomere repeats [88]. The single-stranded DNA binding protein RPA localizes to APBs [173] and is involved in DNA repair. The SUMO2/3 variant is found conjugated to proteins at the interior and in the shell of PML-NBs and APBs [123, 175, 176]. To examine the distribution of these proteins in ALT-positive U2OS cells, nuclei and nuclear ALT-FISH spots were segmented and intensities were quantified in each spot area across all staining channels (DAPI, ALT-FISH, IF candidate 1, IF candidate 2). Always two candidate proteins were stained in parallel (RPA with SUMO2/3 and TRF2 with PML) and together with TelC or TelG ALT-FISH (**Figure 17 A, C**). In each SSTR spot, the fold-enrichment of an IF signal over the nuclear background was used to classify spots as co-localizing with a given signal or not. The DAPI channel was used as a background measure. Using this strategy, enrichment values of 2,700-4,500 individual SSTR spots from 365-450 U2OS cells were quantitated for each condition (**Figure 17 B, D**) and used to call co-localization events (**Figure 17 E, F**). For SUMO2/3 and RPA, respectively, co-localization was observed in 60.8% and 64.7% of TelC spots and 69% and 90.3% of TelG spots. An overall higher proportion of spots was enriched for TRF2 and PML. 91.1% of TelC and 98.6% of TelG spots displayed significant PML enrichment, with maximum enrichment values reaching up to 50-fold (compare **Figure 17 D**). While less than <1% of TelG spots were not associated with PML, this fraction was higher for the TelC spots (8.9%). TRF2 was found in 70.4% of TelC and 76.6 % of TelG spots. This pairwise co-localization analysis revealed that candidate proteins were enriched to varying extent and nearly all SSTR foci coincided with PML protein. Another evidence that ALT-FISH detects SSTRs came from the clear association seen with RPA signals.

Next, I specifically looked into how PML/ALT-FISH co-localization events were distributed between APBs and non-APB PML-NBs. I first identified APBs by co-enrichment for TRF2 and PML signals. APBs co-localized with 65.3% of TelC spots and 76.1% of TelG spots (**Figure 17 E, F**). Independently, TRF2 foci were used to segment telomeres on the same images and signal enrichment was determined for the other channels. Segmented telomeres were binned into groups with and without significant PML enrichment, and their ALT-FISH signal was compared (**Figure 17 G and H**, violin plots). Telomeres with PML contained significantly more SSTRs. However, the majority (63.8-68.8%) of telomeres lacked PML and ALT-FISH (**Figure 17 G and H**, stacked bar graphs), which is consistent with only some telomeres engaging in APBs. Overall, the findings uncovered that APBs and two thirds of SSTR foci overlapped in U2OS cells. It can be concluded that the majority of APBs can be identified by ALT-FISH, without the need for PML/telomere co-staining. 23-26% of spots matched with PML-NBs that were not present in APBs, while 0.9-8.9% of spots had no spatial link with PML. Thus, a portion of SSTR foci are distinct from APBs.

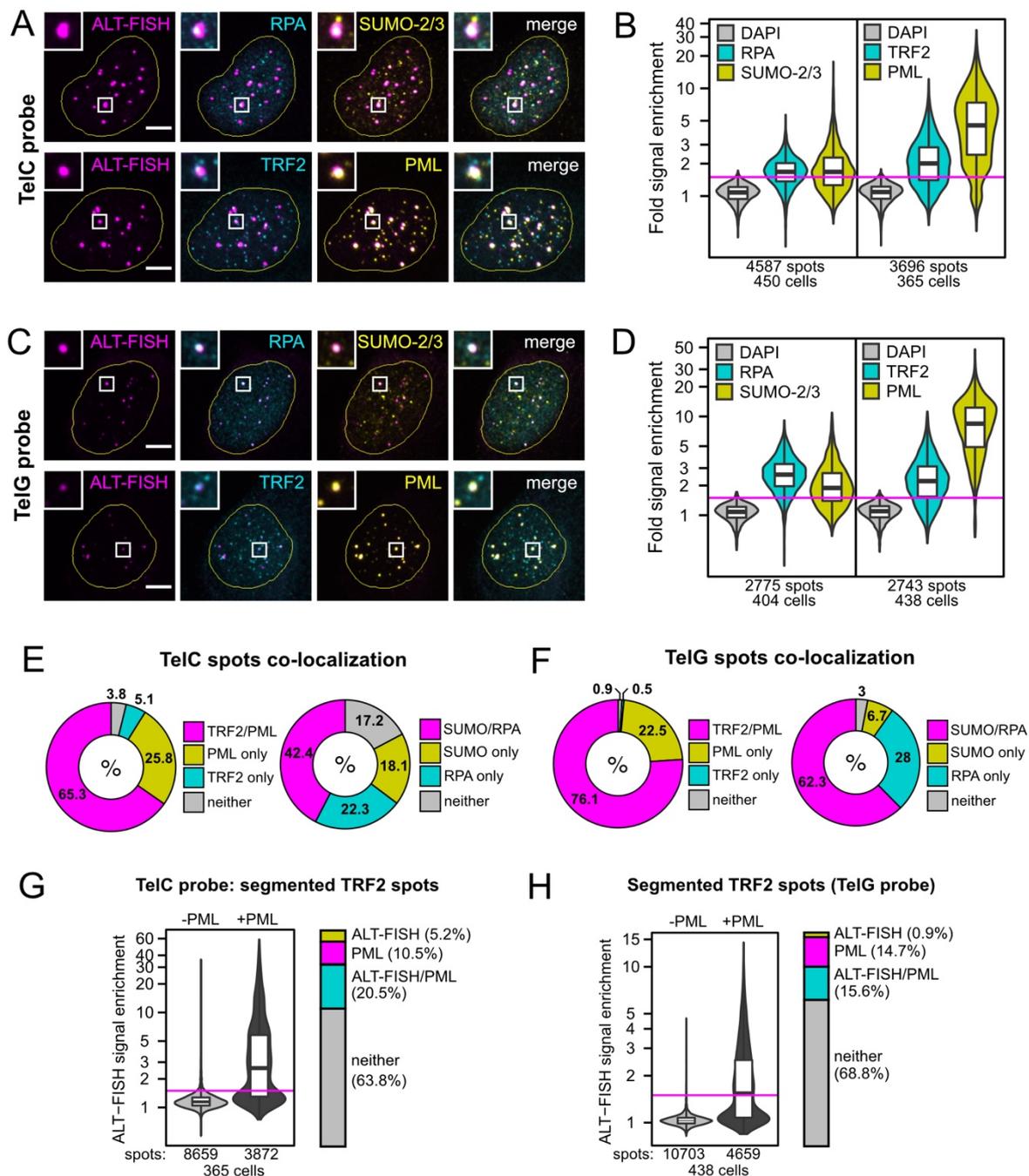


Figure 17. ALT-FISH signals co-localize with components of ALT-associated PML bodies. (A) TelC ALT-FISH and immuno-staining of APB components and TRF2 in U2OS cells. RPA was co-stained with SUMO2/3 and TRF2 together with PML. (B) Quantification depicting the fold enrichment of each signal in the spot area relative to the nucleoplasm (magenta line: co-localization threshold). (C-D) Same as A and B, but for TelG. (E-F) Binned data from A and B depicting the fraction of spots co-localizing with either of the IF signals, both or none. (G-H) Segmented TRF2 spots analyzed with respect to their ALT-FISH signal enrichment in dependence of PML co-localization. The fractions of TRF2 spots co-localizing with the other signals are additionally shown as barplot. All boxplots depict median, first and third quartiles of the data with the number of analyzed spots/cells indicated below each plot. Scale bars, 5 μ m. Figure adapted from ref. [165].

2.7 ALT-FISH stratifies ALT-positive and ALT-negative cells with high accuracy

In the previous section, I found a spatial link between some of the SSTR foci detected by ALT-FISH and APB components, relating them to ALT activity at telomeres. As of present, the most practical choice to identify ALT-positive cells *in situ* is to stain for APBs. However, the antibody-based detection, the need of co-localization analysis for their identification and their limited dynamic range [177] make them unattractive as quantitative ALT activity readout. In part due to these shortcomings, most studies only report the use of APBs as a binary classifier (ALT, non-ALT) at the level of bulk samples and do not use it for mapping ALT tumor heterogeneity with single cell resolution [104, 124, 178]. In contrast, the simpler principle and single fluorescent channel used in ALT-FISH would potentially qualify for such applications.

I therefore sought to determine whether the ALT-FISH assay was sensitive and specific enough to separate ALT-positive from ALT-negative cells in heterogeneous cell populations. I designed an experiment, in which ALT-positive U2OS cells stably expressing nucleolin (NCL)-RFP were co-cultured with ALT-negative HeLa cells stably expressing histone H2A-YFP and subsequently stained by ALT-FISH (**Figure 18**). The different fluorescent proteins enabled their distinction using an independent image channel (**Figure 18 A**). Utilizing these markers, U2OS and HeLa cells were identified from the mixed population (**Figure 18 B, E**) and ALT-FISH signals could be determined for each group of cells. While TelC and TelG spots were readily discernible in U2OS cells, they were mostly absent in HeLa. For HeLa cells stained with either the TelC or TelG probe, the median spot number was 0. U2OS cells showed a median of 12 (TelC) and 7.5 (TelG) spots (**Figure 18 C, F**).

An ALT-FISH spot count cutoff was identified that yielded clear separation. A cutoff of >2 spots per nucleus achieved a very high (91-98%) detection rate of the ALT-positive cells, while retaining a low false positive rate (FPR) of 0.8-3.4% (ALT-negative HeLa cells detected as ALT-positive) (compare **Figure 18 C, F**). To confirm robustness of separation, I also verified these results using the coefficient of variation (CV) as an alternative image feature derived from the ALT-FISH channel. The CV reflects image contrast by the standard deviation to mean ratio of the nuclear pixel intensity and is thus uncoupled from intensity threshold based spot detection. Also using CV value cutoffs of >0.2 (TelC) and >0.3 (TelG), the detection accuracy for ALT-positive U2OS cells remained high (91-100%, FPR 1.6-4.1%) (**Figure 18 D, G**). Clear separation of U2OS and HeLa cells (99.4-100%, FPR 0-1.3%) using similar CV/spot count cutoffs (>2 spots or >0.3 CV value) was also observed when an independent co-culture was stained in suspension (data shown in ref. [165]). It can be concluded that both the relative and absolute ALT-FISH signal differences between U2OS and HeLa are robust towards changes in the sample format. Cell mixing demonstrated that ALT-FISH is a precise technique for separating individual cells depending on their ALT activity.

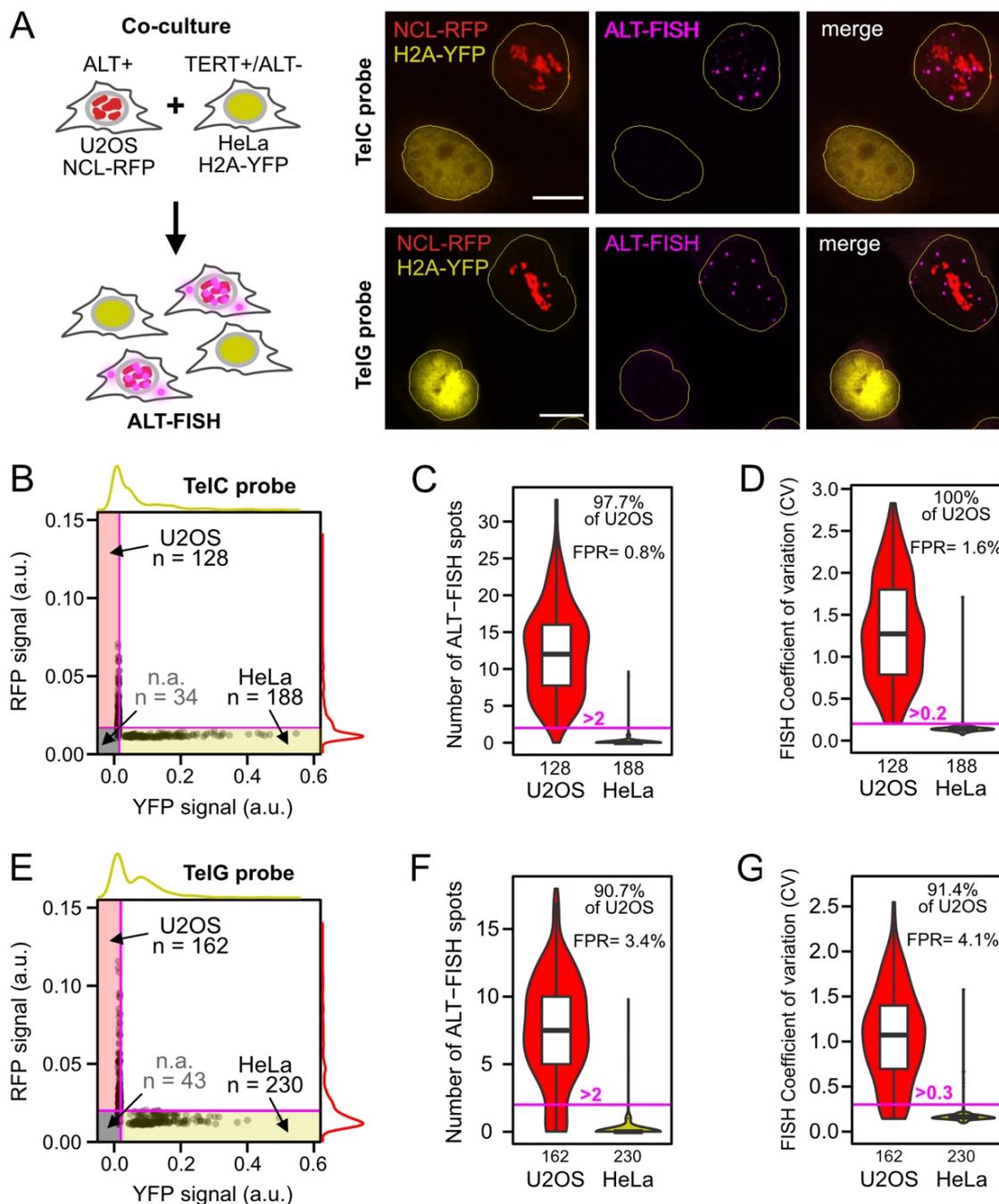


Figure 18. ALT-FISH enables accurate distinction of ALT-positive and ALT-negative cells in heterogeneous cell mixtures. (A) Experiment with fluorescently labeled U2OS and HeLa co-cultures to assess the sensitivity and specificity of ALT-FISH. Nucleus segmentation from DAPI is shown as outlines. Scale bars, 5 μ m. (B) Scatterplot of RFP (U2OS) versus YFP (HeLa) nuclear intensity with corresponding density diagrams for n = 350 cells pooled from two independent co-culture and TelC staining experiments. (C) Distribution of nuclear TelC spots for the two cell lines separated by their markers. The fraction of correctly identified U2OS cells and FPR are indicated for a spot count cutoff of >2. (D) Same plot as C, but depicting the coefficient of variation (CV) of ALT-FISH intensity in the nucleus and separation using a cutoff of >0.2. (E-G) Same as panels B-D, but for the TelG probe staining of n = 435 cells from two replicate experiments. All boxplots depict median, first and third quartiles of the data with the number of analyzed cells indicated below each plot. Scale bars, 5 μ m. Data was jointly acquired with Anne Rademacher. Figure adapted from ref. [165].

The assay's reliance on a single channel offers a substantial improvement over APB detection, both experimentally and in terms of data interpretation. As compared to previously reported APB count differences, ALT-FISH spot count differences were significantly greater. Prior to this, our group has applied a stringent 3D co-localization approach to quantify APBs from microscopy images [177]. Using this method, the examination of ALT-negative HeLa cells had shown that around 34% of the cells still exhibited at least one co-localization event and 0.5 APBs/cell on average, while ALT-positive U2OS cells contained 4-5 APBs/cell on average [177]. This 10-fold difference contrasts to the only 0.3-3.4% of HeLa cells with two or more ALT-FISH spots in the present dataset, where U2OS to HeLa differences were almost an order of magnitude higher (40-fold for TelG probe, 124-fold for TelC probe). Thus, more sensitive ALT activity measurements at larger dynamic range may be possible with ALT-FISH.

2.8 The ALT activity profiles of cancer cell lines are predicted by ALT-FISH spot counts

A crucial requirement for application of ALT-FISH is to confirm that it reliably tracks ALT activity, regardless of the cell line model or tumor entity. I therefore conducted ALT-FISH on a diverse panel of 13 cell lines (**Figure 19**). The panel covered osteosarcoma (HOS, MG-63, Saos2, CAL72, U2OS), cervix carcinoma (HeLa wt, HeLa ST, HeLa LT) and pediatric glioblastoma (SF188, NEM165, NEM157, MGBMI) cell lines. Cell lines selection was based on pre-existing TMM classification data and predisposition to ALT [179-181]. Both osteosarcomas and pediatric glioblastomas (pedGBMs) frequently activate ALT (~44%-47 ALT-positive cases) [104, 108, 124, 182]. One HeLa clone of unknown telomere length (HeLa wt) and two isogenic HeLa clones with short (ST, 9kb) or long (LT, 20 kb) telomeres were included as telomerase-positive cervix carcinoma models. Despite the fact that ALT is rare in this entity, ALT induction had been previously reported for the HeLa LT clone when artificially depleting the histone chaperones ASF1A/B [181, 183]. HeLa ST and LT clones are indistinguishable when assayed for APBs, CC level and telomerase activity in their unperturbed state. Because of this, they were considered as potential model for the investigation of emerging ALT activity. In addition to the cancer cell lines, I furthermore included human umbilical cord endothelial cells (HUVECs) as a cell line with no active TMM. In a first step, the TMM status of all cell lines was independently validated by two complementary assays. Bulk ALT activity was measured using the C-circle (CC) assay [120] (**Figure 19 A**). Telomerase activity was qualitatively assessed by Telomerase Repeated Amplification Protocol (TRAP) [184] (**Figure 19 A**). Both confirmed the previous TMM status. Very low to undetectable CC levels were present in cell lines that exhibited telomerase activity patterns in the TRAP assay and TRAP negativity was mostly linked to increased CC levels in the cancer cell lines, but not in HUVECs.

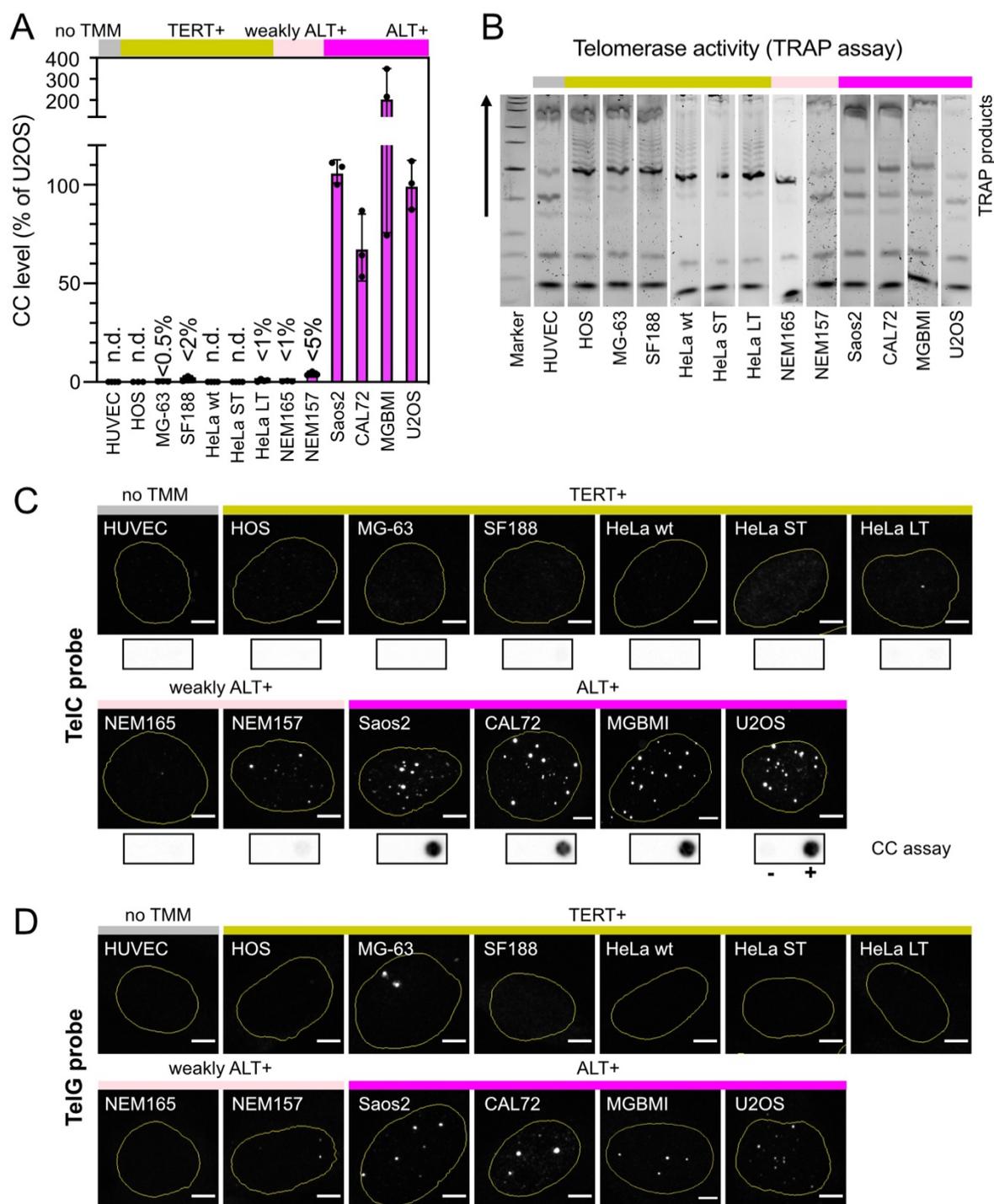


Figure 19. ALT-FISH profiles of a cell line panel in relation to established TMM markers. (A) Quantification of C-circle (CC) levels for a cancer cell line panel and HUVEC cells with previously determined TMM status (color code). CC levels are normalized to U2OS CC levels on the same blotting membrane. Bar plots depict the mean and s.d. of 3-4 independent replicates. (B) Qualitative telomerase activity analysis by telomerase repeated amplification protocol (TRAP) assay. Telomerase activity is shown by a ladder of TRAP products that appear above the control bands (arrow) in the gel. (C) Representative images of cell lines stained by TelC ALT-FISH. Representative CC assay blot results (-/+ phi29 polymerase reaction) above and below, respectively. Outlines show nucleus segmentation. (D) Same as panel C, but for TelG. Scale bars, 5 μ m. Figure adapted from ref. [165].

The CC level range displayed by CC-positive/TRAP-negative cell lines was broad. The MGBMI cell line had the highest level (~200% of U2OS) followed by U2OS, Saos2, CAL72 with similar and NEM165, NEM157 with low levels (1-5% of U2OS). It is noted that higher relative CC levels had been documented for NEM165 and NEM157 in a previous study [180]. Due to their low CC levels and lack of detectable telomerase activity, these two cell lines were categorized as weakly ALT-positive.

In a second step, ALT-FISH staining was performed using either the TelC or the TelG probe (**Figure 19 C, D**). Confocal images were recorded from three to four independently stained replicates per condition and analyzed using the *Telosegment* toolkit developed in *section 2.3*. Afterwards, the quantification results of all cells belonging to the same condition were pooled. After filtering, 477 to 1,942 high quality cells were obtained for each condition, totaling to about 26,000 cells for all samples (**Figure 20 A**). For ALT-positive cell lines, the median nuclear spot counts ranged from 7–17 (TelC) to 3-6 (TelG). For telomerase-positive and HUVEC, they were primarily between 0-1 or ranged between 0-2 for the weakly ALT+ cell lines. One notable exception was the MG-63 cell line, which regularly showed 1-2 spots for the TelG probe. Both TelC and TelG counts significantly distinguished the group of ALT-positive cell lines from telomerase-positive cell lines or HUVEC (Padj < 2e10-16, Wilcoxon rank sum test).

According to the co-culture experiment presented in *section 2.7*, a cutoff of >2 nuclear spots had been identified as optimal for separation. When relating this threshold to the population average spot count for each cell line, the ALT-positive group was clearly above (Average spot counts, TelC: 9.11-18.11, TelG: 3.30-6.38). Of the weakly ALT-positive cell lines, only NEM157 exceeded this cutoff for the TelC spot count (3.00) and contained ~42% of cells classified as ALT-positive. NEM165 cells fell below the TelC and TelG cutoffs and the proportion of ALT-positive cells was within the previously estimated FPR range (3.8% and 0.8%, respectively, compare *section 2.7*). For comparison, 2.9% (TelC) and 0.3% (TelG) of the ALT-negative HeLa wt cell line were above the cutoff. Out of the two weakly ALT-positive cases, one was identified as containing a significant fraction of ALT-positive cells using ALT-FISH. Nevertheless, it fell below the typical threshold of 10% of U2OS levels for being considered as ALT-positive by the CC assay. Overall, I found very good concordance between CC levels and the average numbers of ALT-FISH spots. The quantitative relationship of CC levels and ALT-FISH signals was examined using a larger sample set at later time and is presented in *Results section 3.3*. Although cancer cell lines in culture are typically considered to be homogeneous, the ALT activity analysis conducted here showed significant heterogeneity. The count distribution was right-skewed, with maximum counts reaching up to 11 times the median of the population (compare **Figure 20 A**).

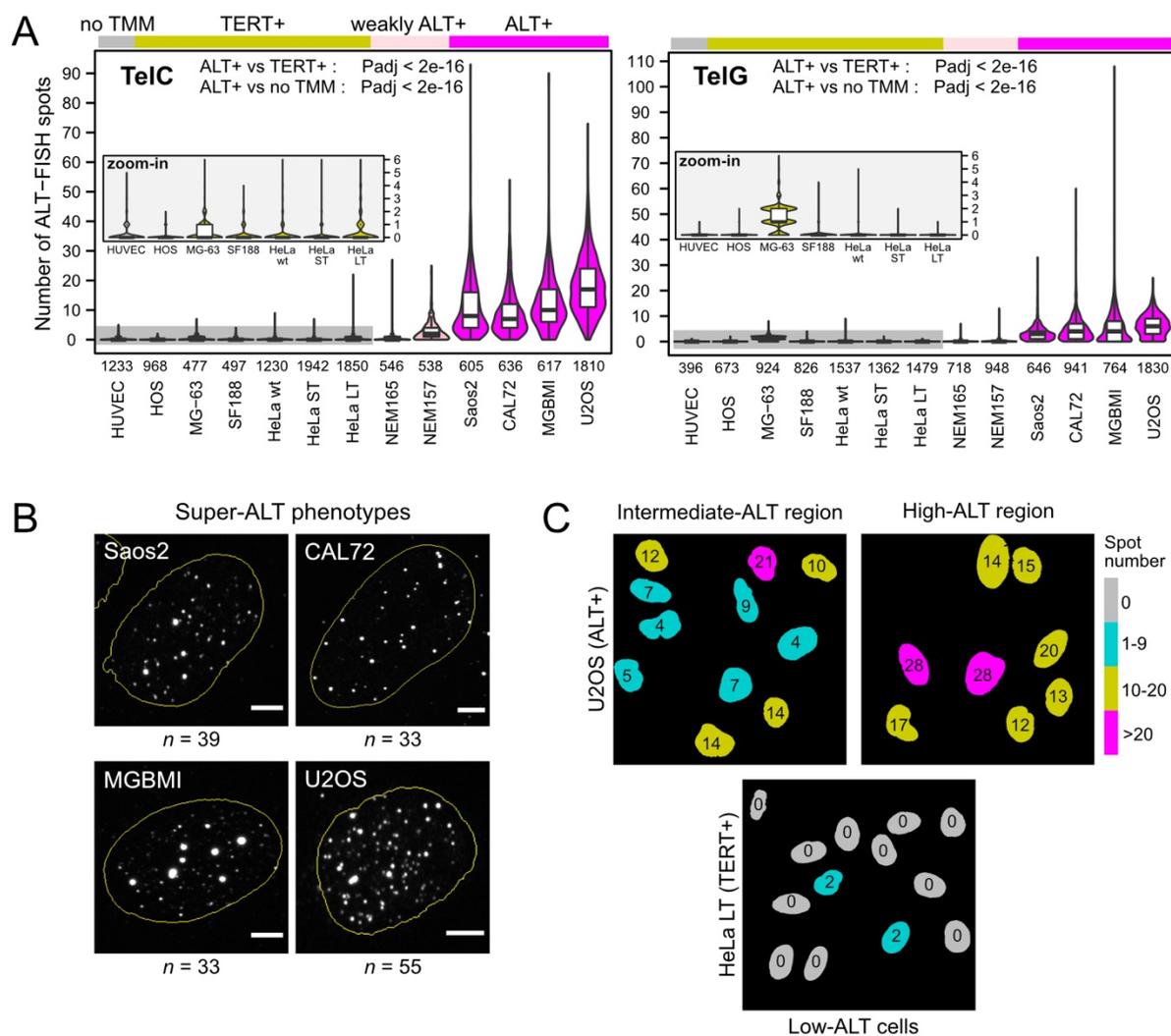


Figure 20. ALT-FISH resolves single cell ALT activity distribution across cell lines. (A) Quantification of nuclear TelC and TelG spot counts for the cancer cell line panel. Data was pooled from 2-4 independent replicates per cell line. All boxplots depict median, first and third quartiles with the number of analyzed cells indicated below each plot. Zoom-ins into no TMM/TERT+ samples are shown for better visualization. Adjusted P values from a pair-wise Wilcoxon rank sum test with BH correction are indicated for comparisons between TMM groups (weakly ALT+ included into ALT+ group). (B) Examples of super-ALT phenotypes (>95th percentile of the count distributions) in Saos2, CAL72, MGBMI and U2OS ALT-positive cell lines with spot count n given below each image. Scale bars, 5 μm . (C) Spatial ALT activity patterns shown for selected image positions of U2OS and HeLa LT cells, built from TelC spot counts. Examples of intermediate to high ALT activity clusters in U2OS (top) and rare cells with low ALT activity found within the HeLa LT population (bottom). Figure adapted from ref. [165].

Such subpopulations with high ALT activity were consistently present in the ALT-positive cell lines. Defined by counts above the 95% distribution percentile, I termed them 'super-ALT' cells (**Figure 20 B**). Whether this observation links to an intrinsic biological property of ALT needs to be tested. High local accumulations of SSTRs could potentially stem from bursts in telomere processing. But also telomere dysfunction related polyploidization [185] or telomere trimming events [186, 187] could drive variations in telomere nucleic acid content.

Single cell spatial resolution is a key benefit of ALT-FISH that the CC assay does not offer. For the underlying cell line panel, it was exploited to explore more subtle differences in ALT activity. I compared the ALT-prone HeLa LT cell line to its isogenic counterpart HeLa ST and found the distribution of TelC signals in the HeLa LT population to be more similar to the weakly ALT-positive cell line NEM165, with even slightly more cells classified as ALT-positive (6.2% vs. 3.8%) (**Figure 20 A**). In contrast, HeLa ST and the non-isogenic HeLa wt clone only had 1.1% and 2.9% of cells with more than two spots. The difference was stronger when the ALT classification stringency was reduced to cells that had two or more counts. HeLa LT and NEM165 had 10.2% and 9.1% cells falling into this category, while HeLa ST and HeLa wt only had 2.7% and 5.2%. These results demonstrated that the identification of subpopulations with low ALT activity profiles can inform about the predisposition of a cell line to engage in ALT.

I furthermore found that some of the ALT-FISH heterogeneity appeared to be spatially encoded. Cells with identical spot counts were frequently observed close together, and there were hubs of low ALT activity in ALT-negative neighborhoods as well as areas with overall higher or lower ALT activity (**Figure 20 C**). This could at least in part be explained by a model in which the SSTR abundance is maintained shortly after cell division. An in-depth investigation of this phenotype and other factors linked to heterogeneity was conducted on the high content ALT-FISH datasets presented in the *Results* sections 3.6 to 3.9. In conclusion, accurate mapping of ALT activity across three cancer types was shown by the study of the cell line panel. The single cell resolution can be utilized to identify the dominant TMM of a cancer cell population and to visualize (spatial) variations in ALT activity.

3. Mapping the distribution of ALT activity in unperturbed and perturbed cell populations using ALT-FISH

As demonstrated in the previous chapter, ALT-positive cancer cells were found to contain large amounts of SSTRs from different sources. These SSTRs compartmentalized into nuclear and cytoplasmic foci and could be visualized by the ALT-FISH assay. The abundance of SSTR compartments was found to be a highly sensitive and specific readout for ALT activity in single cells valid for different tumor entities. ALT-FISH profiles of cell lines furthermore uncovered heterogeneity in ALT activity across single cells and between cell lines. How TMM diversity is established and how ALT activity is re-distributed when cell populations are confronted with perturbances that evoke TMM changes are open questions in the field, yet highly relevant to the understanding of telomere maintenance regulation in tumors.

In the following chapter, I employed ALT-FISH to characterize changes in ALT activity in three established cell line models of perturbation that include artificial ALT induction, ALT potentiation and ALT suppression. Furthermore, I adapted the ALT-FISH staining procedure, image acquisition and automated data analysis workflows to a microplate sample format that enables the screening of ALT activity in hundreds of thousands of individual cells. High-content datasets obtained by this approach were used to assess the influence of cell cycle and spatial neighborhoods on the variability of ALT-FISH signals and to test the feasibility of the here developed setup for conducting perturbation screens.

3.1 Capturing changes in ALT activity upon telomerase-to-ALT switching

For classification, it is a common and needed simplification to assign a dominant TMM samples based on the joint assessment of various telomere-related features (CC levels, TERT expression, presence of APBs, telomere length, etc.). However, the resolution of ALT-FISH has the potential to uncover TMM switching or bifurcation events which involve the ALT pathway. Evaluation of the cell line panel in *section 2.8* did suggest the presence of ALT mosaicism in some cell lines with low CC levels. I therefore tested next, if ALT-FISH could also resolve TMM switching in cancer cells (**Figure 21**).

To this end, I focused on the artificial ALT induction by histone chaperone ASF1A/B knock-down in the previously introduced telomerase-positive HeLa LT cell line [181] and its isogenic control HeLa ST, which does not induce ALT. As in the original study, ASF1A and ASF1B were co-depleted in HeLa ST and HeLa LT by siRNA transfection and cells were grown for 72 hours before assayed. Western blot analysis of samples from five independent knock-down experiments confirmed a strong reduction (~80-90%) of both ASF1A and ASF1B protein levels compared to a non-targeting siRNA (siCtrl) (**Figure 21 A**). To assess ALT induction, CC levels were measured for all five replicates and matched TelC and TelG ALT-FISH samples of three knock-down experiments were obtained in parallel (**Figure 21 B**). Relative CC levels were significantly increased from $1.5\% \pm 0.5$ to $14.3\% \pm 6.5$ ($P = 0.0079$, Wilcoxon rank sum test, siASF1 vs. siCtrl) in HeLa LT, but undetectable all HeLa ST samples, indicating the induction of ALT activity in HeLa LT.

ALT-FISH data obtained from these samples indicated an increase in the proportion of cells with both TelC and TelG spots in ALT-induced HeLa LT (**Figure 21 C, D**). To capture the smallest changes in ALT activity, the data was binned into four groups with 0 spots (no ALT), 1-2 spots (low ALT activity), 3-4 spots (robust ALT activity) and >4 spots (high ALT activity). A cross-comparison to the datasets for HeLa ST and LT shown in *Results section 2.8* revealed no major differences between unperturbed cells and the siCtrl condition. The increase in TelG spot number strictly followed the changes in CC levels of HeLa LT (compare **Figure 21 B, D**). Upon knock-down, the number of cells with at least one TelG spot increased from <5% to about 25%, while this proportion stayed constantly below 5% for all HeLa ST samples. This correlation supported the hypothesis that C-circles are what the TelG probe primarily detects. While the most dramatic increase in cells with TelC spots was observed for HeLa LT depleted for ASF1, the fraction of cells with more than two TelC spots also increased from >3% to about 10% in HeLa ST. ALT induction has not been described for this cell line and these findings implied the detection of minor changes in ALT activity that had previously been missed by the CC-based ALT detection. Further experiments are needed to determine whether ALT activity changes in HeLa ST cells are also related to the same phenotypic features originally described

for ALT induction in HeLa LT [181]. In summary, investigation of this ALT induction system demonstrated that the method is suited to detect telomerase-to-ALT switching. In the future, the kinetics of this process can be better resolved by conducting ALT-FISH on samples collected throughout the switching phase.

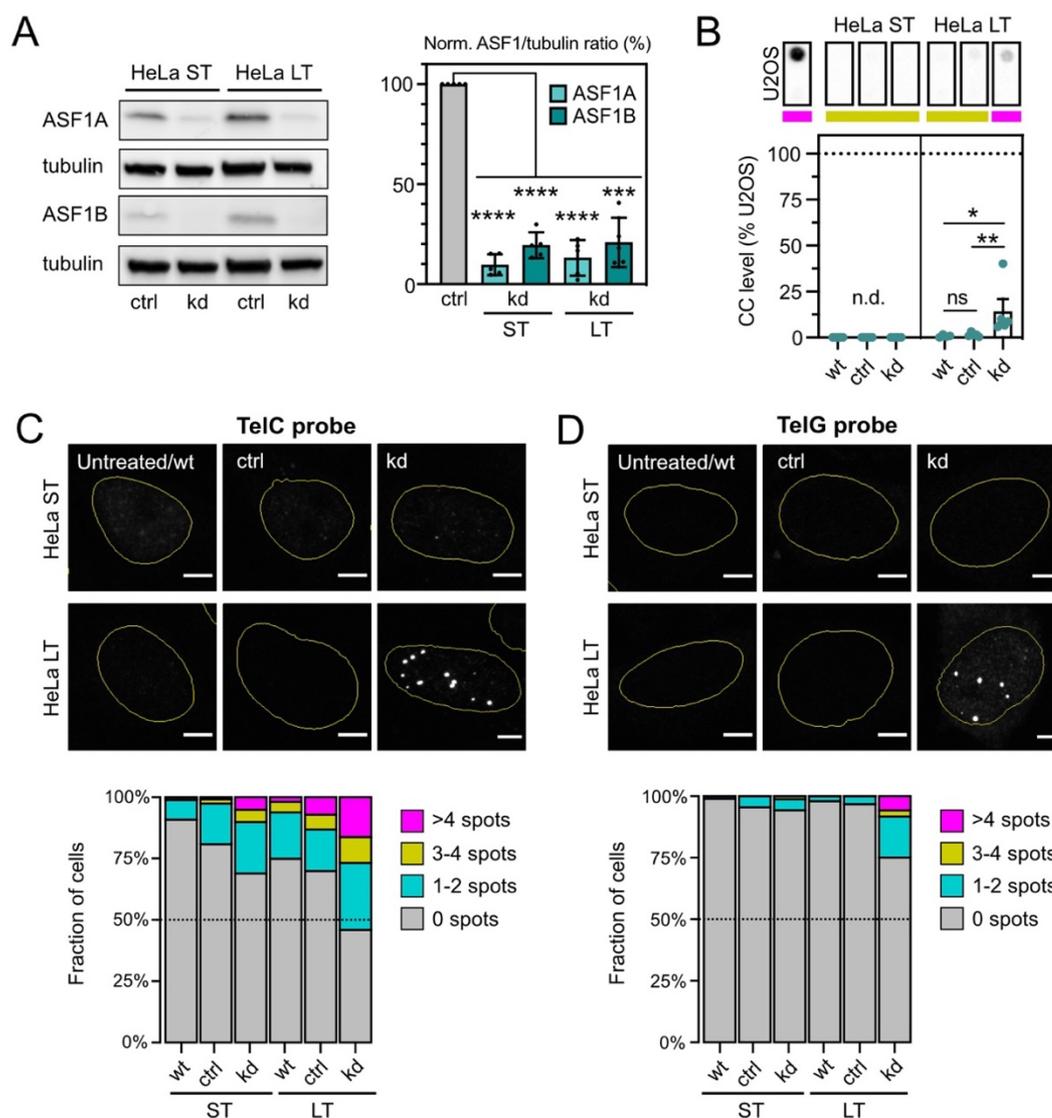


Figure 21. Mapping telomerase-to-ALT switching in HeLa LT cells depleted for the histone chaperone ASF1. (A) ASF1A/B western blot and quantification of siRNA-mediated ASF1 knock-down (kd) in HeLa LT and HeLa ST (three days post-transfection). Protein levels were normalized to tubulin and a matched non-targeting siRNA (ctrl) sample. Bar plots depict the mean and s.d. plus individual values of $n = 5$ independent experiments. One sample Student's t test, *** $P = 0.0001$, **** $P < 0.0001$. (B) Quantified CC level changes upon ASF1 depletion (representative dot blots on top) matched to the samples in A. Color bar labels mean ALT activity above 10% of U2OS CC levels in pink, below in yellow. Bar plots show mean and standard error plus individual values (n.d., not detectable). Pair-wise Wilcoxon rank sum test, $P > 0.05$ (ns, not significant), * $P = 0.015$, ** $P = 0.0079$. (C) Images of TelC ALT-FISH signals and matching quantification results for the ASF1 knock-down. Count data was pooled from three independent experiments. Data from untreated HeLa cells (wt) were included as reference and are the same as in Figure 20. Scale bars, 5 μm . (D) Same as C, but for the TelG probe. Sample collection for western blotting and CC assay were carried out by Caroline Knotz. Figure adapted from ref. [165].

3.2 Perturbing ATRX to study genetic ALT activity modulation

ASF1 depletion is an attractive system to study the induction of ALT *in vitro*. However, ASF1 loss-of-function is not an established feature of ALT-positive tumors. In order to validate my findings in another genetic context that has been associated with ALT, I turned to studying perturbations of the chromatin remodeler and ALT suppressor ATRX, which is recurrently mutated in various entities [108, 127, 188]. ATRX depletion has been demonstrated to occasionally activate ALT in somatic or telomerase-positive cells [189, 190], while ectopic ATRX re-expression in ALT-positive cell lines has been associated with a reduction of ALT-related characteristics [191, 192]. I compared two different perturbations of ATRX with regards to their effects on single cell ALT activity. First, ATRX loss was investigated in an isogenic cell line model (**Figure 22**). Second, ATRX-deficient U2OS cells with doxycycline-inducible ATRX expression (U2OS-ATRX) [191, 192] were utilized (**Figure 23**).

Our group had previously characterized two isogenic knock-out clones of the weakly ALT-positive pedGBM cell line NEM168 harboring ATRX gene indels (B5 and F5) that had entirely lost the ability to express ATRX protein and displayed elevated CC levels [180]. I was able to confirm this phenotype by western blot and CC analysis (**Figure 22 A, B**). NEM168 wt cells had low CC levels ($3.2\% \pm 0.1$ of U2OS levels), classifying them as weakly ALT-positive. CC levels were about 10 times higher in clone B5 ($35.4\% \pm 6.3$) and more than 40 times higher in clone F2 ($141.7\% \pm 52.1$), demonstrating that ATRX loss had increased ALT activity.

Their analysis by ALT-FISH staining revealed that the average spot count in wt cells was remarkably close to that of the previously studied weakly ALT-positive cell line NEM165 (compare *Results section 2.8*, ~ 0.5 spots/nucleus for TelC and ~ 0.2 spots/nucleus for TelG), consistent with its CC levels being low. For the B5 and F2 clones, respectively, the average number of TelC spots per nucleus rose markedly by 5 to 6 fold upon loss of ATRX and the average TelG spot count was even 8 to 10 times greater ($P < 2e-16$, pair-wise Wilcoxon test with BH correction) (**Figure 22 C**). Additionally, the CC abundance differences between the two knock-out clones were clearly in line with the different ALT-FISH spot distributions. I next characterized the effect of ATRX re-expression on ALT-FISH signals in U2OS cells (**Figure 23**). As previously measured by CC assay in this cell line, ALT activity decays when ATRX is re-expressed upon addition of doxycycline. CC levels have been found to be reduced by $\sim 80\%$ after three days, reaching the detection limit at seven days induction, indicative of a potent ALT suppression [180, 191]. I chose three days to circumvent confounding effects caused by long-term ATRX expression in U2OS. Three days post-induction, ATRX protein in U2OS-ATRX cells was still detectable, but at a lower level than in ATRX wt HeLa cells. No ATRX protein was detectable in uninduced U2OS-ATRX and in non-transgenic ATRX-deficient U2OS wt cells (**Figure 23 A**).

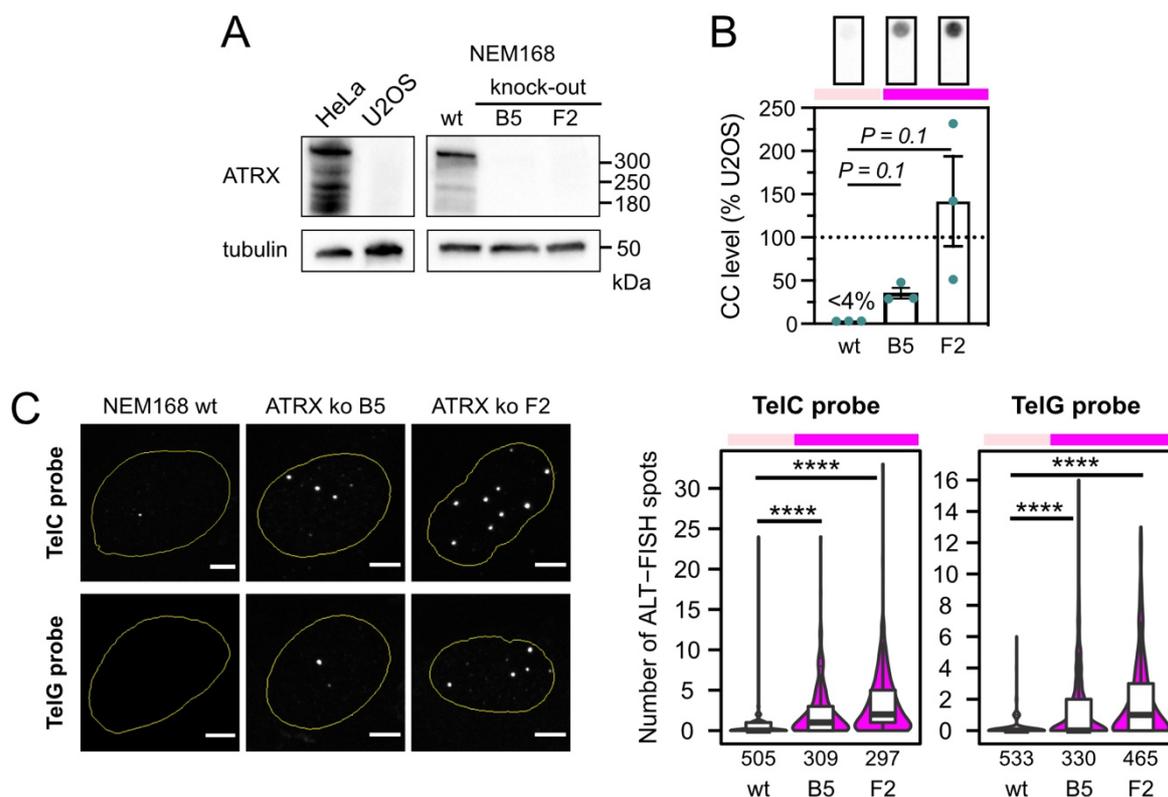


Figure 22. Mapping ALT activity changes upon ATRX loss-of-function in pediatric glioblastoma cell lines. (A) ATRX western blot for the two NEM168 knock-out clones (B5, F2) versus their parental line (wt). ATRX wt HeLa and ATRX deficient U2OS cells served as controls. (B) Quantification of CC level changes for three replicates (mean and s.e.m plus individual values, dotted line: U2OS). P values: one-sample Student's t test. (C) Images of TelC and TelG staining in NEM168 wt, B5 and F2 with matched quantification of nuclear spot numbers pooled from 2-4 independent replicates. Pair-wise Wilcoxon rank sum test with BH correction: adjusted P values $*** < 2e-16$. All boxplots depict median, first and third quartiles with the number of analyzed cells indicated below each plot. Scale bars, 5 μ m. Western blotting and CC assay were carried out by Caroline Knotz. Figure adapted from ref. [165].

In line with the published data on this cell line, CC levels significantly decreased to $13.1\% \pm 3.8$ of the U2OS wt control in the induced sample ($P = 0.0019$, one sample Student's t test) (**Figure 23 B**). However, a moderate reduction to $69\% \pm 14.5$ ($P = 0.16$, not significant, one sample Student's t test) was also observed when comparing the uninduced sample to U2OS wt, possibly indicating leaky ATRX expression or variable CC abundance between the two U2OS clones. ALT-FISH staining revealed a striking loss of TelG spots upon induction (median spot number; uninduced: 6, induced: 0, $P < 2e-16$, Wilcoxon test with BH correction) (**Figure 23 C**). Thus, and consistent with the ASF1 depletion and ATRX knock-out clones, the TelG signal showed a strong correlation with CC abundance. The TelC spot number did not follow this trend (compare **Figure 23 C**). Instead, both the induced and uninduced cells consistently had more spots. I hypothesized that differences in TERRA levels were likely responsible for the observed differences and thus repeated the ATRX induction experiment with an RNase A/H pre-treatment (**Figure 23 D**).

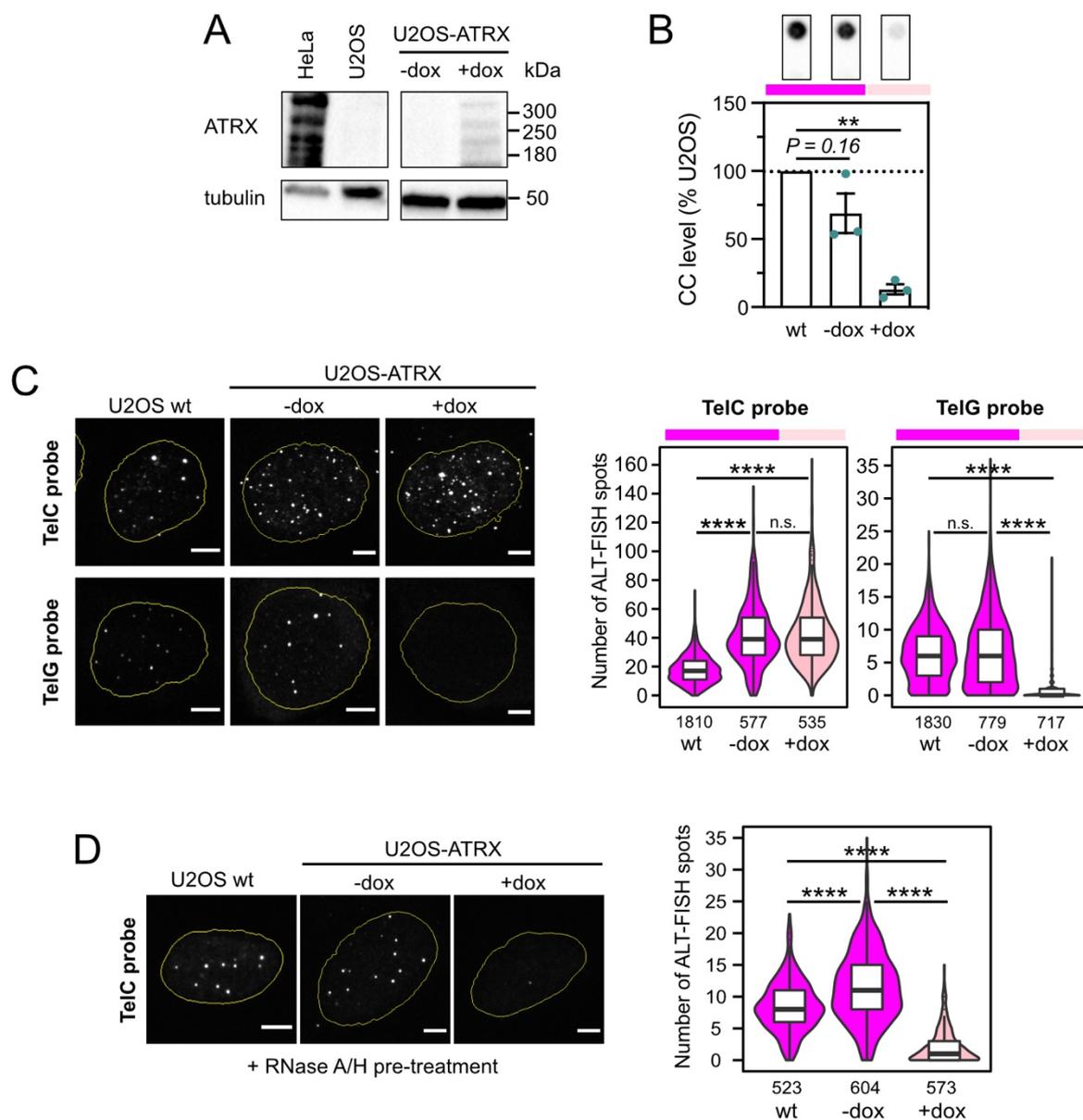


Figure 23. Monitoring ALT suppression by ATRX using ALT-FISH. (A) Western blot to confirm doxycycline-inducible ATRX expression in the U2OS-ATRX cells. ATRX wt HeLa and ATRX deficient U2OS cells served as reference. (B) Corresponding CC level reduction quantified for three independent replicates. Bar plot shows mean and standard error plus individual values. Dotted line: U2OS reference. One-sample Student's t test, $P = 0.16$, $**P = 0.0019$. (C) Images and quantification of TelC and TelG ALT-FISH staining in U2OS-ATRX cells +/- dox (three days). U2OS wt data from Figure 20 A is shown for comparison. Data was pooled from three independent induction experiments. (D) Same as in C, but only for the TelC probe and with pre-treatment by RNase A+H to remove TERRA. Pair-wise Wilcoxon rank sum test with BH correction, $P > 0.05$ (n.s., not significant), $**** P < 2e-16$. All boxplots depict median, first and third quartiles with the number of analyzed cells indicated below each plot. Scale bars, 5 μm . Western blotting and CC assay were carried out by Caroline Knotz. Figure adapted from ref. [165].

Similar to the TelG probe signals, RNase-resistant TelC spots were massively reduced upon ATRX expression (median spot number; uninduced: 11, induced: 1, $P < 2e-16$, Wilcoxon test with BH correction), confirming the hypothesis. The cause of the higher TERRA levels in U2OS-ATRX compared to U2OS wt was not investigated. It might have appeared during the clonal selection process used to create the cell line. This example highlighted that the combination of the TelC and TelG probes with RNase pre-treatment provides the most complete evaluation of ALT activity.

3.3 Integrative comparison of ALT-FISH and C-circle data reveals robust correlation

For the various cellular model systems examined here, the ALT-FISH results coincided with data obtained by the CC assay, which can be regarded as the current gold standard for ALT detection [120]. In this section, I aimed at providing a quantitative side-by-side comparison of the two methods by integrative analysis of the ALT-FISH and CC data obtained so far (**Figure 24**). This comparison sought to assess overall robustness of ALT-FISH as a new method and any differences between using the TelC or TelG probe.

I collected ALT-FISH spot counts from all previously examined samples. For each condition, the replicates were combined and counts were averaged. Matching CC values were obtained by normalizing the values measured in triplicates to a calibrator sample of U2OS gDNA on the same membrane as percent level to the U2OS reference. The resulting CC values and ALT-FISH spot counts were log-transformed and plotted against one another, separately for the TelC and TelG probe. Before log-transformation, a pseudo count value of 0.1% was added to all zero CC values to also include these samples. Spearman's rank correlation coefficient R was calculated for both the TelC and TelG dataset. Correlation coefficients of $R = 0.89$ for TelC (**Figure 24 A**) and $R = 0.88$ for TelG (**Figure 24 B**) indicated a very strong positive correlation between CC abundance and ALT-FISH spot counts. This result demonstrated that averaged TelC and TelG spot counts were both accurate predictors of the overall CC abundance, suggesting that both probes can be used interchangeably for estimating ALT activity.

The comparison also highlighted a number of outliers. The telomerase-positive cell line MG-63 was clearly separated from most samples in the TelG graph (**Figure 24 B**), since it had barely detectable CC levels, but consistently displayed 1-2 TelG foci (see *Results section 2.8*). As expected, samples from the ATRX re-expression experiment also deviated by low CC levels and high TelC counts caused by elevated TERRA. ALT-positive samples could be well separated from ALT-negative samples using a CC abundance cutoff as low as 6%. Bulk ALT positivity was correctly assigned when using a cell-population averaged ALT-FISH spot count cutoff of ~ 1 for the TelC probes and ~ 0.08 for the TelG probe. Comparison of the TelC and TelG probe results suggest that the TelC signals offer a larger dynamic range for single-cell

ALT activity analysis. The TelG probe, on the other hand, may provide a more sensitive bulk separation of ALT-positive and ALT-negative samples. The integrative analysis furthermore suggested that the techniques used to detect and evaluate ALT-FISH signals enable more direct comparisons of absolute ALT activity levels between samples without the need for normalization to a calibrator sample as in the case of the CC assay.

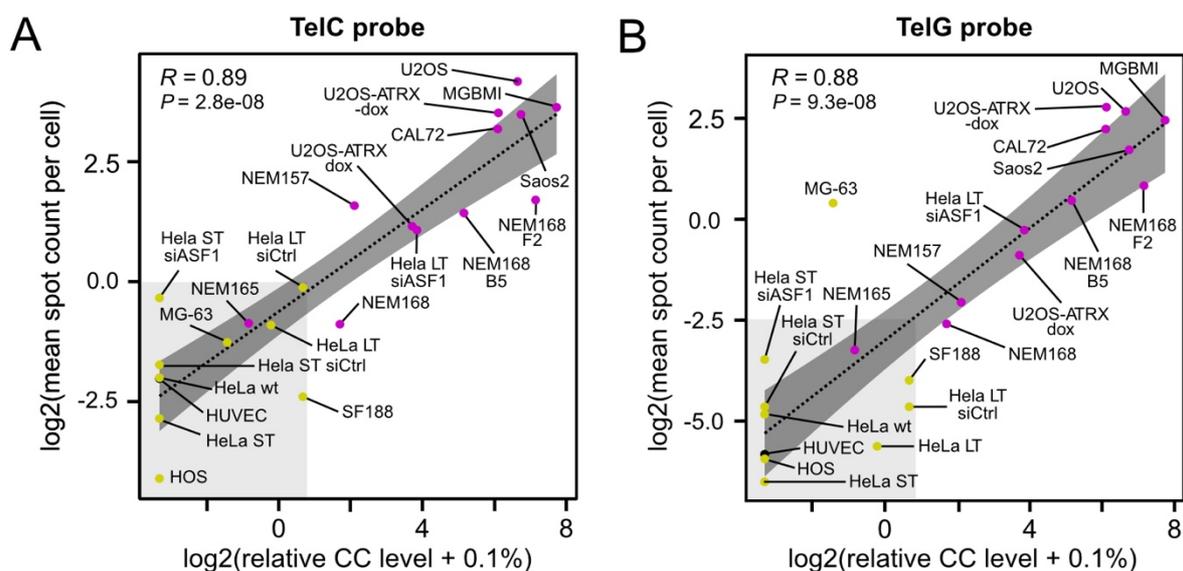


Figure 24. Correlation of C-circle levels and ALT-FISH signals across conditions in cell lines. (A) Scatter plot of CC levels (expressed in % U2OS from the same membrane) versus mean TelC spot count per cell. CC values were averaged across biological replicates. A pseudo-value of 0.1% was added to include samples with undetectable CC levels. Mean spot count per cell represents the mean of all analyzed cells across biological replicates. R corresponds to the Spearman correlation coefficient with P value below. Dotted line with 95% CI as ribbon provides a guide to the eye. Grey shaded area comprises ALT-FISH spot count and CC level values that yield a good separation between ALT-positive (pink) and ALT-negative (yellow, black) samples. (B) Same as A, but for the TelG probe dataset.

3.4 Adaptation of the ALT-FISH assay to high-content screening formats

The effect of drugs or genetic perturbations on survival of ALT-positive cells has been tested by various bulk cell approaches to discover vulnerabilities that might be exploited for targeting ALT-positive tumors [193-196]. One limitation of these strategies is their limited throughput and often little insights into the links between ALT activity and cellular traits. These limitations were addressed here by developing a high-content ALT-FISH screening setup for studying a high number of perturbations in parallel. The ALT-FISH assay is compatible with screening due to three main features: (i) the single hybridization-based detection principle, (ii) the minimal liquid handling effort and cost-effective reagents, and (iii) the automatable data analysis. Furthermore, it can be integrated with other microscopic readouts of cell phenotypes. ALT-FISH screening is introduced here for 96-well glass bottom microplates, which are

commonly employed in high-content screening applications (**Figure 25**). Some adjustments were made to the original staining procedure to reduce liquid handling time. Image acquisition was optimized to image both ALT-FISH and DAPI signals at reasonable throughput and resolution with a 20x air objective. Each microplate well of 6 mm diameter was recorded as 11x11 tile scans of 15 μm z-stacks (**Figure 25 A**). Automated imaging of one microplate comprising 60 usable wells could be accomplished in below 24 hours. At a maximum density of 20,000 cells per well, this amounted to a possible throughput of screening ~ 1 million cells per microplate per day.

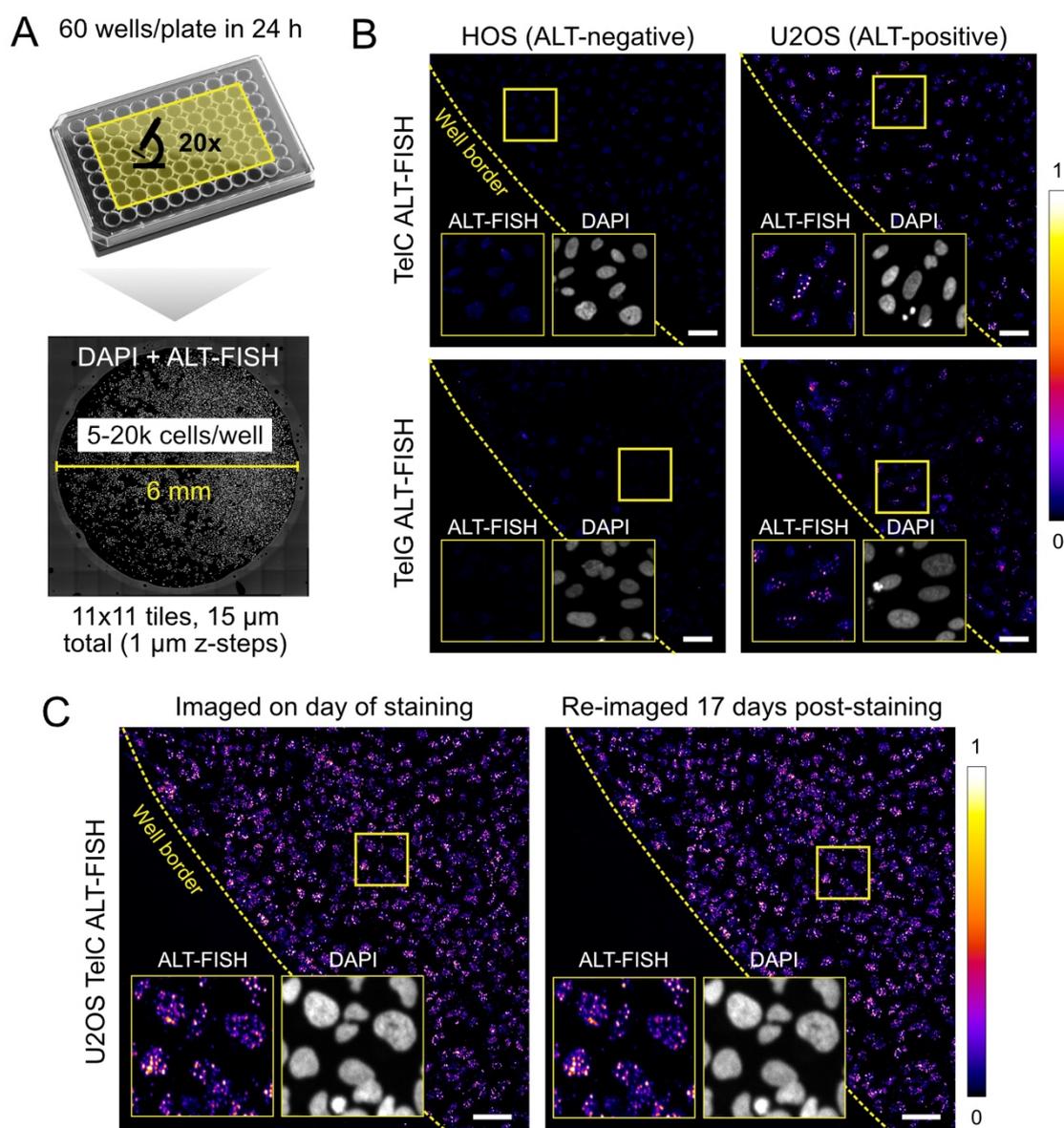


Figure 25. ALT-FISH in screening-compatible microplates. (A) Overview of the plate layout and imaging strategy to acquire ALT-FISH and DAPI signals for large cell numbers and multiple conditions. (B) Representative staining images for ALT-negative HOS or ALT-positive U2OS cells grown in microplates. (C) Comparison of TeIC ALT-FISH signal stability for the same sample of U2OS cells grown in a microplate. Between the two imaging rounds with the exact same acquisition parameters, the sample was stored in 2xSSC buffer at 4°C for 17 days. Scale bars, 50 μm .

As an initial validation experiment, ALT-positive U2OS and ALT-negative HOS osteosarcoma cells were stained by either the TelC or TelG ALT-FISH protocol (**Figure 25 B**). The staining of U2OS clearly displayed distinct spots for both the TelC and TelG probe, whereas HOS were largely devoid of SSTR foci. These differences in signals were in concordance with the previously acquired high resolution ALT-FISH images of these two cell lines (see *Results section 2.8*). Since image acquisition periods are long for microplates and temporal separation of staining and imaging is often desirable for screens, sample stability must be ensured. I therefore compared ALT-FISH signal quality in U2OS for the same sample region at the same day of staining to 17 days after staining (**Figure 25 C**). Between the two rounds of imaging, the microplate had been stored at 4°C in 2xSSC imaging buffer. From this comparison, no significant deterioration of ALT-FISH signal became apparent and cell nuclei positions and integrity were unaffected. In conclusion, I could confirm compatibility of the ALT-FISH assay with microplate screening formats with this line of experiments.

3.5 Automated analysis and quality control of high-content ALT-FISH data with *Telosegment-HT*

The *Telosegment* software toolkit presented in *section 2.3* was developed for automated analysis of datasets comprising hundreds of cells. High-content data derived from microplates has different requirements with regards to the analysis. To account for this, I developed the *Telosegment-High Throughput (Telosegment-HT)* workflow (**Figure 26**), which relied on additional software for increased data processing speed. The fundamental steps of the workflow were similar: (i) Pre-processing (stitching, z-projection), (ii) Primary analysis comprising nuclei segmentation and spot detection, (iii) Secondary analysis consisting of feature extraction and spot assignment and, finally, (iv) data inspection and QC. Two main improvements significantly accelerated data processing. Cellpose [197], a neural network based algorithm was used to segment nuclei. It is more flexible with regards to segmenting nuclei of different size and shape and more accurate at separating closely adjacent nuclei. RS-FISH [198], a faster, segmentation-free spot detection algorithm was employed. While the secondary analysis steps were conducted in R and were conceptually similar to the previous workflow, I established additional methods to remove image artifacts that specifically occur at microplate well borders and introduced supplementary scripts to generate automated reports for inspection and quality control of whole plates. Using the *Telosegment-HT* workflow, the duration for processing ALT-FISH data of on microplate (60 wells) distributed as follows: 24 hours for pre-processing, 2 hours for primary analysis and 24 hours for secondary analysis and QC report. For reference, ALT-FISH signals of about 1 million single cells can be acquired and evaluated in less than one week utilizing the workflows described here.

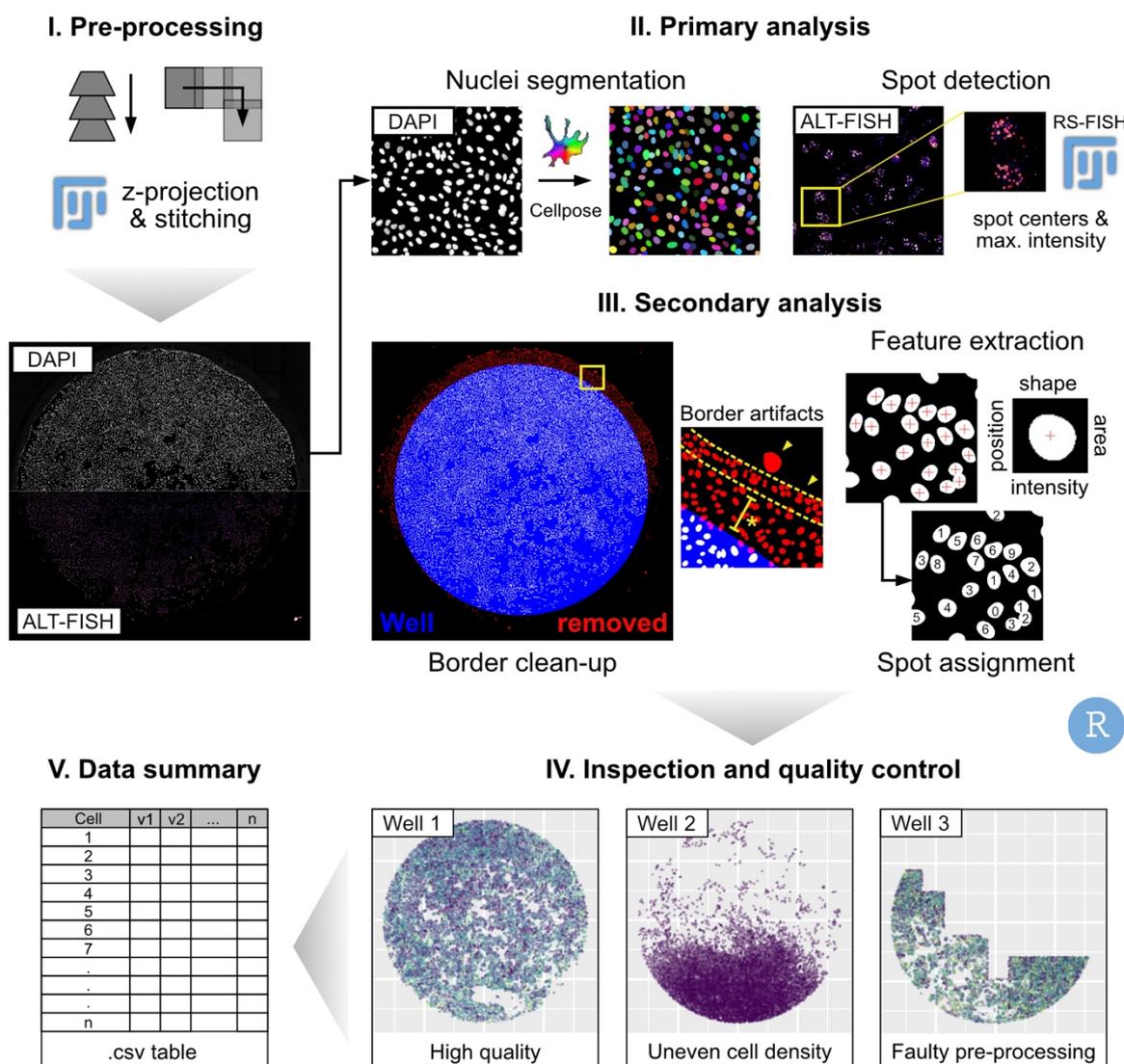


Figure 26. Automated high-content ALT-FISH data analysis using the *Telosegment-HT* toolkit. The workflow enables fully automated processing, analysis and quality filtering of ALT-FISH data from microplates. It involves (I) stitching tile-scans and creating maximum z-projections using FIJI. (II) Using Cellpose [197] and RS-FISH [198] for nuclei segmentation and ALT-FISH spot detection. (III) Removing nuclei near well borders and segmentation artifacts, and extracting image features and spot assignments. (IV-V) Generating a QC report for user-based inspection and data filtering, and final quantification results summary in a universal tabular format.

3.6 Sensitive and specific detection of ALT activity in hundreds of thousands of cells grown in microplates

Next, to evaluate the specificity and sensitivity of ALT-FISH in microplates, I quantified ALT-FISH signals of U2OS and HOS cells across 10 microplates that had been independently stained and imaged on different days (**Figure 27**). For simplicity, I focused on the TelC probe staining. Each cell line had been grown in six wells on each plate to examine technical differences between and within plates. After quality filtering, the final dataset contained 869,081 U2OS and 608,359 HOS cells (**Figure 27 A**). Assessment of the average ALT-FISH spot count per well revealed that <12% of HOS cells had one or more spots, with <0.6% having three or more spots, while 88% had no detectable spots (**Figure 27 A, B**). In contrast, >95% of the U2OS cells displayed more than one spot, with a median spot count of five. As expected, the dynamic range of spot counts was smaller than for the previously acquired high resolution ALT-FISH images. Nevertheless, the U2OS mean spot count was 35-fold higher than that of HOS cells (5.41 and 0.155 spot/cell) and enabled reliable separation of these two ALT-positive and ALT-negative cell lines. This was also the case when using the cumulative spot intensity of ALT-FISH spots per cell as an alternative measure for ALT activity (**Figure 27 C**), which correlated well with the number of spots (**Figure 27 D**). Spot counts and cumulative spot intensities were largely comparable between populations of U2OS and HOS cells grown in different wells. Some moderate batch effects were observed between plates that could reflect both biological and technical variation. I concluded that including an ALT-positive and an ALT-negative control cell line on each plate would be useful to control for variations during screening and when treatments are employed.

Next, I analyzed the high quality TelC ALT-FISH U2OS dataset of 869,081 cells presented in **Figure 27** to explore which cellular and spatial features might be linked to variations in ALT activity. Visual inspection suggested three potential sources of variability. First, it was noted that the number of SSTR compartments seemed to increase with nucleus size. Second, it seemed that densely growing U2OS cells had different SSTR abundance from sparsely growing U2OS cells. Third, hypotheses about the mechanisms responsible for the spatial patterning of SSTR compartment distribution between neighboring cells, which were previously observed in one of the earlier datasets, were tested using isolated microcolonies of clonal origin found in some of the U2OS microplate wells.

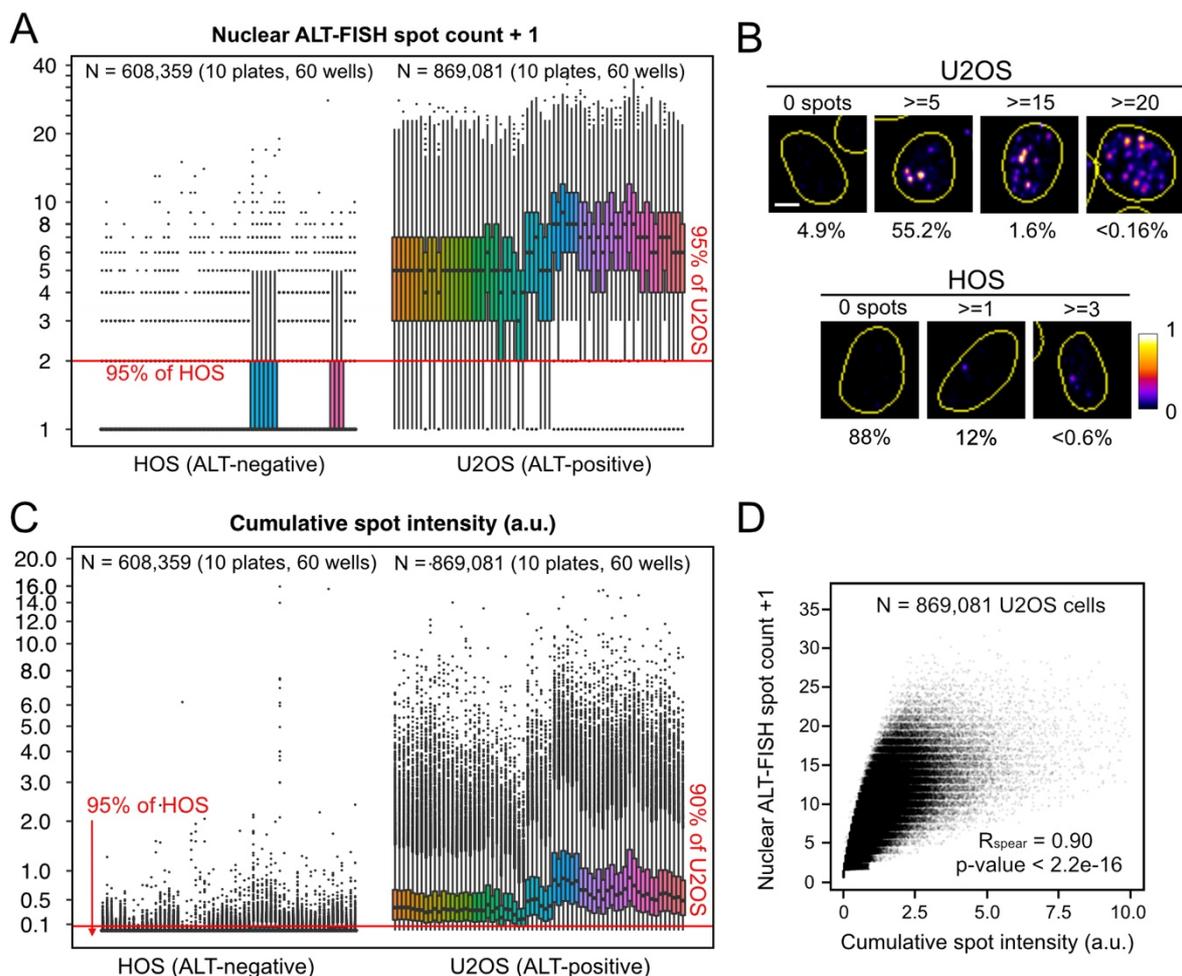


Figure 27. Dynamic range, specificity and robustness of ALT-FISH in microplates. (A) Nuclear ALT-FISH spot counts (TeIC) quantified for a total of 608,359 HOS and 869,081 U2OS cells across 120 individual wells and 10 microplates stained and imaged on separate days. A pseudo-count of 1 was added for log-representation. Each boxplot shows data from a single well (color: replicate wells from same plate). (B) Representative images of ALT-FISH signals in U2OS and HOS cells, with the frequency of specific spot counts or count ranges given below. Scale bar, 5 μm . (C) Same as A, but depicting the cumulative spot intensity (sum of all spot center intensities of a nucleus). (D) Scatterplot and Spearman correlation analysis of spot count versus cumulative spot intensity for U2OS cells revealing a strong positive correlation. Each data point corresponds to one cell. Points were jittered along y for better representation.

3.7 High-content ALT-FISH data uncovers links between ALT activity and the cell cycle

Among many factors, cell cycle stage and karyotype abnormalities are well established contributors to nuclear size [199, 200]. To follow up on the potential correlation seen between ALT activity and nuclear size in U2OS cells, I turned to analyze the relationship between ALT activity, cell cycle stage and ploidy (**Figure 28**). Both parameters can be inferred from the DNA content reflected by the integrated DAPI signal per nucleus [201], which is co-quantified during the analysis of the plates. To account for technical differences between plates, DAPI content histograms were normalized to the inter-peak distance between the G01 and G2M peaks of each plate distribution and then pooled afterwards. The center of the G01 peak was set to 1. Two normal distributions were then fitted to the normalized and pooled DNA content histogram data to model the G01 and G2M peaks of the entire population (all 10 plates). The center and spread of these two Gaussian fits was then used to group the cells into six bins based on specific normalized DNA content ranges: subG01, G01, earlyS, lateS, G2M and HP (high ploidy), as depicted in **Figure 28 A**. Definition of the groups is detailed in the *Materials and Methods* section. When examining the ALT-FISH signals for U2OS cells in the different groups, I found a gradual increase in ALT activity with progression towards G2/M phase (**Figure 28 B, C**). Both the spot count and cumulative spot intensity were lowest for cells in the subG01 group. HP cells displayed the highest ALT activity values. The gradual increase was in line with the fact that SSTRs are mostly generated during or after HR processes act on telomeres. DNA repair pathways are differentially regulated along the cell cycle. Non-homologous end-joining (NHEJ) is active throughout, but HR is nearly absent in G01. Instead, the highest HR activity is found in S phase, and then drops as the cell cycle progresses [202]. Having observed the highest SSTR abundance in the G2M group may indicate that telomeric HR intermediates or by-products are stable and persist after HR activity in S phase. The fact that HP cells had even higher levels of SSTRs than cells in the G2M group could relate to their aberrant karyotype and the higher absolute number of telomeres. Karyotype anomalies, however, may also signify greater DNA damage burden, which may be linked to more replication stress and defective telomeres, ultimately exposing more SSTRs in these cells.

Lastly, I found that early G1 cells that had just underwent cytokinesis (subG01) displayed the lowest level of ALT activity. Based on this observation, one can speculate that mitosis represents a bottleneck for the abundance of SSTRs. For extrachromosomal SSTRs, expulsion from the nucleus and subsequent loss in the cytoplasm could be one mechanism by which nuclear SSTR levels decrease, before being replenished as HR becomes active again in later cell cycle stages. In conclusion, my findings showed that links between ALT activity and the cell cycle (and HR activity) can be retrieved from high-content ALT-FISH datasets and explain part of the variation observed in ALT-positive cancer cell lines.

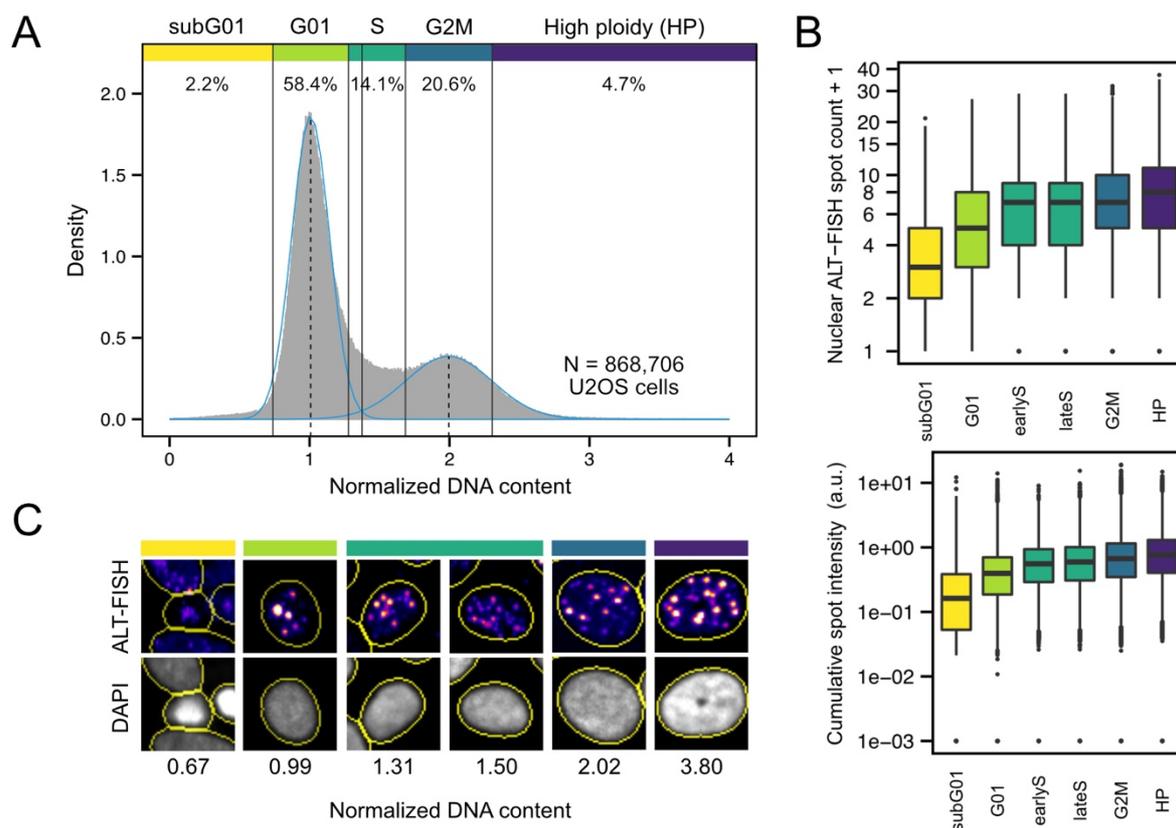


Figure 28. Links between cell cycle and ALT activity in high-content ALT-FISH data. (A) Normalized DNA content distribution derived from DAPI signals of U2OS cells from 10 microplates (same as Figure 27) and used to assign cell cycle stages. G01 and G2M peaks were modeled by two Gaussian functions (light blue curves) and assignment (vertical lines) was done as described in the *Materials and Methods* section. The fraction of cells in each stage is indicated. Note that due to technical reasons, cells with a DNA content >4 (0.04% of cells) were removed from the analysis. (B) ALT-FISH spot counts and cumulative spot intensities for cells grouped into the different cell cycle stages, showing an increase in ALT activity towards later stages and higher ploidy. (C) Representative images of ALT-FISH signals across the different cell cycle stages. Corresponding DAPI images and quantified DNA content is given below for direct comparison.

3.8 Spatial cell isolation correlates with higher ALT activity

Visual inspection of the high-content datasets suggested that U2OS cells growing at lower density had more ALT-FISH signal. I thus devised an analytical approach to test whether local cell density had an influence on ALT activity and would explain some of the observed variation (**Figure 29**). To this end, the center coordinates of the cell nuclei that had been retrieved during the image analysis procedure, were used to derive neighborhood maps (**Figure 29 A**) for six representative U2OS wells from three plates stained on different days. Two cells were defined as neighbors, when the distance between their centers was between 9-60 μm (**Figure 29 B**). Based on this definition, cells were grouped by their number of neighbors to compare their ALT activity in relation to the surrounding cell density. It revealed that both the number of ALT-FISH spots and the cumulative spot intensity were markedly higher in cells with less neighbors (**Figure 29 C**), suggesting a positive link between spatial isolation and higher ALT activity.

Since the previous analysis of the DNA content had established a link between SSTR abundance and the cell cycle stage (**Figure 28**), I wondered whether isolated cells would also show a cell cycle stage bias. I therefore integrated the cell cycle stage assignment into the spatial neighbor analysis and found in fact a trend towards later cell cycle stages and higher DNA content with increasing isolation (**Figure 29 D**). When specifically comparing cells with zero neighbors to the remaining population, G2M and HP cells were significantly overrepresented (G2M: 47.4% vs. 22.5%; HP: 10.5% vs. 1.8%; P value $<2.2\text{e-}16$ Fisher's exact test). Therefore, the cells grown in isolation may represent non-cycling cells arrested in these cell cycle phases or cells that experience a higher level of stress due to the lack of neighbor contacts. Additional experiments are needed to examine the underlying mechanisms and determine cause and consequence in relation to the cell cycle.

In summary, this analysis highlighted that ALT-FISH microplate data also contains biologically relevant spatial information, which can be quantitatively explored due to the high numbers of available cells. Such analyses can reveal links between ALT activity and cellular communities that low throughput techniques would overlook.

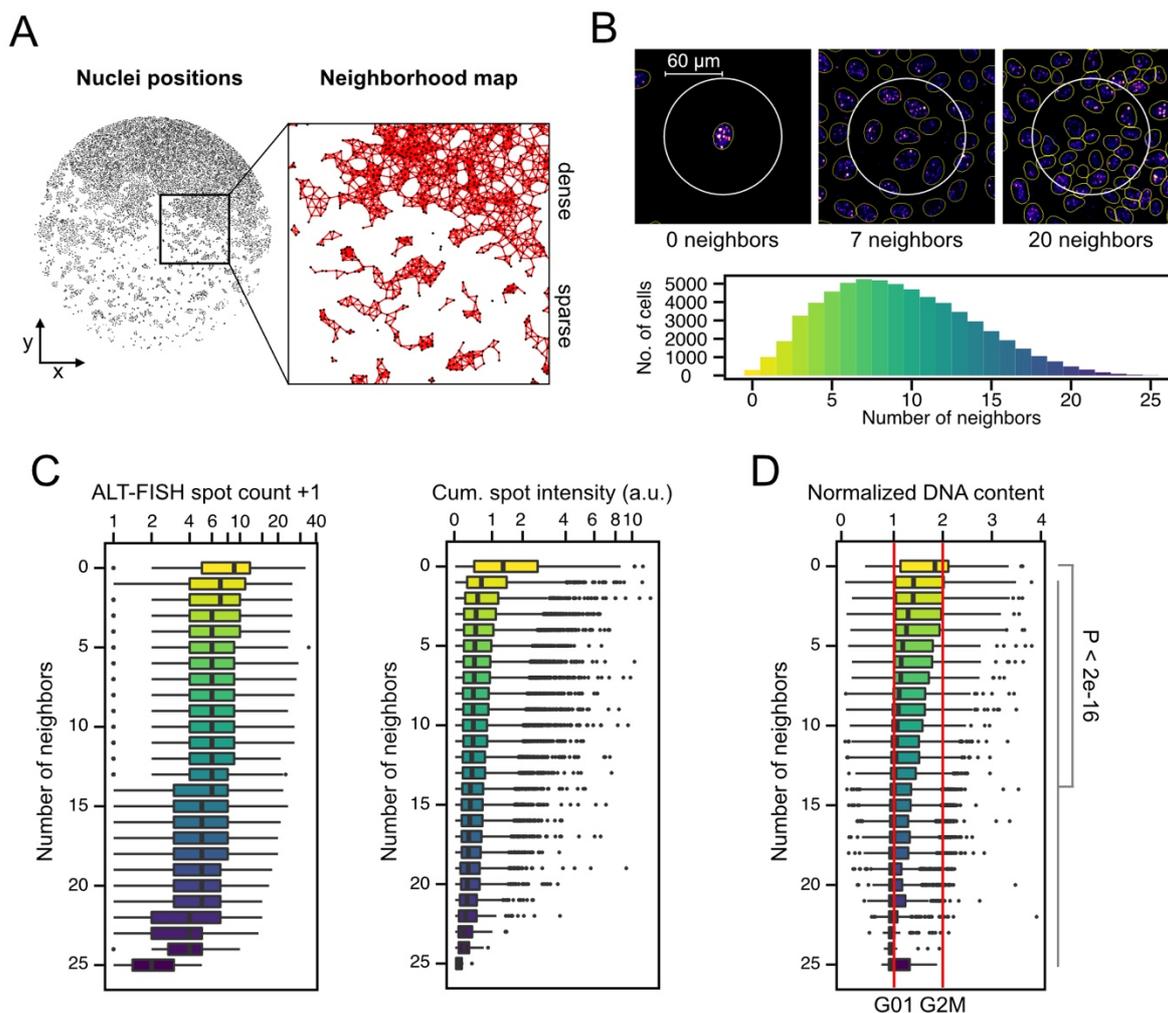


Figure 29. Contribution of local cell density to ALT activity variation. (A) Spatial neighborhood map created from linking the positions of segmented U2OS nuclei in microplate wells at a specified adjacency radius. (B) Top: representative images (ALT-FISH, nucleus outlines) of U2OS cells in neighborhoods of different density. Adjacency radius: 9-60 μm (circle shows 60 μm). Bottom: distribution of neighbor links for 63,584 U2OS cells pooled from 6 wells of comparable density across 3 microplates. (C) ALT activity measured by spot count (left) and cumulative spot intensity (right) for cells across groups of 0 to 25 neighbors. (D) Normalized DNA content of cells in relation to their neighborhood density. Centers of the G01 and G2M peaks are marked by solid red lines. The rare cell subpopulation with zero neighbors ($n = 304$ cells across all 6 wells) displayed the highest ALT activity. It differed significantly in its cell cycle stage composition from cells with more than zero neighbors, and showed an enrichment in G2M and high ploidy cells. Fisher's exact test, $P < 2e-16$.

3.9 Stochasticity of ALT activity revealed by spatial analysis of micro-communities

The timing and extent by which ALT is active on telomeres of a single cancer cell influences the spatial patterning of SSTR abundance in a cycling cell population. For example, if extrachromosomal SSTRs are generated, they may be distributed to the daughter cells upon cell division. The partitioning mechanism can hereby affect the SSTR levels of the daughter cells. Similar levels are expected in the daughter cells if these species are partitioned equally. Instead, correlation between cells that descended from the same mother cell is eliminated in the case of random partitioning. Even though interpreting such relationships for chromosomal SSTRs (which ALT-FISH also detects) is more challenging, a telomere state with high SSTR exposure may also be passed on to daughter cells. For example, if a specific telomere persists in a critically short state, it may permanently increase its likelihood of engaging in ALT and exposing SSTRs. I investigated these aspects by taking advantage of a feature present in the U2OS microplate dataset, which was the existence of cell microcolonies that grew in isolation. Cells growing in these isolated communities can be assumed to have a clonal origin and therefore can be used to study aspects of SSTR inheritance. Using the adjacency maps developed in *section 3.8*, I devised a strategy to extract U2OS microcolonies comprising 2-20 cells for their individual evaluation (**Figure 30 A, B**). Examination of the ALT-FISH spot counts and cumulative spot intensities in 105 microcolonies revealed extensive variability of ALT activity between individual clones (**Figure 30 C**). Up to 10-fold differences in SSTR abundance were observed between similar-sized colonies. These results indicated that single U2OS cells can give rise to clones with vastly different SSTR levels and underlined the stochasticity of ALT-related telomere processing [203]. I next restricted my analysis to microcolonies made up of only two cells (doublets) (**Figure 30 C**, left) to follow up on the question if some SSTRs were equally partitioned between daughter cells. Due to their limited number per well, 235 doublets from six wells on three plates were considered for the analysis.

To determine if there was more similarity between cells belonging to one doublet as compared to two cells from different doublets, I calculated a pair-wise dissimilarity index for each cell and its doublet partner, based on either the spot count or the cumulative spot intensity (see *Materials and Methods*). This was compared to a dataset in which the same cells were randomly shuffled into artificial doublets (**Figure 30 D**). A dissimilarity index value smaller one indicates greater similarity to the doublet partner than to the population median of non-partners. While the differences observed between the original and shuffled dataset were statistically not significant (P value >0.05 , Wilcoxon rank sum test), the values were slightly lower in the doublets. Thus, the analysis found no strong evidence for equal inheritance of SSTRs. Instead, the data pointed towards a stochastic model of ALT activity and SSTR distribution, which may also contribute to amplify heterogeneity in a tumor setting.

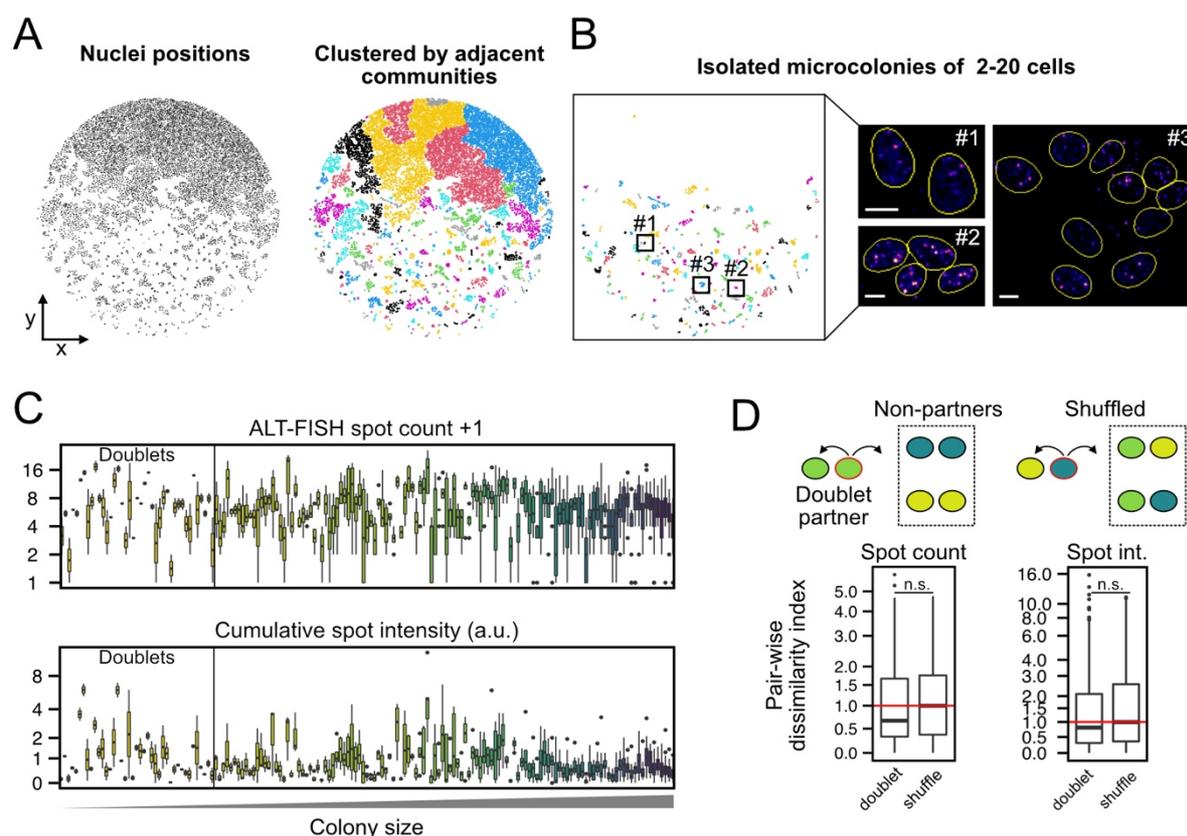


Figure 30. Analysis of ALT variation and inheritance in spatially isolated microcolonies. (A) Example of U2OS microcolonies grown in one well of a microplate. Coordinate-based adjacency graphs were used to identify clusters of adjacent communities (colored) (B) Size filtering of clusters (2-20 cells) enabled the extraction of single colonies (examples #1,2,3) and their features. Scale bars, 5 μ m. (C) ALT activity parameters for the 163 U2OS microcolonies identified in B (one boxplot per colony). (D) Doublet colonies were used to test equal versus unequal partitioning of ALT activity. A pair-wise dissimilarity index of spot counts and cumulative intensity was calculated for 235 doublets from 6 wells across 3 plates or randomly matched cell pairs (shuffle) (see *Materials and Methods*). Values center around 1 (red line) if a cell in a doublet is equally similar to its partner than to its non-partners. Wilcoxon rank sum test, $P > 0.05$ (n.s., not significant).

3.10 Drug effects on ALT activity and viability are simultaneously tested in microplates

In addition to the spatial context and cell cycle information that can be gathered from high-content ALT-FISH staining in microplates, its combination with perturbations can reveal connections between ALT activity, biological pathways and cell viability. In a proof-of-concept application, I next tested the effects of two different drugs (JQ1, Mirin) on ALT activity in U2OS cells in the here established microplate setup. These drugs were selected based on their putative functional links to telomeres or ALT-related processes.

The bromo-domain inhibitor JQ1 was selected because it had been previously found to affect telomere elongation through an unknown pathway distinct from telomerase [204]. One of the drug's major targets, BRD4 has furthermore been linked to the regulation of RNA:DNA hybrids, replication stress and DNA damage [205] and TERT expression [206]. Mirin, an Mre11-Rad50-Nbs1 (MRN) complex inhibitor, was selected as a second compound for two reasons. Activity of the MRN complex is needed for ALT-mediated telomere elongation [111, 207]. Among other functions, Mirin blocks the MRN-dependent activation of ATM signaling, which has recently been demonstrated to be essential for maintaining xenografts and cell lines derived from relapsed ALT-positive neuroblastoma tumors [195].

For testing the effect of Mirin and JQ1 on ALT activity, U2OS cells were treated for 48 hours with different concentrations (2, 4, 8, 16, 32 μM) before conducting TelC ALT-FISH and DAPI staining (**Figure 31 A**). Untreated and solvent treated cells were included on the same plate and handled in parallel. Cell viability was assessed by quantifying the cell number in each well (average of triplicate wells for each concentration and drug). The highest treatment concentration of both drugs (32 μM) caused a comparable loss in cell number (**Figure 31 B**), although JQ1 generally had a larger effect on viability at lower concentrations (**Figure 31 C, E**). I next assessed changes in ALT activity by quantifying both the ALT-FISH spot count as well as the cumulative spot intensity per cell (**Figure 31 D, F**). I found different effects of the two drugs. While Mirin treatment was not associated to marked changes in ALT activity, JQ1 treatment caused a concentration-dependent increase of both ALT activity parameters. These results showed that high-content ALT-FISH screening in microplates was sensitive at detecting even moderate changes in ALT activity caused by drugs, while at the same time providing a straight-forward readout for cell viability. The assay conducted in this fashion could thus differentiate between different modes of actions of putative ALT-targeting compounds in a screening compatible format. All here tested concentrations affected viability. By further extending the treatment regime to lower concentrations, the assay may further be used to test both viability linked and viability independent effects on ALT activity. In addition, including ALT-negative cell lines treated in parallel on the same plate could aid in the identification of compounds that affect viability in an ALT status dependent manner.

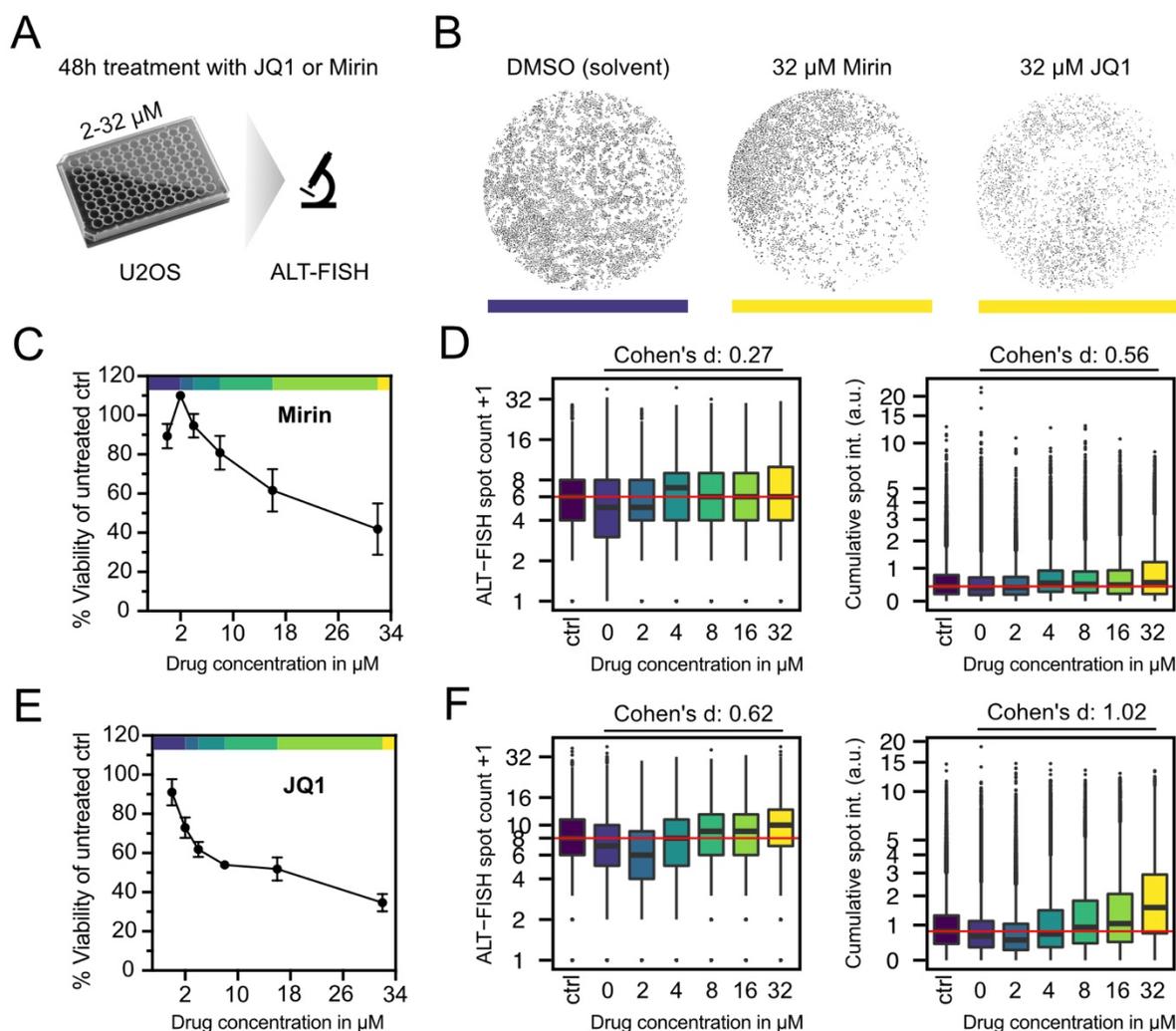


Figure 31. High-throughput detection of ALT activity and cell viability changes in response to drugs. (A) Treatment of U2OS cells with JQ1 or Mirin in microplates. Treatment was done in triplicate wells per concentration or 6 wells (untreated, DMSO). (B) Representative solvent or 32 μM drug treated wells showing a reduction in viability (cell number). (C) Relationship between cell viability and drug concentration for Mirin. Viability was normalized to the mean cell number of the untreated control wells. Points show mean and s.d. of replicate wells. (D) ALT activity parameters measured for treatment with different concentrations of Mirin, the untreated control (ctrl) and DMSO (0 μM). Boxplots depict median, first and third quartile of $n = 12,139$ to $58,054$ cells pooled from replicate wells. Cohen's d is indicated as measure of effect size between solvent and the highest drug concentration. (E) Viability plot and (F) corresponding ALT activity features for JQ1 treatment. $11,533$ to $66,557$ cells were analyzed per condition. Cell treatment and staining were conducted by Maria Panayotova, a master student supervised by me.

4. Spatial mapping of ALT activity compartments in primary tumors at single cell resolution

In the previous chapter, the ALT-FISH method was successfully used to dissect ALT activity changes in response to perturbances applied to various independent cancer cell line models. In addition, a better biological understanding of ALT heterogeneity was gained through technical advancements of the method that enabled high throughput measurements of spatially resolved ALT activity in hundreds of thousands of single cells. The question of whether the SSTR compartments detected by ALT-FISH also exist in primary tumors and whether they offer a likewise sensitive way to map ALT activity in patients was not addressed.

In the following chapter, ALT-FISH was applied to tumor tissue sections from leiomyosarcoma and neuroblastoma patients with the goals of (i) evaluating the assay's capability to determine the TMM status of patient samples, (ii) examining ALT variability and its spatial relationship to other signals in intact tumor tissues, and (iii) developing a method to obtain both ALT activity and transcriptome readout from the same single cell in its intact spatial context.

4.1 Considerations for ALT-FISH staining and image analysis on tumor tissue sections

In comparison to cancer cell lines, primary tumor tissue specimens have other requirements when it comes to their processing for fluorescence microscopy based analysis. The preservation method strongly influences the integrity of the biomolecules of interest. For this thesis, I decided to develop ALT-FISH on fresh-frozen material. Tissue sections from unfixed fresh-frozen material are substantially more difficult to collect, handle, and store. However, they reflect tumor tissue in its most native state due to the high degree of protein and nucleic acid preservation, thereby also offering the possibility of fluorescent readout multiplexing from the same sample. Since tumors are typically dense tissues, cryo-sections were prepared at a thickness of 5 μm , which corresponded to about 1-2 cell layers for the tumor entities studied here. This was crucial to minimize cell crowding, which can obstruct later image data interpretation. Like for the cell line samples, tissue images were recorded as 3D z-stacks, but reduced to 2D maximum projections for image analysis. In its general procedure, ALT-FISH staining on tumor tissue comprised the same steps as conducted for cell lines: fixation, isothermal probe hybridization and washing. However, some critical modifications were necessary to stabilize the tissue during the staining procedure and also obtain additional high quality immuno- or pseudo hematoxylin/eosin (pseudo H&E) staining from the same sample. The protocols and their critical steps are described in the *Materials and Methods* section.

For pre-processing and analysis of image data from tumor tissue, I used the methods and workflows that I had established for the high-content ALT-FISH data analysis (compare *Results section 3.5*). To account for tissue-specific artifacts, such as elevated autofluorescence in necrotic and/or highly vascularized tissue areas, I adapted the workflow by including an additional manual data curation step. The limited amount of material and the fact that the TelC ALT-FISH probe had produced a slightly better separation of ALT-positive and ALT-negative cells in the cell line experiments as well as a larger dynamic range, led me to only use the TelC probe for all staining experiments performed on tissues.

4.2 The ALT-FISH assay is compatible with primary tumor material

In order to evaluate the overall quality and intensity of ALT-FISH signals in primary tumors and explore their spatial distribution, I first conducted an experiment on leiomyosarcoma (LMS) tumor tissue (**Figure 32**). LMS is the entity with the highest documented ALT prevalence (78%) to date and is marked by substantial mutational heterogeneity and chromosomal aberrations. LMS tumor samples were obtained through a collaboration with the groups of Priya Chudasama and Stefan Fröhling at the German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT).

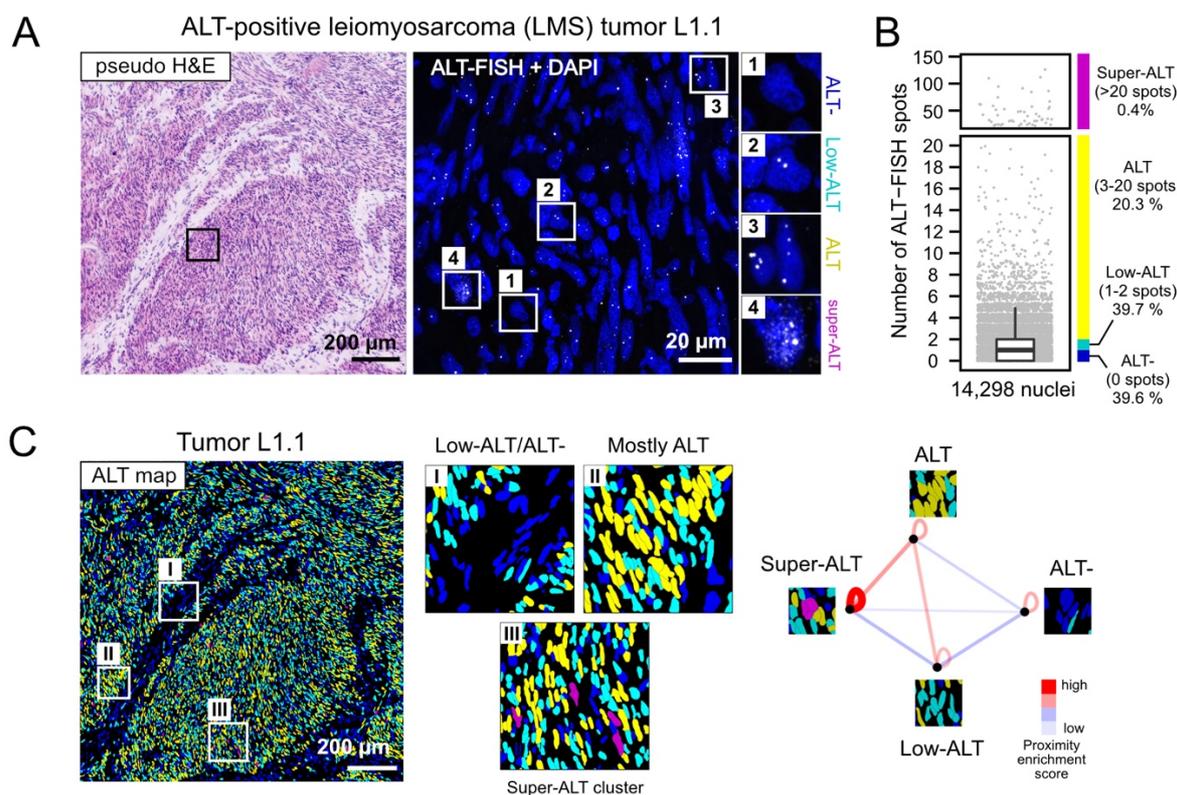


Figure 32. ALT-FISH spatially resolves ALT activity in primary tumor tissue sections. (A) 1.3x1.3 mm region of an ALT-positive leiomyosarcoma tumor tissue section stained by pseudo H&E and TelC ALT-FISH. Zoom-in shows merged ALT-FISH and DAPI of a smaller region (box) with example cells for the different ALT activity groups highlighted. (B) Spot count distribution in the region displayed in A with jittered single values of the 14,298 nuclei shown in addition. The y-axis is interrupted for visualization of the super-ALT cells with >20 spots. Bar plot to the right shows binning into the ALT activity groups. (C) Left: ALT activity map visualization that projects the grouping from B back onto the nuclear masks of the tissue region. Regions of spatial ALT heterogeneity are illustrated in the zoom-ins I-III. Right: Nearest neighbor analysis for the ALT activity groups in this tumor, suggesting preferential clustering within groups of same ALT activity. Image analysis was jointly conducted with Stephan Tirier and Anne Rademacher. Figure adapted from ref. [165].

The samples were part of a large patient cohort that had already undergone detailed molecular characterization [107]. The characterization included the assignment of ALT status by CC assay and mutational profiling by whole genome sequencing. As a positive control for assessing the ALT-FISH staining, I first selected LMS tumor L1 harboring a missense mutation in the ATRX gene and being among the samples with the highest CC level in the aforementioned study cohort (sample ID LMS28 in ref. [107]). I conducted both ALT-FISH and pseudo H&E staining [208] on a section of L1 to visualize ALT activity and at the same time assess overall tissue morphology (**Figure 32 A**). In agreement with the previous experiments, ALT-FISH staining produced distinct fluorescent SSTR foci in the majority of the ~14,000 nuclei segmented in the image region. Consistent with the cell line results, some spots were also detected outside the nuclei (data not shown). However, overall spot counts per cell on average

were lower than in cell lines (**Figure 32 B**). Because the raw ALT-FISH signals across a large image area are challenging to interpret visually, I devised a visualization technique that may be applied for more intuitive examination of heterogeneity in tissue samples. It generates color-coded ALT activity maps by projecting pre-defined groupings of ALT activity back onto the spatial cell location in the tissue. Such an ALT activity map representation for tumor L1 is shown in **Figure 32 C**. Four categories of ALT activity (ALT-negative, 0 spots; Low-ALT, 1-2 spots; ALT, 3-20 spots; super-ALT >20 spots) were defined based on the nuclear spot counts (**Figure 32 B**). The previously identified cutoff used to distinguish ALT-positive from ALT-negative cells in cell mixtures (>2 spots, compare *Results section 2.7*) guided the classification of the first three categories. Cells having a spot count greater than 20 were defined as super-ALT cells. The ALT activity map uncovered segregated regions of low and high ALT activity in tumor L1 (**Figure 32 C**), located in the low and high cell density areas, respectively. The matching pseudo H&E data shown in **Figure 32 A** confirmed that the two areas also displayed distinct morphology. A nearest neighbor analysis of the groups (**Figure 32 C**, rightmost panel) revealed preferential clustering of cells with similar ALT activity levels, matching the spatial patterns that had been observed to some extent in the previous datasets from cell lines.

In conclusion, the analysis of LMS tumor L1 demonstrated that ALT-FISH staining on sections from tumor material was feasible and could be coupled to additional imaging readouts. The spatially resolved view on ALT activity uncovered significant heterogeneity in this tumor with respect to the amount and local distribution of SSTR compartments that reflected differential engagement in the ALT pathway.

4.3 Spatial mapping of ALT activity in leiomyosarcoma and neuroblastoma tumors

The staining conducted on a single tumor sample served its purpose to confirm the presence of SSTR compartments in ALT-positive primary tumor tissue. Next, I wanted to validate the specificity and sensitivity of the assay using a representative sample set of primary tumors. To this end, the LMS set was expanded by eight additional LMS tumors L2-L9 (four ALT-positive, four ALT-negative) and one matched normal tissue control from the same cohort described in ref. [107] (**Figure 33**). A second section of LMS tumor L1 (L1.2) was included as technical control. I furthermore included seven neuroblastoma (NB) samples N1-N7 (three ALT-positive, four ALT-negative) to cover an additional tumor entity (**Figure 34 A**). The NB samples were obtained through collaboration with the group of Frank Westermann at the DKFZ and had undergone a rigorous molecular and omics-based characterization, which is described in ref. [109]. Similar to the LMS samples, two replicates were stained for one ALT-positive NB tumor (N2.1, N2.2) as a technical control. All LMS and NB samples and their relevant associated

clinical and mutational data are listed in ref. [165]. ALT status had been assigned by positivity for the CC assay. For unbiased sampling of regions for ALT-FISH imaging, an area of each tumor tissue with ideally 1 mm² in size and if possible located at the section center, was selected for imaging. In total, ALT-FISH spots in ~87,000 (LMS) to ~106,000 (NB) individual cell nuclei across 20 tissue sections were quantified (**Figure 33** and **Figure 34 A**). ALT activity maps were generated using the grouping strategy previously established on tumor L1 (**Figure 32 A**). For each sample, also the fraction of cells allocated to each of the four ALT activity groups (ALT-, Low-ALT, ALT, and super-ALT) was determined (**Figure 33** and **Figure 34 A**, stacked bar graphs next to the maps). Analysis of the complete dataset revealed generally lower spot counts in tissues as compared to cell lines. Nevertheless, when LMS and NB samples were combined, sections from ALT-positive tumors showed a 16-fold higher average count per nucleus (1.21; 95% CI: 0.66–1.75) than did those from ALT-negative tumors and normal tissue (0.077; 95% CI: 0.016–0.14) (**Figure 34 B**).

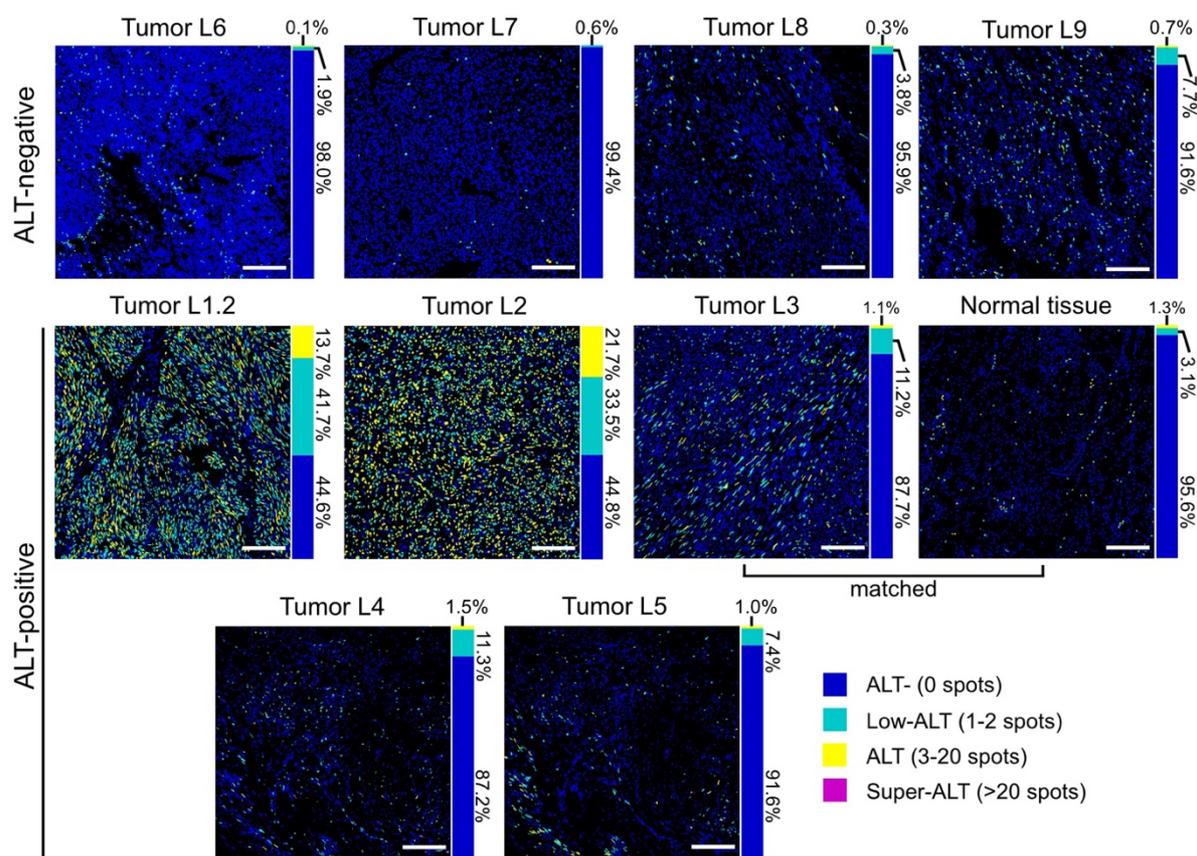


Figure 33. ALT-FISH analysis of leiomyosarcoma (LMS) tumors. ALT activity maps generated from the ALT-FISH signals and nuclear segmentation of four ALT-negative and five ALT-positive LMS tumor samples and one matched normal tissue control (~1x1mm regions). Note that tumor L1.2 corresponds to an additional section of the tumor analyzed in Figure 32. Scale bars, 200 μ m. ALT status is based on the CC assay analysis reported in ref. [107]. Supplementary tumor information is provided in ref. [165]. Figure panels adapted from ref. [165].

When comparing the ALT-positive to ALT-negative cases within each entity, the mean spot number per nucleus was still significantly different (LMS: P value = 0.047, NB: P value = 0.002, unpaired Student's t test), indicating a reliable separation by the spot count data alone. Next, I used the normal tissue matched to LMS tumor L3 (Figure 33, middle row), for which absence of ALT activity was expected, to designate criteria for the classification of bulk ALT tumor status that were informed by the ALT-FISH signal. The normal tissue had an average nuclear spot count of 0.1 spots/nucleus. As 95.6% of the cells showed no spots, the few spots seen in the remaining <5% of the cells were considered unspecific background signal.

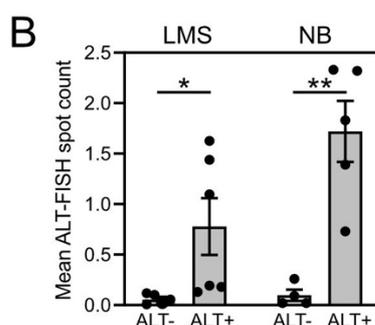
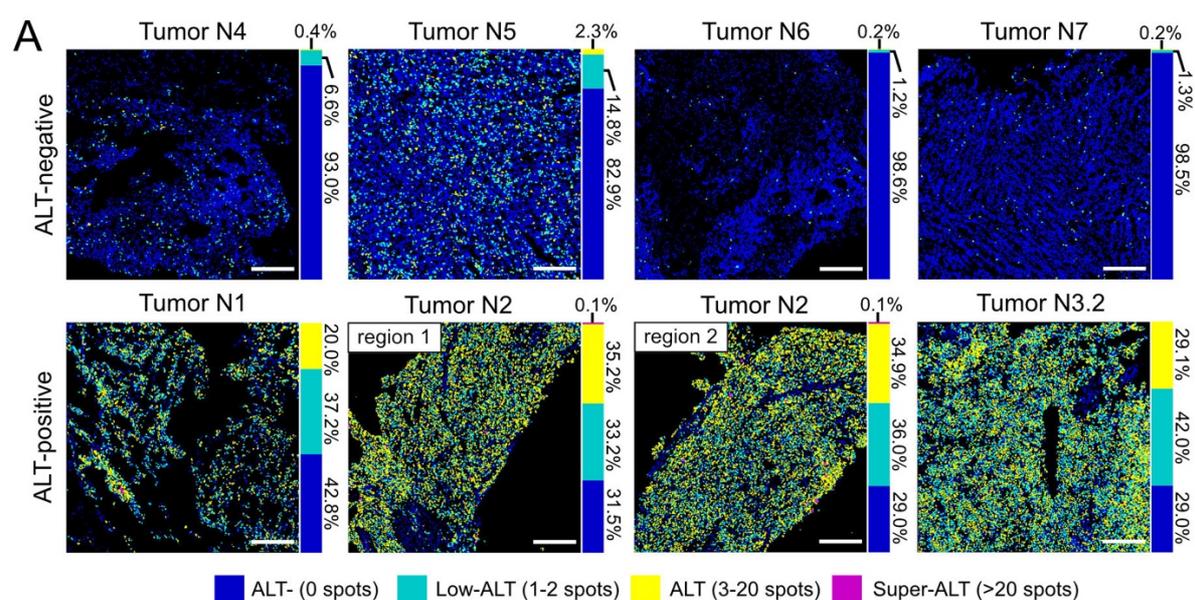


Figure 34. ALT-FISH analysis of neuroblastoma (NB) tumors and cross-entity comparison. (A) ALT activity maps generated from the ALT-FISH signals and nuclear segmentation of four ALT-negative and three ALT-positive NB tumor samples (~1x1mm regions). Note that two adjacent regions from the same tissue section were recorded for Tumor N2. Scale bars, 200 μ m. ALT status is based on the CC assay analysis reported in ref [109]. Supplementary tumor information is provided in ref. [165]. (B) Mean ALT-FISH spot count and standard error across all LMS and NB samples with individual values shown and grouped by ALT status. Unpaired Student's t test: *P = 0.047, **P = 0.002. The normal tissue sample was included in the ALT-negative group. Figure panels adapted from ref. [165].

Based on these features, I tested two independent classification strategies. In the first, the cutoff for ALT positivity was defined as the mean spot count per nucleus being >1.5 times above the mean spot count of the normal tissue. With this cutoff, the assay's sensitivity and specificity were 0.91 and 0.89, respectively (using the CC data as ground truth). As a second strategy, the threshold to ALT positivity was declared to >5% of the cells containing at least one spot. Using this approach, sensitivity and specificity were 1.00 and 0.75, respectively. From comparing the two methods, I deduced that the ALT-FISH assay had excellent overall specificity and sensitivity for classifying bulk ALT status. However, the analysis also highlighted the significance of using normal tissue for reliable estimation of the background staining when screening unknown samples.

Next, I examined the inter-and intra-tumor heterogeneity of ALT activity in the two entities. First, the percentage of cells in the four ALT activity groups in both the NB and LMS tissue sections stained in duplicates were similar, demonstrating a high technical reproducibility. Second, tumors from both entities displayed significant heterogeneity with respect to their ALT activity profiles. For the LMS samples this was particularly evident. ALT-negative LMS tumors L6-L9 were comparable to the normal tissue, while LMS tumors classified as ALT-positive displayed a range of SSTR compartment abundances (**Figure 33**). Compared to LMS, the ALT-positive NB tumors displayed higher ALT activity and less inter-tumor heterogeneity (**Figure 34 A**). However, the sections from tumors N2 and N3 contained clearly discernable areas of low or absent ALT activity, similar to those previously seen in LMS tumor L1 (compare **Figure 32 C**).

In general, a strong signal on the CC assay blots was correlated to higher average ALT-FISH counts as inferred from visual inspection of the images in the original studies. More CC signal and higher spot counts were matched to samples harboring ATRX mutations, corroborating the previous relationship seen between ALT-FISH signals and ATRX expression. NB tumor N5 stood out among the CC-negative tumors, since it contained 14.8% of nuclei with 1-2 and 2.3% of nuclei with >2 spots (**Figure 34**, top row). This could be interpreted as the co-existence of ALT and telomerase or ALT and no TMM in this specific tumor. While cases like tumor N5 are likely to be missed by standard bulk ALT readouts, they are uncovered by the single cell resolution of the ALT-FISH assay. Detection of such cases is particularly valuable for understanding the emergence of ALT and ultimately provide a more fine-grained stratification of patients.

4.4 Combining ALT-FISH with a tumor marker to distinguish between the heterogeneity of tumor cells and non-malignant infiltrations

The ALT-FISH analysis conducted on tumor samples from two representative entities showed that in some cases there were clear boundaries between tissue regions of low and high ALT activity. This would be explained by two co-existing and spatially segregated tumor cell populations that employ different TMMs. However, tumors can contain a significant amount of non-malignant cells, such as those found in the tumor-associated vasculature and the immune microenvironment. In fact, the studies that had previously characterized the specific tumors studied here, had indicated variable tumor purity, ranging between 55 and 100% for the LMS samples and being greater than 60% for the NB samples [107, 109]. To assess the contribution of non-malignant cells and demonstrate that ALT-FISH signals were in fact highly selective for tumor cells, I conducted ALT-FISH on an additional section of the heterogeneous NB tumor N3 (N3.1) (compare **Figure 34**, bottom row) and combined it with immuno-staining for the well-established neuroblastoma tumor cell marker NCAM [209, 210] (**Figure 35 A**).

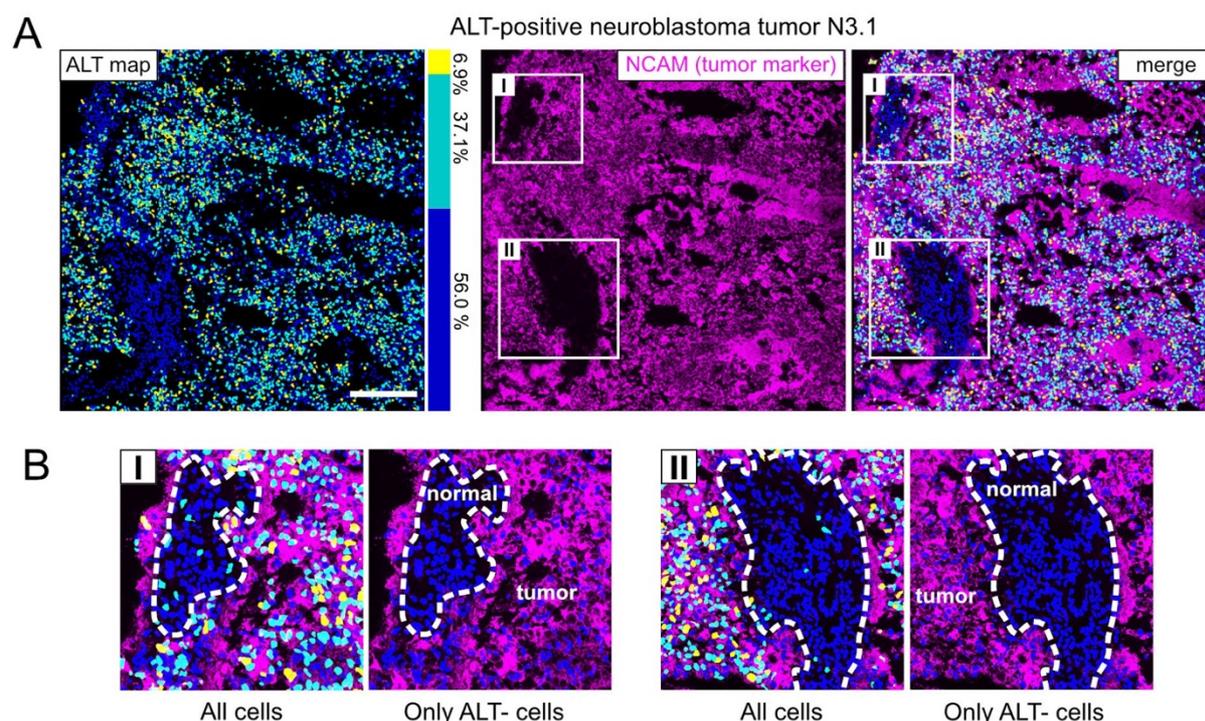


Figure 35. Combined ALT-FISH and tumor marker co-staining of neuroblastoma tissue. (A) ALT activity map of a selected region of an ALT-positive neuroblastoma tissue sample. Immuno-staining against the tumor cell marker NCAM is shown in magenta and as merge with the map. Note that sample N3.1 shown here and sample N3.2 in Figure 34 are regions from two independent sections of the same tumor. Scale bar, 200 μ m. (B) Two selected regions of N3.1 with heterogeneous ALT activity (I,II): The lack of NCAM signal identified the subregion lacking ALT activity as infiltrating non-malignant cells (dotted outline). Figure adapted from ref. [165].

The ALT-FISH signal and the tumor marker had a strong correlation. In turn, nearly all of the non-tumor cell nuclei identified by NCAM staining were categorized as ALT-negative (**Figure 35 B**, region "normal"). This result clearly corroborated the specificity of ALT-FISH and was in agreement with the previously observed lack of SSTR compartments in non-malignant HUVEC cells (compare *Results section 2.8*). The mean spot count per nucleus over all the NCAM-negative areas in section N3.1 was furthermore comparable to the mean spot count quantified for the normal tissue sample of the LMS sample set (~0.1 spots/nucleus). Notably, cell nuclei from all ALT activity groups were present in NCAM-positive areas (**Figure 35 B**, region "tumor"), suggesting that part of the observed heterogeneity was in fact ALT activity variation between individual tumor cells. In conclusion, improved heterogeneity assessment of this NB tumor was provided by identification of non-malignant tissue infiltrations with simultaneous measurement of ALT activity. For the most complete ALT-FISH based profiling of tumors, it is therefore advisable to combine the assay with a tumor marker immuno-staining or any other readout that will reliably identify non-malignant infiltrations.

4.5 Single cell transcriptome information and ALT activity obtained from the same cell in intact neuroblastoma tumor tissue

The previous results showed that ALT-FISH can measure ALT activity of single cells in tumors, even in combination with readouts of tissue morphology or tumor cell identity. However, to identify ALT-specific cell traits, this readout needs to be paired with more comprehensive measures of cell phenotypes. Recent advances have made it possible to probe transcriptomes of single cells in their spatial context [211]. Like ALT-FISH, many of these methods use imaging-based detection of transcripts by hybridizing specific probes *in situ*. As a proof-of-concept in this thesis, I thus explored the technical feasibility of simultaneously detecting ALT activity and single cell transcriptomes in tumors (**Figure 36**).

To this end, spatially resolved expression of 100 genes in a single section of an ALT-positive neuroblastoma tumor was obtained using the Molecular Cartography platform from Resolve Biosciences. It uses multiplex single molecule FISH in sequential rounds of hybridization to map the expression of 100 genes that were selected to represent the relevant cell types based on a preceding single cell RNA sequencing analysis [212]. The transcript detection runs were conducted on a specialized instrument from Resolve Biosciences. Automated data acquisition was conducted in the DKFZ Single-Cell Open Lab. It involves automated rounds of hybridization and imaging to decode the location of barcoded gene-specific probes (**Figure 36 A**). To incorporate ALT-FISH, the TelC probe was co-hybridized with the gene-specific probes before the decoding rounds. After the run, I recorded ALT-FISH and DAPI signal using the previously described confocal setup and applied the established workflows to segment

nuclei, detect spots and generate ALT activity maps. Linking the transcript data to the confocal images required custom image registration, which was both established and carried out by Anne Rademacher from our group.

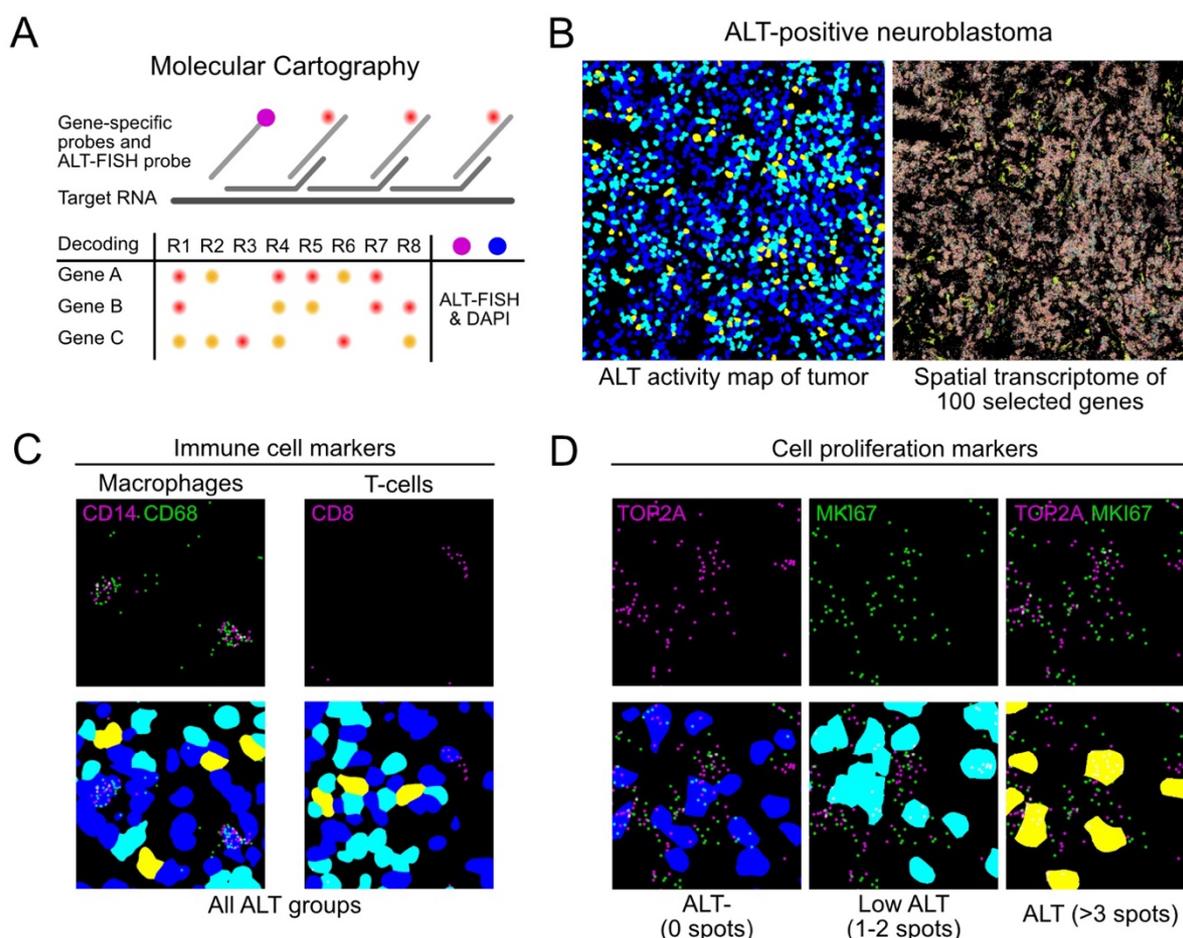


Figure 36. Integration of ALT-FISH and spatial transcriptome detection in single cells. (A) Spatial transcript detection by Molecular Cartography (MC) from Resolve Biosciences together with ALT-FISH. The TelC ALT-FISH probe and gene-specific MC probes with barcoded overhangs are hybridized to their targets (transcripts, SSTRs). Eight rounds (R1-R8) of fluorescent barcode detection by probe hybridization and removal, decode transcript identities and positions. ALT-FISH and DAPI are acquired at the end and after transcript decoding. (B) Representative region of the ALT-positive neuroblastoma tumor tissue section used for the MC/ALT-FISH integration. ALT activity map (grouping shown in D) is shown next to an overlay of all transcripts detected by MC in this region. (C) Detected macrophage and T-cell marker transcripts from MC (top) overlaid with the corresponding ALT activity map (bottom). (D) Proliferation marker transcripts TOP2A and MKI67 (top) overlaid with the corresponding ALT activity map (bottom), with cells from the three ALT activity groups displayed separately.

Single cell ALT activity could be combined with the transcript locations of the 100 target genes by the integrated ALT-FISH and Molecular Cartography approach devised here. A representative region of a tumor section is shown in **Figure 36 B**. The gene panel included immune cell and proliferation markers, as well as specific genes that represented specific neuroblastoma features (not analyzed here). As exemplarily shown for macrophages and

CD8-positive T-cells, the integration of ALT-FISH and Molecular Cartography was able to detect lowly abundant immune cells and relate them to the ALT activity pattern of the surrounding tumor cells (**Figure 36 C**). Likewise, the cell proliferation markers TOP2A and MKI67 were detectable in cells of different ALT activity groups (**Figure 36 D**). This integrative approach opens numerous possibilities to relate transcriptome profiles and ALT activity in tumor tissue sections. It will allow to resolve and characterize tumor cell states associated with no, low or high ALT activity and will reveal differences in the interactions with immune cells that are related to telomere maintenance by ALT.

III. Discussion

The compartmentalization of the cell nucleus is a fundamental aspect of cellular organization and function. However, despite significant research advances, several critical questions remain unanswered: What is the mechanism that drives compartment formation at specific sites in the genome? How are compartment boundaries maintained to concentrate or exclude biological activities? Is the role of mesoscale structural rearrangements a cause or a consequence of biological function? Multivalent interactions and phase separation of factors have been proposed as mechanisms for creating steep concentration boundaries in the nucleus, but how is functional specificity maintained if phase separation drives non-stoichiometric enrichment of factors around chromatin? Addressing these questions is critical to reveal the relationships between altered nuclear organization and diseased cell states to ultimately target detrimental biological processes that arise from aberrant compartment assembly and chromatin organization.

This thesis set out to address some of these key questions by developing experimental tools and functional readouts for the mechanistic dissection and detection of nuclear compartments by imaging. To address shortcomings of prior approaches, I focused on developing tools and cellular models that can investigate the functional interplay between processes such as phase separation, transcription, and chromatin changes in intact cells with improved control over the underlying molecular features. In addition, I exploited the detection of compartments for measuring disease-related biological activities in a high throughput fashion. Specifically, a method was developed to microscopically detect activity of the ALT pathway in cancer cells, for which no scalable single cell readout was available so far.

Technology development was conducted in three areas (**Figure 37**): (i) A toolbox of chromatin effector constructs with fluorescent tags, tunable binding and activation properties was created to enable flexible perturbation and imaging of compartments formed around silenced and transcriptionally active chromatin domains. (ii) Several features reflecting the altered abundance, processing, and transcription of telomeric repeats in cancer cells were integrated to develop the ALT-FISH assay. It maps compartments of single-stranded telomeric repeats (SSTRs) found almost exclusively in ALT-positive cancer cells. (iii) Quantitative and automated image analysis workflows were devised to measure compartment characteristics, activity, and abundance across cells.

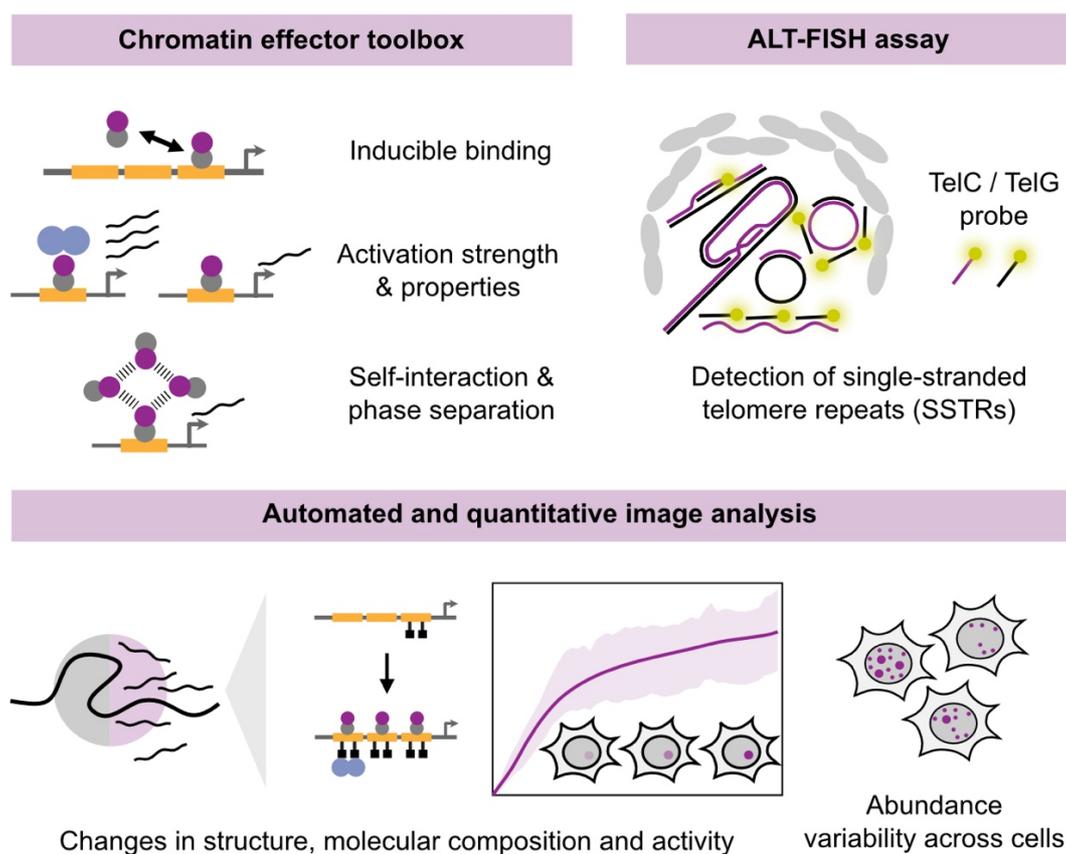


Figure 37. Overview of the developed methods used to study nuclear compartments.

These approaches to perturb and characterize nuclear organization were applied to three different types of nuclear compartments (**Figure 38**): (i) The transition of mouse pericentric heterochromatin (PCH) from a compacted to a relaxed state was studied by targeting an activator to major satellite repeats (MSRs). It revealed a switch-like transition of the compartment that was independent of the H3K9me3-HP1-SUV39H axis and linked to deposition of H3K27ac marks. (ii) By employing various activation domains (ADs) and construct topologies to target chromatin binding sites, distinct functions such as H3K27ac deposition, transcriptional activity to different levels, and BRD4 co-activator enrichment could be modulated on an inducible reporter gene cluster. Differential reliance on the H3K27ac-BRD4 co-activation pathway was found for different ADs. Using multivalent optogenetic protein domains, the formation, composition, and activity of transcriptional activator condensates was probed in living cells. The resulting observations challenged the notion that TF condensates generally boost transcription and underscored the significance of condensate composition in establishing transcriptionally active nuclear compartments. (iii) Detection of compartments defined by accumulation of SSTRs up to several kb in length was established for various sample types that ranged from cell lines in different culture formats to patient tumor samples. These structures emerge during the ALT process, and were found to spatially link to

ALT-associated PML bodies (APBs). The regulation and variability of these compartments as indicator of ALT activity was investigated in cancer cell lines with varying TMM statuses and from different tumor entities, as well as in models of telomerase-to-ALT switching and ATRX deficiency. The results proved sensitive, specific, and quantitative measurement of ALT activity, comparable to the widely used CC assay. Application of the ALT-FISH assay in a high-throughput screening format demonstrated its versatility and identified cell cycle, ploidy and spatial context as factors that link to ALT heterogeneity in cell lines. Finally, it was shown that SSTR compartment detection by ALT-FISH is applicable to primary tumor tissue specimens from leiomyosarcoma and neuroblastoma patients, enabling precise differentiation of ALT-positive and ALT-negative cases, as well as spatial resolution of TMM heterogeneity. The assay's detection principle was furthermore multiplexable with different standard or more advanced microscopic readouts.

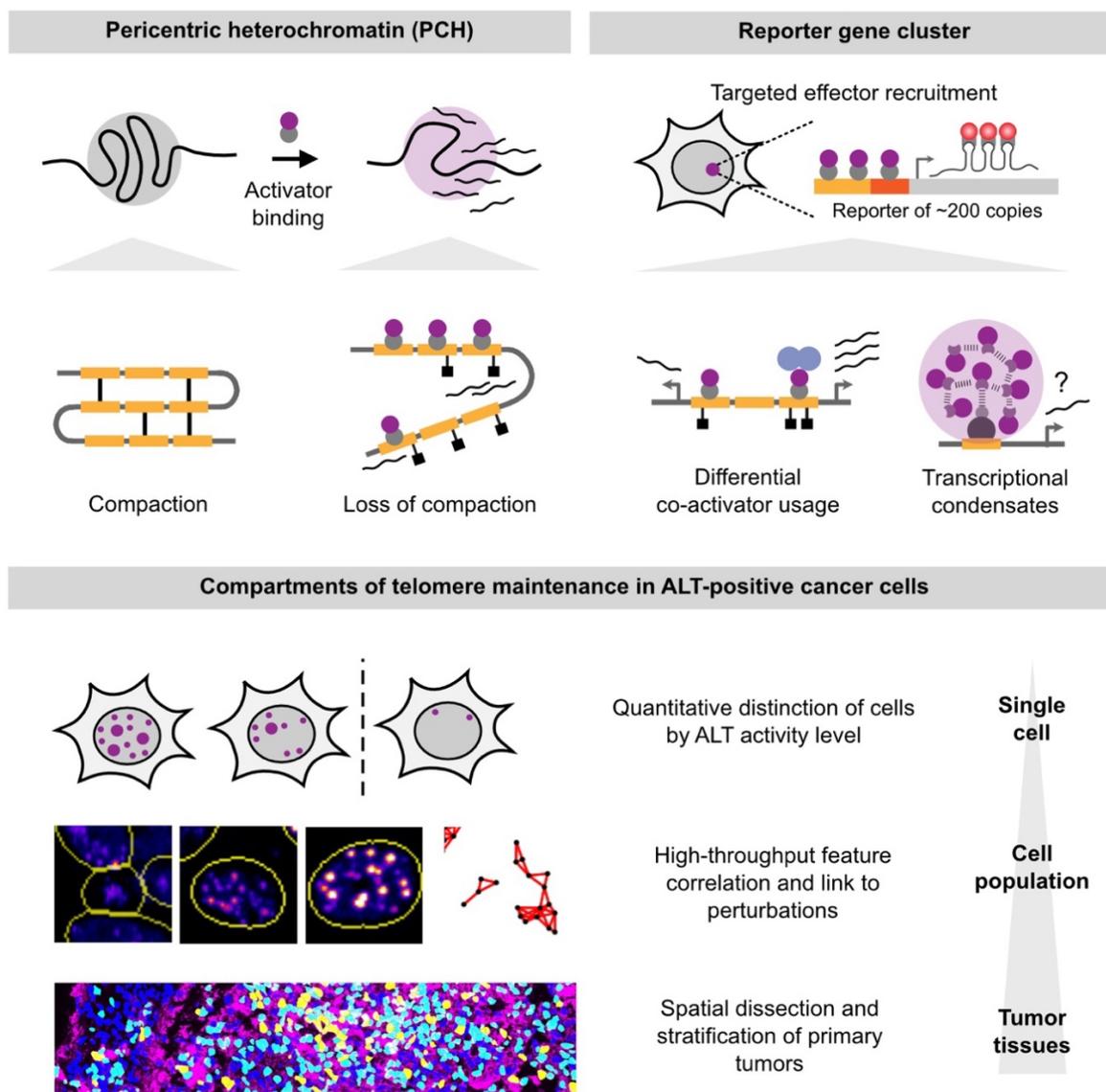


Figure 38. Overview of the nuclear compartments and their specific aspects studied here.

As discussed below, the findings of this thesis challenge various aspects of recently proposed mechanisms that invoke phase separation to describe structure-function relationships of compartments in the cell nucleus. The results obtained by the toolbox perturbations are specifically discussed regarding their contribution to our mechanistic understanding of compartment transitions, co-factor interactions and phase separation in transcription. To evaluate the ALT-FISH method, its strengths and limitations are compared to other ALT detection methods. Furthermore, an integrated description of the mechanistic conclusions on the regulation of ALT derived from the data is developed. The final part of the discussion focuses on the potential of using compartment perturbation and detection techniques to describe cellular phenotypes and nuclear functions by imaging beyond the here demonstrated cases.

1. Probing nuclear compartment transitions, co-activator interaction and phase separation by perturbation and imaging

Two model systems, silenced mouse PCH and an inducible reporter gene cluster, were studied by the perturbation strategies devised here. For mouse PCH, it was found that targeting the strong activator VPR to MSR was sufficient to cause its large-scale reorganization into a de-compacted and highly H3K27-acetylated state. Low levels of bound activator were sufficient to induce a substantial reorganization. Chromocenters that lacked H3K9me3 and HP1 due to *Suv39h* deficiency responded equally drastic to the activator, implying that these constitutive heterochromatin factors have little effect on the transition process. This switch-like domain transition is consistent with a model of a collapsed polymer globule that loses cross-linking interactions between its segments [213]. Histone acetylation can reduce bridging interactions and relax chromatin compaction [62, 214]. It is therefore likely the main driver of the de-compaction observed in the cell system studied here. VPR-induced acetylation is most probably brought along by its subcomponents p65 and VP64, both of which interact with various histone acetyl-transferases [152, 159].

The modulation of PCH compaction via the acetylation level is in line with the previously seen, yet milder, chromocenter reorganization phenotypes upon global inhibition of histone deacetylases in murine cells [215, 216]. Changes in transcription were only assessed for the reporter gene cluster. However, the enrichment of elongating CTD-Ser5 phosphorylated form of RNA Pol II at VPR-perturbed chromocenters is a strong indicator of their transcription [57]. A better understanding of MSR transcription will be gained by combining the here established methods with FISH readouts to detect MSR RNA, which is inherently difficult to study at endogenous levels. To which extent H3K9me3, HP1 and other heterochromatin marks persist

upon chromocenter de-compaction was not assessed. Maintenance of heterochromatic features after activation is interesting to follow, because they could serve as a cue to re-establish the silenced state after the activating stimulus has ceased. The inducible toolbox developed here makes it possible to toggle between the active and silenced state, which will help to further dissect the underlying mechanisms. Moreover, it will be informative to recruit less potent and more specific activators, or histone acetyl-transferases (HATs) and methyl-transferases, to elucidate the precise factors that are sufficient and required for triggering chromocenter de-compaction. Gaining insight into the sequence of events during de-compaction has significant implications for understanding how heterochromatin defends itself against unlicensed activator attacks and arranges its mesoscale configuration in the nucleus.

The perturbation and microscopic analysis performed on chromocenters were employed in a conceptually similar manner to investigate the activities of an inducible reporter gene cluster as a model of a large, transcriptionally active chromatin domain. Transcription, H3K27ac, and BRD4 enrichment could be induced at the reporter in a temporally controlled manner with synthetic activators and even in isolation from each other. This was made possible by comparing different constructs that differed in size and turnover of the chromatin bound state and interaction affinity with co-activators like HATs and BRD4. Thus, the toolbox developed in this thesis allowed it to differentiate interactions that were dependent on transcription from those that were independent of it as well as identifying the factors that govern activation strength. The self-interaction propensity of the optogenetic PHR domain fused to ADs was exploited to artificially control phase separation of chromatin effector proteins in living cells. The objective was to generate assemblies of activators on the reporter gene cluster that resemble transcriptional condensates, and to examine their activity in comparison to assemblies that are not formed through a phase separation mechanism.

The difference in activation strength between VP16 and VPR was found to rely, at least in part, on the type of interactions that these activators establish with HATs and BRD4. A reduced ability of VP16 to induce H3K27ac at the reporter was found. By conducting live cell measurements of the interaction between BRD4 and the reporter gene cluster in the presence of VP16 or VPR, it was discovered that a pool of BRD4 interacted with VPR directly in a manner that was not dependent on its bromodomain. The low level of BRD4 localization induced by VP16 was bromodomain-dependent. Notably, transcription was not required for inducing the accumulation of BRD4 and H3K27ac marks. Although not examined in this thesis, the direct interaction between BRD4 and VPR may be due to multivalent interactions between the large C-terminal IDR of BRD4, as previously observed for other TFs [42]. It is possible that the strong ability of VPR to form PHR droplets through self-interaction may also aid in enabling this type of interaction. However, bromodomain-dependent interactions were also observed

and demonstrated to influence transcriptional activity. The latter could be due to the binding of the BRD4 bromodomains to the lysine 310-acetylated variant of p65 [217]. This complex is known to facilitate transcription by recruiting additional CDK9 to phosphorylate RNA Pol II CTD. Thus, it may also serve as a potent mediator of VPR's ability to establish and maintain a transcriptionally active chromatin state. VP16 but not VPR activity could be enhanced by artificially pre-depositing histone acetylation at the reporter via a dCas9 p300 acetyltransferase fusion construct. Thus, one possible for VPR's robustness as an activator could be its ability to efficiently induce histone- and self-acetylation and thereby partially circumventing the requirement for acetylation-based BRD4 interaction to boost transcription. This view is also supported by the capability of VPR to stimulate histone acetylation in constitutive heterochromatin, as seen from the chromocenter experiments. More experiments, for example, inhibition of HAT activity or other co-activation pathways, are needed to isolate the events that co-occur during the transition of chromatin compartments from a silenced to an activated state. Moreover, methods for the time-resolved observation of protein interactions on chromatin need to be advanced to resolve the dynamic protein co-association occurring during transcription.

As introduced earlier, phase separation has been proposed as a mechanism for the formation and function of both heterochromatic and transcriptionally active compartments in the nucleus [38, 41, 157]. Phase separation does not consume energy, and partitions cell components into dense and dilute phases above a critical concentration. The volume of the phases is proportional to the overall concentration, and other molecules can selectively dissolve in them based on their chemical properties [26]. Chromatin binding is not necessary for maintaining condensates, and they may form in regions with low chromatin density or associate with chromatin via surface condensation [218]. The physical state and function of nuclear condensates is a subject of debate. Recently applied *in vivo* methodologies, such as measuring viscosity and internal mixing [57, 219] or quantifying assembly dynamics with single molecule precision [74], address the growing needs for a better biophysical definition.

Due to their high IDR content, which can promote self-interactions, many TFs, components of the general transcriptional machinery and co-activators have been proposed to undergo phase transitions in the cell. But measuring the biological relevance of a phase transition for a given biological process in cells, has been a challenge. Studies that compare the activity of TFs with versus without its fusion to a foreign IDR to induce condensate formation, do not account for the diverse functions of IDRs beyond phase separation. In addition, over-expression of endogenous TFs or the way by which phase-separated droplets are detected can confound interpretation of the underlying relationships [74, 220]. This problem was addressed in the present thesis in a bottom-up approach with a set of synthetic transcription factors, whose

association into condensates on the reporter gene cluster could be controlled by light and the PHR domain. Five different PHR-AD fusions (STAT2, VP16, p65, Rta, VPR) were expressed and anchored to the *lacO* or *tetO* sites of the reporter gene cluster via CIBN [65] to induce its transcription. Imaging revealed that anchored PHR-AD assemblies in cells above the critical concentration for droplet formation (C_{crit}) were larger and organized into subdomains that could engage in fusions among each other and with unbound droplets. Additional experiments, not included in this thesis, were conducted to further analyze these assemblies by dSTORM super resolution microscopy and time-lapse imaging [146]. They supported a phase-separated state of PHR-ADs at the reporter clusters, consistent with the previously reported characteristics of PHR optodroplets [78].

To examine how the formation of condensates would affect transcription at the reporter gene cluster, three different types of experiments were conducted. First, transcriptional activity of the five studied ADs was evaluated separately for two groups: "droplet formers" (VPR, p65, Rta) and "non-droplet formers" (VP16, STAT2), which were distinguished based on their critical concentration. The critical concentration was determined by analyzing the relationship between droplet abundance and concentration. This analysis revealed a positive correlation between stronger and faster activation and higher phase separation propensity (or low C_{crit}). Second, by utilizing the single cell resolution of the established imaging assays, cells expressing the same PHR-AD construct, were classified as containing droplets or not. The classification showed that droplet formation *per se* did have a neutral effect on transcription across all of the droplet-forming activators VPR, p65 and Rta. Third, synthetic multivalent linkers were designed to modulate the formation threshold and composition of PHR-VP16 droplets. Here, two out of three cases showed inhibition of transcription, while attaching the N-terminal IDR of FUS [79] robustly enhanced transcription, in agreement with previous reports on the effect of FUS [217, 221]. However, and in line with the results obtained for the other activators, the FUSN-VP16 droplet formation *per se* did not affect transcription.

In summary, the findings made here, challenge the notion that transcriptional condensates act as general amplification hubs for gene expression. Simply accumulating more TF molecules by a phase separation processes, was insufficient to enhance transcription. Instead, the specific droplet composition was critical. Inhibition of transcription, as observed for the VP16 droplets induced by LacI and GBP multivalent linkers, could likely be caused by changes in droplet-internal mobility [222]. A more solid (cross-linked) versus a liquid-like (mobile) state can potentially constrain the movement of the activator and thereby limit its efficient interaction with the core promoter, and/or impede the dynamic interaction with co-activators. An alternative explanation is provided by changes in chemical properties introduced by the linkers (e.g. changes in charge patterns [223]). In turn, these altered chemical environments could

influence the solubility of molecules that are critical for the activation process. A selective exclusion of general factors, such as RNA Pol II, was not found for the inhibitory droplets studied here (see ref. [146]). But other transcription-related factors may be de-enriched and explain the inhibition. Another limitation of the approaches used here is that ectopic expression of PHR-ADs fusions was used to model transcriptional condensates on chromatin. It may not recapitulate the properties of full-length TFs expressed at endogenous levels and binding to endogenous target sites.

The observation that FUSN enhanced activation independent of droplet formation points to a crucial role of IDRs outside the phase separation regime, which has already been suggested by recent studies [33, 81]. Multivalent self-interaction of (IDR-containing) proteins on chromatin can fulfil other functions, such as facilitating interactions with co-activators or stabilizing TF residence time [29, 224, 225]. This could in part also explain how the increased self-interaction observed for the strong ADs studied in this thesis related to stronger activation, without the need of droplet formation. These aspects were addressed in a separate line of experiments, whose results were not included in this thesis, but are presented and discussed in ref. [146]. In brief, they uncovered the presence of an indirectly reporter-bound fraction of VPR but not VP16 molecules below C_{crit} that could possibly help in stabilizing TF residence time and buffer loss of TF molecules upon weakened binding to the promoter. A recent study came to very similar conclusions for transcriptional activity of endogenous oncogenic EWS-FLI1 TFs [81]. It is thus vital to distinguish the function of the complex of directly and indirectly bound TFs and co-activators (through multivalent self-interactions), which is determined by the stoichiometric binding of its components, and the non-stoichiometric accumulation of factors into condensates.

2. An integrative assay quantifies ALT activity by mapping compartments of single-stranded telomere repeats

Telomere extension by the alternative lengthening of telomeres pathway is active in many cancer entities [104]. Despite its importance and prevalence, there is limited knowledge regarding the heterogeneity of ALT activity in tumor cell populations. The investigation of tumor heterogeneity has become more accessible and quantitative through numerous advances in sequencing and imaging based single cell phenotyping [226]. However, for ALT cancer research to profit from these developments, ALT detection needs to be adapted to the growing needs for spatial resolution, robustness, and scalability. The current lack of robust detection assays with high-throughput *in situ* readouts hampers research in this area. To address this issue, ALT-FISH was developed that exploited the altered organization, abundance, and

extrachromosomal accumulation of single-stranded telomere repeats (SSTRs) into ALT-specific compartments. Prior to the development of ALT-FISH in this thesis, multiple independent studies had already reported that ALT-positive cancer cell lines of different origin and genetic background accumulated G-rich and C-rich single-stranded telomeric nucleic acids [115, 168, 173, 227, 228]. This marker had been disregarded for specific mechanistic studies, which is likely due to the fact that SSTRs arise ubiquitously along the ALT process. Instead, the well-established assays that detect C-circles (CCs) [120] and APBs [121] as single defined targets were preferentially used. The CC assay only measures bulk ALT activity in isolated genomic DNA. The APB assay has a limited dynamic range and requires image co-localization analysis. Moreover, a number of studies have challenged the ALT specificity of both assays. CCs were recently detected in telomerase-positive cancer cell lines with long telomeres [229] and their release was shown to occur as a consequence of telomere trimming [187]. Complete lack of APBs was found in liposarcomas that use ALT [230]. Reorganization of chromatin into large PML bodies with many features of APBs was observed in cell lines from ICF syndrome patients, independent of ALT [231]. Ultimately, PML-telomere co-localizations have been observed in some non-neoplastic human cells with no obvious ALT activation [232].

In contrast to these two assays, ALT-FISH represents an integrative approach of measuring multiple ALT hallmarks at once that overcomes several limitations of the previous ALT detection methods. ALT-FISH staining of cancer cell lines revealed that SSTRs organize into distinct nuclear and cytoplasmic compartments that can contain up to tens of kilobases of repeats. Their organization into foci is likely determined by localized production and/or their sequestration into pre-formed compartments. Sequestration into PML-NBs or containment within a PML protein shell assembled at the site of production are two putative explanations for their compartmentalization. They fit well to the previous observations that PML associates with ssDNA during DNA damage, replication stress or viral infections [233-235]. On average, nine out of ten SSTR compartments contained PML protein. Six out of ten SSTR spots co-localized with APBs, the number of which was previously determined to be 4-5 per cell [177]. This observation suggests that ALT-FISH co-detects most APBs under the conditions tested here. However, as the quantification of SSTR spots follows standard methods for spot detection on a single image channel, it provides a significant advantage over APB detection. In my thesis I could establish scalable and fully automated workflows to analyze ALT-FISH signals in hundreds up to hundreds of thousands of cells grown on coverslips or in microplate formats. The microplate data demonstrated that staining, data acquisition and analysis are easily adaptable for large screening setups. This potential has recently also been recognized by an independent research group that reported the use of native telomere FISH in multi-well plates to detect ALT [236].

Based on the results of this thesis, I found that the number of spots per cell nucleus above a defined intensity cutoff was generally a sufficient measure to distinguish ALT-positive from ALT-negative cells with high (~91-98%) accuracy and count differences were 4-12-fold higher as reported for APB numbers [177]. To initially determine a suitable spot detection threshold, it is recommended to perform cell-mixing experiments similar to those described in here. In addition, a side-by-side comparison of an ALT-positive and ALT-negative sample on the same microscopy setup was found to be important. For the microplate data, the cumulative spot intensity was introduced as an additional parameter to quantify ALT activity. It displayed an excellent correlation with the spot count in a near-linear relationship, which suggests that most of the ALT-FISH are similar in their intensity. It does, however, contain more information about the actual SSTR abundance and will capture changes in spot intensity upon unchanged spot number. For the proof-of-concept experiment that examined the effect of the two drugs Mirin and JQ1 on ALT activity, a comparison of both parameters revealed subtle changes in SSTR abundance could be uncovered when considering both parameters.

Similar results were obtained by staining with the TelC and the TelG ALT-FISH probe, suggesting their interchangeable use is possible in most cases. However, the TelC signal offered a larger dynamic range and detection of TERRA, which may be of advantage for specific biological questions. Although not explored to full extent here, cytoplasmic ECTR abundance can be quantified from ALT-FISH signals as well, provided a cytoplasm marker is available. Further experiments that explore this feature in combination with perturbations will be of value to better characterize the dynamics of the extrachromosomal SSTR pool in ALT-positive cells. Only little is known about ECTRs in single cells, but the immunogenicity [117] and inheritance patterns of extrachromosomal DNA has recently been recognized as relevant driver of tumor heterogeneity [237].

By integrative analysis of multiple independent samples collected along the whole study, it was found that ALT-FISH spot counts for both probes averaged across all cells of a sample correlated exquisitely well with bulk ALT activity measured by CC assay. Notably, it was surprising to see that the strength of the correlation was comparable between the two probes, despite the TelC probe not detecting CCs. It can be concluded that CC abundance seems to follow the abundance of other SSTR species, independent of their strandedness or source. It is in line with the idea that SSTR levels detected by ALT-FISH provide a more integrated view on overall ALT activity and are inter-correlated. By comparing different cell lines, validity of ALT detection by ALT-FISH could be confirmed for osteosarcoma, cervix carcinoma and pedGBM as cancer entities. Moreover, ALT-FISH signals largely followed CC levels in models of artificial ALT induction by ASF1 knock-down, ALT de-repression by ATRX knock-out and ALT repression by ectopic ATRX overexpression. The ubiquitous applicability of the assay was

furthermore corroborated on primary tumor samples from neuroblastoma and leiomyosarcoma patients of known ALT status. Patient stratification was possible with a sensitivity of 91% and specificity of 89% (using spot counts). These values are in principle acceptable for diagnostic purposes and were similar or higher than for ALT detection assays already used in a clinical setting, such as qPCR-based based telomere length measurement [238] and serum-based CC assay [239].

One open question that needs to be addressed in the future is how the signals measured by ALT-FISH relate to actual recombination-based telomere elongation. Side-by-side assessment of ALT activity and telomere length in inducible cell line models that undergo rapid changes in ALT activity [166, 181] will be crucial tools to clarify this relationship. The same applies to understanding the relative contributions of ECTRs (t-circles, CCs), TERRA, stalled replication forks and other chromosomal intermediates to the ALT-FISH signal. This can for instance be addressed by nuclease pre-treatments that degrade non-circular components, or strategies that induce replication stress, as recently demonstrated for G4-stabilizing drugs in cancer cell lines [166]. The targeting of pathways that regulate ECTR maintenance and excision, such as SMC5/6 complexes and TZAP [119, 240], will address the contribution of extrachromosomal SSTR species. The here established high-throughput approaches can be applied to investigate cues that trigger the formation or dissolution of SSTR compartments and will lead to a better understanding of where and how SSTRs arise during the ALT process and are related to telomere lengthening.

3. Resolving ALT activity in single cells in their spatial context

Bulk assignment of TMM status (ALT, TERT, none) to tumors or cell lines is common practice and needed for sample stratification. It is based on the joint assessment of various features that include CC levels, telomerase expression/activity, telomere repeat content, composition, TERT promoter mutations [241, 242], and mutational patterns [107]. Recurrent transcriptome and proteome signatures aid in assigning TMM activity, but are likely tumor entity specific [109] and need further validation by more cross entity studies. Assigning a single TMM to a sample disregards two important aspects: (i) the potential co-existence of telomere extension by ALT and telomerase in the same tumor cell population and (ii) different "activity levels" of the two TMMs across cells. However, resolving these aspects is crucial to describe and understand switching or bifurcation behaviors that may occur in evolving tumors in adaptation to therapeutic intervention, genetic or microenvironmental factors.

The quantitative criteria established by careful comparisons to appropriate controls of telomerase-positive cells or cells lacking TMM activity demonstrated that the ALT-FISH method developed here can precisely detect the ALT status of individual cells in cell lines and

primary tumor tissues. In the datasets studied here ALT heterogeneity was observed on two levels. First, heterogeneous (mosaic) samples with a low percentage (6-40%) of clearly ALT-positive cells were identified. Second, substantial variability in spot count and intensity was found within ALT-positive subpopulations or cell lines for which most cells (>95%) had been classified as ALT-positive. In some of the latter cases, the maximum ALT activity values reached up to >10 times the median of the population. Low to undetectable CC levels were observed in mosaic samples, but some displayed detectable bulk telomerase activity by TRAP assay. In the dataset analyzed here, SSTR content mosaicism was also seen upon artificial ALT induction in HeLa LT cells, in line with the original study that had reported a reduction, but not a full suppression of TERT mRNA expression and telomerase activity [181]. Since simultaneous activation of ALT and telomerase in the same cell is unlikely [243], these findings potentially support the use of ALT and telomerase (or no TMM) by distinct cell subpopulations. The results hereby provide the first quantitative single cell description of such a phenomenon, although previous studies have proposed it by observing the co-occurrence of ALT and telomerase signatures by bulk readouts [139-142]. Variable ALT activity was also found in tumors from neuroblastoma and especially leiomyosarcoma patients and could be explored through the generation of spatial ALT activity maps. It was furthermore demonstrated that ALT-FISH can be combined with technologies that spatially detect single cell transcriptomes by *in situ* hybridization. Further exploitation of such integrative solutions will yield unprecedented insights into the transcriptional programs that correlate with ALT activity and how the tissue microenvironment links to it. A more systematic exploration of a larger patient cohort will determine how these features link to treatment decision-relevant clinical parameters and unravel determinants of ALT regulation in tumors. In this context some limitations of the ALT-FISH assay are noted. For instance, cells lacking ALT-FISH signals may still engage in ALT-mediated telomere extension. SSTR levels could be delayed from the actual recombination event, or cells could go through periodic bursts of ALT activity, followed by a phase where no recombination is needed due to sufficiently long telomeres or a lack of cues to initiate it. These phases could strongly depend on the cycling time or other cell line specific traits like telomere length, giving the impression of TMM heterogeneity. To investigate such models in the future, it is important to combine ALT-FISH with spatially resolved single cell measurements of *de novo* DNA synthesis at telomeres. The co-existence of separate cell subpopulations using telomerase could additionally be probed by the integration of emerging microscopic assays that detect TERT mRNA expression or telomerase activity in single cells [244].

This thesis utilized microplate-based ALT-FISH analysis of hundreds of thousands of cells to investigate factors influencing the abundance of SSTR compartments. By inferring the cell cycle stage from the DNA content, it was found that the number and intensity of SSTR spots

increased towards G2/M phase, with even higher levels in polyploid cells. These findings are supported by previous studies linking ALT-related telomeric DNA synthesis to G2/M phase [110, 111, 245, 246] and an increased APB number in G2 [247]. Increased DNA content and higher absolute number of telomeres or telomere clustering in APBs may contribute to the observed differences, in addition to increased engagement in DNA synthesis. More SSTR exposure in polyploid cells may be due to elevated genotoxic stress. Alternatively, these could represent cells that fail to exit G2 due to massive telomere dysfunction or failed resolution of HR intermediates, which would also manifest in more SSTRs being exposed. Notably, the lowest SSTR abundance was seen for cells that had just entered G1 phase. The loss could be explained by multiple mechanisms, one being passive expulsion of extrachromosomal species upon nuclear envelope breakdown, or their degradation rates exceeding their production rate due to lack of HR activity in G1. Since PML bodies undergo significant changes during mitosis [248], cell cycle dependent containment or shielding from degradation in APBs or PML-NBs could also explain the variable number and size of SSTR foci. **Figure 39** summarizes a putative model explaining how differential SSTR foci patterns could arise in ALT-positive cancer cells and how this may relate to the formation of APBs and productive telomere elongation.

Using the spatial information from the microplate data, the influence of clonal origin and cell-cell contacts on SSTR variability was probed. By analyzing isolated microcolonies, I tested the hypothesis of SSTR inheritance between mother and daughter cells. No strong evidence was found to support this mechanism and the data rather suggested stochastic engagement in ALT activity that was independent of the mother cell's SSTR foci pattern. A model of stochastic ALT engagement has important implications for a tumor setting, where it could potentially amplify cellular heterogeneity to promote adaptation. In line, a rare population of cells with particularly many and intense SSTR spots (termed super-ALT cells) was seen across all ALT-positive cell lines, and in some of the studied leiomyosarcoma and neuroblastoma tumors. The biological significance of these outliers in tumor tissues awaits further characterization. The presence of these cells is similar to the phenomenon of "ultrabright foci". These large focal accumulations of telomeric nucleic acids were previously observed by denaturing telomere FISH on tumor tissues [108]. In the microplates, these cells were associated with a higher DNA content and spatial isolation. Isolation can be a consequence of a cell cycle arrest shortly after the time of cell seeding, which would fit well to the increased DNA content. Further characterizing the features of these cells and determining if they are still cycling, will be crucial to understand if they can contribute to tumor adaptability. For example, these cells could act as dormant cells with in a primed telomere elongation state, that could be reactivated when they start dividing again under certain conditions.

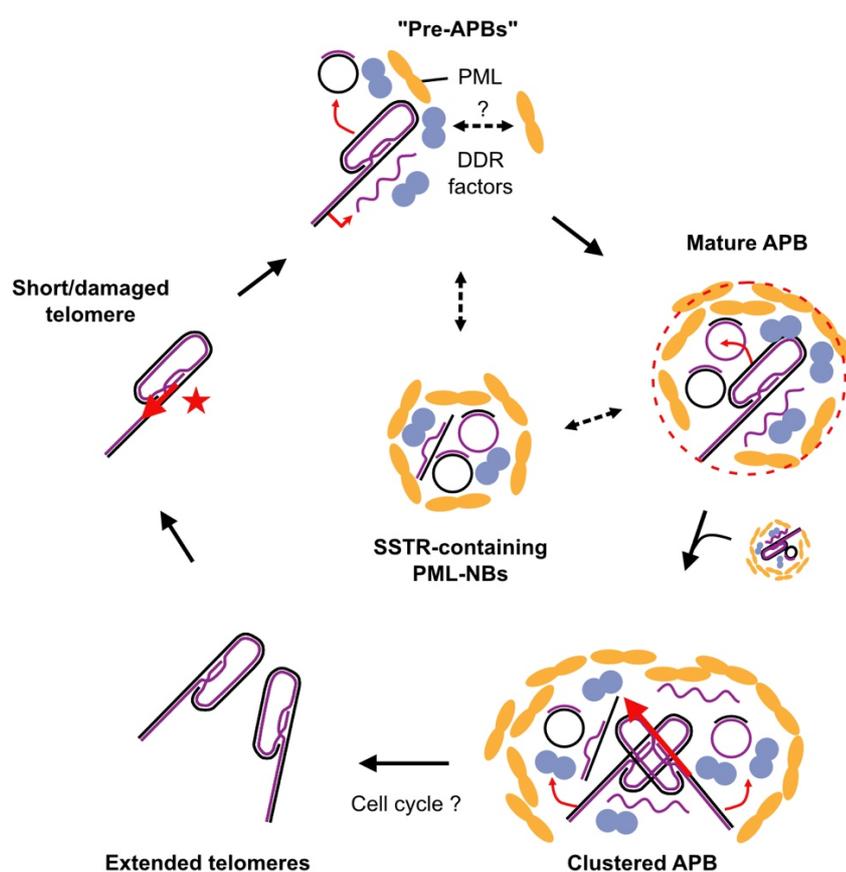


Figure 39. Model explaining how variability in SSTR foci patterns and abundance could arise and relate to ALT-mediated telomere elongation in APBs. Critically short and/or damaged telomeres expose chromosomal SSTRs and produce extrachromosomal SSTRs (ecSSTRs). SSTR exposure and ecSSTR release trigger PML assembly into a pre-APB compartment mediated by DDR factors. If the local SSTR concentration becomes sufficiently high, the pre-APB matures by accumulating more PML through self-interactions, forming a mature APB with a fully assembled and stable PML shell around the single telomere. During the assembly process ecSSTRs may be released and independently nucleate SSTR-containing PML-NBs devoid of telomeres. For productive telomere elongation to occur, two or more mature APBs need to cluster by PML-PML interactions. In the resulting clustered APB, telomeres can engage in HR and be extended through BIR-like processes. The resolution of recombination intermediates is critical for cell cycle progression or may be actively regulated to control cycling. APBs and SSTR-containing PML-NBs may disassemble if SSTR loss/degradation exceeds SSTR exposure/production or PML-NBs change their organization throughout the cell cycle. Successfully extended telomeres may re-enter the SSTR exposure cycle if they become critically short or damaged again.

4. Conclusions and Perspectives

The results of this thesis provide mechanistic insight into how nuclear compartments silence and activate transcription and shed light on the role of phase separation. In addition, it was demonstrated for the ALT pathway that compartmentalized activities within the nucleus can be exploited for large-scale phenotypic profiling by imaging, spanning the full range from cell lines to intact tumor tissue from patients. The targeted manipulation of chromatin marks, transcription and other chromatin processes by synthetic effectors is an independent field of research that has experienced several advancements in recent years. Combining these tools with (time-resolved) imaging- or sequencing-based measurements in single cells as done here but at increased throughput [249, 250], will be crucial to dissect the causal interrelationships of dynamic nuclear processes and how they vary between cells. Furthermore, devising tools that specifically modify material properties of nuclear components is valuable to probe their functional significance and draw a bigger picture of how compartments self-assemble, dissolve and regulate their activity. The common statement "*To fully understand a system, it must be built.*" also applies to nuclear compartments. Future work should therefore focus on utilizing bottom-up approaches to build compartments in the cell, starting from their individual components or induce them in an ectopic environment to probe the requirements for their formation and function. Detection of ALT-specific compartments by ALT-FISH has provided insights into the previously unaddressed heterogeneity of TMMs in cancer. The ability to quantify ALT activity in single cells paves the way for a more systematic and large-scale investigation of factors that regulate ALT. Looking forward, the application of the assay in genetic and drug screening approaches, will largely facilitate the identification of treatment-relevant factors that interfere with this pathway. The integration of ALT-FISH with other imaging readouts and single cell transcriptomic analyses on tumor tissue will lead to a more comprehensive understanding of the molecular and spatial heterogeneity and possibly aid clinical decision-making. Future research should thus aim to expand and refine cellular phenotyping beyond single-cell transcriptomics, epigenomics, and proteomics by exploiting recent advances in imaging technology to map subcellular structures at high throughput [249]. Profiling of nuclear organization can provide crucial insights into how gene activity is regulated, extrachromosomal elements are amplified, and chromatin reorganizes in physiological and diseased processes. By linking compartments formed by altered localization of cellular components to transcriptomic changes, we can improve our mechanistic models of diseases and potentially design more effective treatments. A challenge in using nuclear compartments for cellular phenotyping is accurately quantifying relevant activities and integrating this data with other types of cellular data. Meeting this challenge will require the development of new experimental and computational methods to integrate these large and diverse datasets.

IV. Materials and Methods

1. Materials

1.1 Plasmids

Plasmid constructs were generated by restriction enzyme based cloning. Fusion proteins were expressed under control of a CMV promoter in a pEGFP-C1/N1 backbone (Clontech). sgRNAs with (our without) PP7 loops were expressed in a U6-promoter driven expression cassette located on separate plasmids from dCas9 fusions. sgRNA targeting regions (5'-3') were *tetO* (GACTTTTCTCTATCACTGATA) *lacO* (GTCCGCTCACAATTCCACATG) and mouse major satellites/MSR (GGGCAAGAAAAGTGAATCA). All individual plasmids are described in detail in ref. [146] and made available through Addgene at www.addgene.org/Karsten_Rippe/.

1.2 Oligonucleotides and probes

The following primers (Eurofins Genomics, Germany) were used in qRT-PCR assays to detect beta-actin (ACTB) mRNA and U2OS-2-6-3 MS2 reporter RNA (5' to 3'): ACTB-fwd: TCCCTGGAGAAGAGCTACGA, ACTB-rev: AGCACTGTGTTGGCGTACAG, MS2-fwd: GTCCGGACTCAGATCTCGA, MS2-rev: TTCAAAGCTTGGACTGCAGG. The two ALT-FISH probes used in this study are DNA oligos (Eurofins Genomics, Germany) carrying a single 5'-terminal fluorescent dye label (ATTO 633 or Alexa 647) and a 3'-terminal biotin tag. The biotin tag was not used in the present study. Unless indicated otherwise, Atto 633-labeled probes were used. Alexa647-labeled probes were employed for the microplates and the Molecular Cartography integration experiment. The probes sequences were (5' to 3'): TelC: (CCCTAA)₅ and TelG: (TTAGGG)₅. For depletion of ASF1 in HeLa cells, On-Targetplus Smartpool siRNA pools (Dharmacon) were employed: ASF1A (L-020222-02-0020, lot 200427), ASF1B (L-020553-00-0020, lot 200427), non-targeting control (#D-001810-10-20).

1.3 Antibodies

The following primary and secondary antibodies were used for immuno-staining at the indicated dilutions: rabbit anti-H3K27ac (Abcam, ab4729, lot GR183922-1, 1:1000), mouse anti-PML (Santa Cruz, sc-966, 1:100), mouse anti-RPA (Abcam, ab2175, 1:300), rabbit anti-TRF2 (Novus Biologicals, NB11057130, 1:250), rabbit anti-SUMO2/3 (Abcam, ab3742, 1:200), mouse anti-NCAM/CD56 (Thermo Fisher Scientific, MA1-06801, 1:200), goat anti-rabbit Alexa Fluor 568 (Life Technologies, A11036, 1:300), goat anti-mouse Alexa Fluor 488 (Life Technologies, A11029, 1:300).

1.4 Relevant fine chemicals

Doxycycline (Sigma-Aldrich, D9391) was dissolved at 1 µg/ml in sterile nuclease-free water and used at 1-5 µg/ml to treat cells. The bromodomain inhibitor JQ1 (Sigma-Aldrich, SML1524) was dissolved in DMSO (Sigma-Aldrich) and used at 1 µM for the BRD4 binding experiments and at 2-32 µM for treatment in microplates. The MRN complex inhibitor Mirin (Selleckchem, Cat. no. S8096) was purchased pre-dissolved at 10 mM in DMSO. Ethynyl-uridine (EU) for global nascent RNA labeling by the Click-IT RNA imaging kit (Invitrogen, C10329) was used at 1 mM treatment concentration.

1.5 Cell lines

Wildtype and *Suv39h* double null immortalized mouse embryonic fibroblast (iMEF) cell lines were obtained from the group of Thomas Jenuwein and are described in ref. [55]. U2OS 2-6-3 cells [161] and were gifted by David Spector. U2OS, HeLa wt, CAL72, and Saos2 cell lines were procured from the German Collection of Microorganisms and Cell Culture (DSMZ). HOS and MG-63 cell lines were obtained from CLS Cell Lines Services Germany. The origin of the NEM, MGBM1, and SF188 cell lines is detailed in ref. [180]. HeLa ST and LT cell lines were provided by the group of Jan Karlseder and were originally described in ref. [181]. HUVEC cells were purchased from Lonza. The U2OS NCL-RFP cell line was created using pTagRFP-NCL vector [251] by stable integration of plasmid DNA and subsequent selection of single resistant clones. HeLa H2A-YFP [252], doxycycline-inducible U2OS-ATRX [192], and NEM168 ATRX knock-out clones F2 and B5 [180] were previously described. All transgenic cell lines were constantly maintained in media containing appropriate selection antibiotics (0.75 mg/ml geneticin/G418). The VenorGeM Advance kit (Minerva Biolabs) was utilized to confirm the absence of mycoplasma, and cell line identity was confirmed using single nucleotide polymorphism-profiling (Multiplexion).

1.6 Tumor samples and tissue sectioning

Tumor specimens of leiomyosarcoma (LMS) are described in refs. [107, 165]. They were obtained from nine adult patients who had been diagnosed with LMS at NCT Heidelberg and Heidelberg University Hospital. The Institute of Pathology (Heidelberg University) carried out the histological evaluation. Patient samples were collected under protocol S-206/2011, which was approved by the ethics committee of Heidelberg University, and involved written patient consent. Fresh-frozen LMS samples were cryo-sectioned (5-6 µm) onto SuperFrost Plus microscopy slides (Thermo Fisher Scientific) by the DKFZ/NCT Sample Processing Laboratory. The fresh-frozen neuroblastoma (NB) samples used in this study are described in refs. [109, 165]. NB tissue was embedded in tissue freezing medium (Leica Biosystems), and

sectioned (5-6 μm) onto SuperFrost Plus microscopy slides using an EpreDia CryoStar NX50 cryostat. All tissue specimens and sections on slides were stored at -80°C .

1.7 Microscopy instrumentation

Live cell time course and optodroplet experiments were mostly carried out on a Zeiss AxioObserver Z1 widefield microscope equipped with a HXP 120V fluorescence light source (Leistungs-elektronik JENA GmbH). Images were recorded using a 20x/0.8 Plan Aplanachromat air objective (Zeiss) and a Zeiss AxioCam MRm monochrome camera, using the following emission filters: GFP (F46-002, AHF), tdTomato (43 HE, Zeiss). Light-inducible construct binding kinetics and reporter characterization experiments above/below *Ccrit* were conducted on a Leica TCS SP5 II confocal system and with a 63x Plan Aplanachromat oil immersion objective. ALT-FISH samples, chromocenters and reporter locus radial profiles, area/intensity were imaged on an Andor Dragonfly 505 spinning disc confocal system with a Nikon Ti2-E microscope, equipped with an Andor iXon Ultra 888 EMCCD camera. The following Nikon objectives were used: 100x/1.35 CFI SR HP Aplanachromat Lambda S silicone (ALT-FISH/IF on coverslips, iMEF samples, tumor tissue), 60x/1.4 Plan Apo I oil (ALT-FISH and Molecular Cartography), 40x/1.3 CFI Plan Fluor oil (radial profiles, area/intensity at reporter), 20x/0.75 Plan Apo I air (microplates). The laser diode light sources used for excitation through a quad-band dichroic unit (405/488/561/640) and the corresponding detected fluorescent dyes/proteins and emission filters (EF) were: 405 nm (DAPI; EF 450/50), 488 nm (H2A-YFP, Alexa488, EGFP, eosin; EF 525/50), 561 nm (NCL-RFP, Alexa 568, tdTomato, mCherry, TRITC; EF 600/50) and 637 nm (Atto 633, Alexa 647; EF 698/77 or EF 685/47). Live cell imaging was conducted at 37°C and 5% CO_2 .

1.8 Software

The image analysis workflows developed in this thesis were mainly written in the R programming language [160] and were executed in RStudio build 576 [253] with R v4.2.1. The EBImage R package v4.3.8 [167] was utilized to process images and quantify their features. Two of the three software workflows developed in this thesis are made available as separate publication or elsewhere. The *Telosegment toolkit* provides a full workflow for automated analysis of ALT-FISH images from cell lines and is available at <https://github.com/RippeLab/ALT-FISH> and published in ref. [165]. The workflow to quantify features of chromocenters is published [57, 145] and furthermore made available at <https://github.com/RippeLab/Chromocenters>. FIJI v2.1 [254] was utilized for image format conversion, image pre-processing (stitching, flat-field correction, cropping), intensity quantification and representation purposes. Custom FIJI workflows and the RS-FISH FIJI

plugin [198] were employed for spot assignment and ALT activity map generation. Cell nuclei in tumor tissue and microplate well images were segmented using Cellpose v0.72 [197]. The methods to quantify single cell RNA and BRD4 time courses data were based on the NSSQ and EBIImage R packages and are described in ref. [255].

2. Methods

2.1 Cell culture

Cell lines were grown at 37°C and 5% CO₂ and split every 2-4 days using 0.05% trypsin/EDTA (Gibco) for detachment. The following cell lines were cultured in DMEM medium supplemented with 1 g/l glucose, 10% doxycycline-free fetal bovine serum, 1x penicillin/streptomycin, and 2 mM stable glutamine: U2OS-2-6-3, MG-63, U2OS, Saos2, CAL72, and HeLa wt. HeLa ST, HeLa LT, MGBM1, SF188, iMEF and HOS cells were cultured in DMEM medium containing 4.5 g/l glucose or EMEM, respectively, with the same supplements. HUVEC cells were maintained in EGM-2 medium (Lonza). NEM165, NEM157, NEM168 parental lines and ATRX knock-out clones were cultured in Amniopan complete medium (PAN Biotech). Plating of iMEF cells for the chromocenter perturbation experiments is detailed in ref. [145]. For microscopy or qRT-PCR experiments conducted with U2OS 2-6-3 cells, they were seeded at a density of $1-2 \times 10^4$ cells or 3×10^5 cells per well in 8-well chambered cover glass slides (Nunc LabTek, Thermo Fisher Scientific) or standard 6-well culture plates, before transfection on the next day. For the microscopy samples, the medium was replaced with supplemented FluoroBrite imaging medium (Gibco) before transfection. Cells transfected with optogenetic constructs were grown in absence of light, illuminated with diffuse white LED light for endpoint (90 min, 24 hours) reporter activation, or illuminated with the appropriate light source during imaging for time courses and other live cell experiments.

Cell lines for ALT-FISH alone or in combination with other staining were seeded in standard 24-well plates onto uncoated, sterile 12 mm round glass coverslips at a density of $5-8 \times 10^4$ cells per 1 ml volume per well and stained the next day. For U2OS NCL-RFP/HeLa H2A-YFP co-cultures, a total of 2.5×10^4 cells from each cell line were seeded per well. For ALT-FISH and DAPI staining of drug-treated and untreated cells in microplates, U2OS (6×10^3 per well) and HOS cells (4×10^3 per well) were grown in 0.2 ml culture volume per well in Nuclon Delta Surface 96 well plates (Thermo Fisher Scientific) for 24 hours. Subsequently, fresh medium was substituted with medium containing solvent (DMSO) or the corresponding drug at various concentrations. In the solvent wells, the amount of solvent was adjusted to match the highest drug concentration. Cells were grown for 48 hours in the microplates before assayed.

2.2 Plasmid DNA and siRNA transfection

iMEF cells were transfected with Lipofectamine 3000 (Thermo Fisher Scientific). A detailed experimental protocol is available in ref. [145]. In short, equal amounts of dCas9 and sgRNA-expressing plasmid were transfected and cells were grown for 30 hours before assayed. U2OS 2-6-3 cells were transfected using Xtreme-Gene 9 transfection reagent (Roche) following the manufacturer's protocol. For microscopy experiments conducted in chambered coverglass slides, 200-400 ng plasmid DNA and 0.6 μ l transfection reagent in 20 μ l OptiMEM (Gibco) were used per well. The plasmid DNA mixture contained 100 ng sgRNA plasmid and 100 ng distributed among the remaining constructs. When no sgRNA plasmids were used, equal masses of plasmid DNA were transfected. For qRT-PCR experiments, transfection reactions were scaled up to 2 μ g plasmid DNA per well. Further treatments (doxycycline, JQ1) and the relevant assays were conducted at least 24 hours post-transfection. ASF1A/B and control siRNA pools were reverse-transfected into HeLa ST and LT cell lines using Lipofectamine RNAiMax transfection reagent (Thermo Fisher Scientific), according to the manufacturer's protocol and at a final concentration of 1 nM. siRNAs were diluted to 20 nM in 491 μ l OptiMEM medium (Gibco) and 9 μ l RNAiMax reagent. For isolating genomic DNA and protein lysates of perturbed cells, 10 cm dishes containing 1×10^6 suspended cells in 9.5 ml medium were transfected with 500 μ l of the transfection mix. Transfection was scaled down accordingly for 2.5×10^4 cells per 1 ml transfected in 24-well plates to conduct ALT-FISH staining afterwards. HeLa LT/ST cells were grown for 72 hours post-transfection before assayed.

2.3 Chromocenter perturbation experiments and data analysis

Chromocenter perturbation experiments and associated image analyses were conducted as described in ref. [145]. In short, iMEF cells grown on coverslips were subjected to H3K27ac immuno-staining and DAPI staining. Microscopy images were recorded as z-stacks of 51 frames (10 μ m, 0.2 μ m step size) on the Andor Dragonfly 505 spinning disc confocal system using the 100x objective. Maximum intensity projections were generated in FIJI and then further processed in RStudio. Cell nuclei were segmented in the DAPI channel and using adaptive intensity-based thresholding function *thresh* from the EBImage package [167]. The custom function *makeChromocenterMask* was employed to segment dCas9-bound chromocenter areas. It uses a threshold, which is calibrated by the dCas9 expression levels and contrast features of each cell. The formula used to calculate the adaptive chromocenter threshold for a cell is: $median(nuclear\ dCas9\ intensity) + sat_cutoff \times [maximum(nuclear\ dCas9\ intensity) - median(nuclear\ dCas9\ intensity)]$. The parameter *sat_cutoff* (default 0.1) enables fine-tuning of the segmentation stringency. The relative chromocenter area was calculated by dividing the total chromocenter mask area by the total nuclear area. Signal

enrichment in chromocenters (DAPI, H3K27ac) was computed by dividing the mean intensity across all chromocenter masks of one cell by the mean intensity in the nucleoplasm mask. To group cells based on the amount of MSR-bound VPR activator, the integrated GFP signal across all chromocenters of a cell was determined. Cells were grouped separately for iMEF wt and iMEF dn datasets by the 40th and 60th percentile.

2.4 Light-induced binding to the reporter and measurement of reporter features

rTetR binding to the reporter gene cluster was induced by the addition of doxycycline in absence of light and 15 minutes prior to an illumination phase of 90 min or 24 hours. For endpoint expression measurements by qRT-PCR and radial profiles, cells were illuminated with white diffuse LED light. For time-resolved measurements, the blue excitation light source from the corresponding microscope was used to induce the binding. The resulting images were analyzed FIJI [254] using functions of the NSSQ [255] and EBIImage [167] packages. The NSSQ package and the routines for reporter spot tracking, quantification and radial profile analysis were developed by Jorge Trojanowski.

For acquisition of PHR/CIBN construct binding kinetics, cells expressing the PHR-mCherry-VP16 and CIBN constructs were imaged on the Leica TCS SP5 II confocal system using the 63x objective. The sample was focused using the 594 nm laser to excite mCherry without triggering the optogenetic domains. A 2-3 minutes time series (6 s intervals, single z-plane) was acquired with both the 594 nm and 488 nm (GFP) laser, to trigger PHR/CIBN interaction. Afterwards, the 488 nm laser was switched off to allow for PHR/CIBN dissociation and the same cells were recorded for another 20-30 min (1 min intervals, 2 μ m z-stacks of 4 frames, 0.5 μ m step size). The reporter spot accumulation was tracked and its intensity was quantified. Mean spot intensities were background corrected by subtracting the mean intensity in a ring-shaped area around the spot. Corrected intensities were then normalized to the value of the last (binding, $t = 168$ s) and first (dissociation, $t = 0$ s) frame, respectively, and for each cell.

For generating radial spot profiles of tdMCP-tdTom, H3K27ac and mCh-BRD4 signals, cells activated for 24 hours were treated and stained as described in ref. [146]. Images were recorded as z-stacks of 15 frames (14 μ m, 1 μ m step size) on the Andor Dragonfly 505 spinning disc confocal system using the 40x objective. The NSSQ package was used to generate an optimized mean projection the best three z-slices for quantification. The reporter spot center position was manually selected in each cell utilizing the co-transfected LacI marker. Concentric rings of increasing pixel radius were used to create masks around the center of the spot, and mean intensities were measured within each ring up to a maximum radius of 9 pixels (~ 3 μ m). A ring-shaped background mask around the spot was used for local background estimation. These measurements were used to obtain the raw radial profiles.

Next, these profiles were normalized by subtracting the minimum value and dividing all values by the local background. The single cell profiles of each condition were averaged and their minimum value was again subtracted to obtain the final radial profiles.

For the assessment of reporter features (area and intensity) above versus below *Ccrit*, the same imaging setup was used on cells that had been transfected for 24 hours and were illuminated for 10 min before fixation. Image analysis was conducted with EBIImage, but analysis procedures were modified. Local thresholding was used to segment the spot in the activator and the LacI marker channel separately. The activator area outside the marker area was calculated for each cell by finding the difference between the areas of the marker and activator spot masks. Spot intensity in the activator channel was calculated as described above. Area and intensity analysis was conducted by Jorge Trojanowski.

2.5 Light-induced RNA and BRD4 time course experiments

Light-induced time course experiments to measure transcription and BRD4 dynamics followed existing protocols [255] and were conducted and analyzed by Jorge Trojanowski. Their application to the here presented data is also described in ref. [146]. The main steps of image acquisition and analysis are summarized in the following. After addition of doxycycline in the dark and incubation for 15 min, tdMCP-tdTom (nascent RNA) and mCherry-BRD4 (BRD4) image series of 60 or 90 min (2 min intervals, 3 μm z-stacks, 1 μm step size) were recorded on the Zeiss AxioObserver Z1 widefield microscope using a 20x air objective. Blue light used to excite the GFP-labeled activator constructs (GFP) recorded in parallel served as recruitment trigger. For the JQ1 experiments, cells were pre-treated with 1 μM JQ1 for three hours in the dark before the addition of doxycycline and start of imaging. The experiments involving pre-acetylation of the reporter with dCas9-p300 were carried out in the same manner, only that the additional constructs were co-transfected alongside rTetR-opto and PHR-GFP-AD.

Time course data was analyzed using functions from the NSSQ package [255]. It comprised nucleus and reporter spot segmentation in the GFP channel and tracking of nucleus and spot masks throughout the time series and quantification of their intensity features. Values of area-integrated fluorescence intensity at the reporter spot were then averaged between cells to yield the final enrichment trajectories of RNA and BRD4 signal at the reporter. RNA series were leveled by subtracting the value of the first time frame. BRD4 series were normalized to their maximum value, followed by subtraction of the first frame. Half activation times were calculated from maximum-normalized RNA series. The values for half activation time correspond to the first time point at which the normalized enrichment exceeded or equaled 50 percent. The nascent RNA maximum (plateau) values were determined by averaging the signal intensity over the last five time points in the non-normalized RNA series.

2.6 Optodroplet imaging and quantification

Images of cells forming optodroplets with the PHR-GFP-AD and rTetR-opto constructs were obtained using the widefield setup and settings of the time course experiments. Droplet formation was quantified after 12 minutes of illumination time. For quantification of droplet abundance, nuclear masks were generated by segmentation of the GFP channel. Next, the reporter spot was manually masked in each cell by a circular mask with fixed radius. In the remaining nuclear area, optodroplets were segmented using 1.75 times the median of the nuclear intensity as threshold. Their cumulative segmented area per cell was determined and normalized to the total nuclear area to yield the final values for droplet abundance (in percent) per cell. In addition, the cells were visually classified as containing optodroplets or not. A logistic function (outlined in ref. [146]) was used to fit the data and represent the relative droplet area as a smooth function of nuclear intensity. An empirically determined threshold of 1% matching the visual classification was used to define the critical value for droplet formation. C_{crit} was defined as the nuclear intensity at which the relative droplet area surpassed 1%. For the RNA time courses, the cells were grouped by visual classification. The analysis of image data for calculation of C_{crit} was carried out by Jorge Trojanowski.

High resolution imaging of activator optodroplets in living cells was performed on the Leica TCS SP5 II confocal system using the 63x oil objective. For this, doxycycline-induced cells transfected with PHR-GFP-AD and rTetR-opto constructs as well as a LacI marker were illuminated for 2-5 minutes using the 488 nm laser line to saturate recruitment and droplet formation. Single high resolution images (one z-plane) of entire cells with or without visible optodroplets were taken. In addition, a smaller field of view around the reporter gene array was recorded as time series of up to 10 min (at 0.1 s time intervals) to capture droplet fusions.

2.7 qRT-PCR

Total RNA was isolated using Qiazol lysis reagent (Qiagen), followed by chloroform extraction and isopropanol precipitation. The RNA was treated with RQ1 DNase (Promega) and purified by one round of each phenol/chloroform and chloroform extraction, before precipitation with ethanol, sodium acetate and GlycoBlue (Thermo Fisher Scientific). RNA concentration and purity were measured by absorbance. One microgram of the purified RNA was used for cDNA synthesis using the Superscript IV reverse transcriptase protocol (Thermo Fisher Scientific). The qRT-PCR was performed using 2 μ l of 1:40 diluted cDNA per 10 μ l reaction in technical triplicates. The SYBR Green PCR master mix (Applied Biosystems) was used with a final primer concentration of 500 nM. qRT-PCR data was analyzed using the $2^{-\Delta\Delta C(T)}$ method [256]. Expression levels of the reporter RNA were normalized to the levels of beta-actin mRNA and then expressed as fold-change relative to the mock transfected control.

2.8 C-circle and TRAP assays

Telomerase repeated amplification protocol (TRAP) assays were performed as described in ref. [165]. In short, cell pellets were lysed in NP-40 lysis buffer at a concentration of 500 cells/ μ l for 30 min on ice. For one reaction, a lysate equivalent of 500 cells was mixed with TRAP reaction mix. TRAP reaction products were separated on a 12% TBE/polyacrylamide gel stained with SYBR Green (Thermo Fisher Scientific).

CC assays were conducted with minor modifications to the original protocol in [120] and as described in ref. [165]. Genomic DNA (gDNA) was extracted from cell pellets using the Puregene core kit (Qiagen) and quantified using the Qubit dsDNA broad range kit (Thermo Fisher Scientific). Each amplification reaction was run with 20 ng of gDNA and in triplicates. For each sample, a corresponding negative control without Phi29 polymerase enzyme (New England Biolabs) was processed in parallel. The reactions were incubated for 8 h at 30°C, then at 65°C for 20 min. Using a dot microfiltration apparatus (Bio-Rad, Cat. no. 170-6545), the samples were blotted onto a Roti-Nylon plus 0.45 μ m membrane (Carl Roth) and immobilized by baking at 120°C for 20 min. Blotted amplification products were detected using the TeloTAGGG telomere length assay kit (Roche) with 30 nM of digoxigenin-labeled detection probe hybridized overnight at 42°C. A final concentration of 75 μ U/ml of Anti-DIG-AP Fab fragments (Roche) incubated for 60 min was used to detect the hybridized probes with the CDP-Star chemiluminescent substrate (Roche) on a Bio-Rad Chemi-Doc MP system. Membrane images were quantified using Imagelab 6.0.1 software. CC levels were expressed relative to an U2OS reference sample (20 ng gDNA) from the same membrane. Assay linearity and background levels were assessed using a U2OS gDNA dilution series and a HeLa gDNA negative control. The CC levels of the LMS and NB tumors samples were not measured in this thesis, but have instead been determined in two previous studies [107, 109].

2.9 Western blotting

Cell pellets were lysed in RIPA buffer supplemented with Halt protease inhibitor cocktail (Thermo Fisher Scientific) for 1 h on ice. Protein concentrations were determined by BCA assay (Pierce). 15-30 μ g of lysate per lane was separated on precast 4-20% gradient polyacrylamide gels (Bio-Rad) and transferred to low-fluorescence PVDF membranes (Thermo Fisher Scientific) with the Trans-Blot Turbo system (Bio-Rad). Blocking was performed in 5% skim milk/TBS for 1 h. Membranes were incubated with primary antibodies overnight at 4°C and in 5%BSA/TBST(0.1% Tween-20). HRP-conjugated (ATRX blots, 1h in 5% skim milk/TBS) or fluorescently labeled (ASF1 blots, 1 h in 1%BSA/TBST) secondary antibodies were used detection on a Chemi-Doc MP system (Bio-Rad). Clarity ECL western

substrate (Bio-Rad) was used for luminescence based detection. All antibodies and dilutions are listed in *Materials and Methods section 1.3* and in ref. [165].

2.10 ATRX re-expression

To obtain matching material for the CC assay and samples for ALT-FISH, 1×10^6 U2OS-ATRX cells and 1.5×10^4 cells/ml/well were plated into a T75 flask and a 24-well plate with coverslips. After one day, the medium was exchanged for medium containing 1 μ g/ml doxycycline to induce ATRX expression for 72 hours. In parallel, a corresponding T75 flask and 24-well plate were left untreated, but otherwise processed as the treated cells. The cells growing in the flasks were harvested for gDNA (CC assay) and protein (western blot) isolation. The matching 24-well plates were processed for ALT-FISH staining as outlined below.

2.11 ALT-FISH staining on cell lines

Cells grown on 12 mm glass coverslips were washed twice with phosphate buffered saline (PBS) buffer and fixed for 20 min at room temperature with cold (-20°C) 70% ethanol. Fixation was followed by two short washes in ALT-FISH washing buffer (100 mM Tris-HCl pH 8, 150 mM NaCl, 0.05% Tween-20). For optional RNase treatment, the samples were placed in a solution of PBS containing 50 μ g/ml of RNase A (Thermo Fisher Scientific, EN0531) and/or 10 U/ml of RNase H (New England Biolabs, M0297L) and incubated at 37°C for 30 min, followed by one wash in ALT-FISH washing buffer. Afterwards, excess buffer was removed by tapping the coverslip borders on clean dust-free wipes. The coverslips were then placed with the cell side facing downwards onto 30 μ l of hybridization solution (2xSSC, 8% de-ionized formamide, 5 nM Atto 633 labeled ALT-FISH probe), dispensed onto parafilm in a humidified chamber, and incubated with the solution at 37°C for 20 min. The samples were transferred to 24-well plates, shortly rinsed twice in 2xSSC buffer and then incubated in PBS with 5 μ M DAPI (Sigma-Aldrich) for 15 min. For optional cytoplasm staining, coverslips were incubated with DAPI and 20 μ g/ml TRITC-Phalloidin (Sigma-Aldrich) for 30 min. Subsequently, the samples were washed three times (5 min each) with PBS and rinsed once with distilled water. The samples were dehydrated by one rinse in 70% ethanol, followed by 2 min incubation in 100% ethanol, and finally dried for 5-10 min on a dust-free paper wipe. The dried coverslips were mounted on standard microscopy glass slides using Prolong Diamond antifade medium (Invitrogen). U2OS samples with 0.1 μ m poly-fluorescent TetraSpeck beads (Thermo Fisher Scientific, T7279) used for SSTR copy number estimation were generated by pre-diluting the beads in mounting medium to a final concentration of 1:100. The mounting medium was allowed to cure for at least 24 hours before imaging or long-term storage of samples at 4°C .

Cells grown in 96-well microplates were washed twice with PBS and fixed with 70% ethanol for 15 min at room temperature (200 μ l per well). Residual ethanol was removed without drying the samples and 200 μ l of hybridization buffer (2xSSC, 10% de-ionized formamide, 5 nM Alexa 647 labeled ALT-FISH probe) were added to each well. The microplates were sealed with parafilm to prevent evaporation and incubated at 37°C for 30 min. The hybridization solution was removed and the cells were incubated with 200 μ l 5 μ M DAPI in 2xSSC buffer for 5 min, followed by two washes (5 min each) with 2xSSC buffer. The buffer was replaced once more with 200 μ l of fresh 2xSSC buffer. Plates were sealed with parafilm and imaged, or stored for up to 17 days at 4°C in the dark before imaging.

2.12 Combined ALT-FISH and immuno-staining of cell lines

ALT-FISH was carried out before immuno-staining up to the two 2xSSC buffer washing steps after hybridization (see above). Afterwards, the samples were rinsed twice in ALT-FISH washing buffer and fixed with 4% paraformaldehyde(PFA)/PBS at room temperature for 10 min. The fixed cells were rinsed three times with PBS, permeabilized in 0.2% Triton-X100/PBS for 12 min and washed again for 5 min in PBS. Blocking of samples was carried out in PBS with 10% goat serum (Cell Signaling Technology) for 1 h. Afterwards, the coverslips were incubated for at least 1 h in 30 μ l of primary antibody in blocking buffer on parafilm. Following three washes of 0.002% NP-40/PBS (5 min each), the coverslips were incubated with fluorescently labeled secondary antibodies diluted in blocking buffer for at least 1 h. After three 5-minute washes in PBS, the samples underwent DAPI staining, dehydration, and mounting as described earlier. All antibodies and dilutions are listed in Materials and Methods *section 1.3* and in ref. [165].

2.13 Confocal imaging of ALT-FISH stained cell line samples

Samples on coverslips were imaged on the Andor Dragonfly 505 spinning disc confocal system with the 100x objective. Microplates were imaged using the 20x objective. Unless stated otherwise, the same 637 nm laser intensity, EMCCD camera exposure time and gain settings were used for recording TelC and TelG ALT-FISH signals across samples. Settings were varied between the two different formats (coverslips, microplates) and for the other channels (DAPI, TRITC-Phalloidin, IF signals). However, they were kept constant between conditions when comparison of absolute intensities was necessary (e.g. DNA content estimation by DAPI signal). 16-bit images of 1024x1024 pixels were recorded as z-stacks of 51 frames (10 μ m, 0.2 μ m step size) for coverslips and z-stacks of 16 frames (15 μ m, 1 μ m step size) for microplates. Automated tile-scans with 1% overlap between adjacent images or multiple manually selected positions were acquired for coverslip samples. In microplates, the

full area of each well (~28 mm²) was recorded as 11x11 tile scan with 10% image overlap in a fully automated fashion across all 60 accessible wells of a plate.

2.14 Combined ALT-FISH and other staining of tumor tissue sections

Tissue sections immobilized on glass slides were thawed for a few seconds at 37°C on a PCR cyclor slide holder, then fixed in 100% pre-cooled (-20°C) methanol at -20°C for 30 min. The samples were incubated in isopropanol for 1 min at room temperature, dried for 5 min, and subsequently the section-containing area on the slide was confined with a hydrophobic marker pen (Vector Laboratories). Tissue sections were stained with 5 µM DAPI in PBS for 15 min, washed for 1 min in PBS and then mounted in 70 µl PBS using a 24 x 50 mm cover glass. This step served to inspect tissue integrity (DAPI signal) on a wide-field microscope before proceeding with the staining. The cover glass was removed by brief upside-down immersion of the slide in a beaker filled with PBS. Afterwards, the slides were incubated in ALT-FISH washing buffer for 2 min. The buffer was removed and 300-500 µl of hybridization solution (2xSSC, 8% de-ionized formamide, 5 nM Atto 633 labeled TelC ALT-FISH probe) were added onto the confined slide area. The slides were kept in a humidified chamber inside a hybridization oven (ACD Bio) and incubated at 37°C for 20 minutes without a cover glass. Following hybridization, the slides were subjected to two washes (1 min each) in 2xSSC buffer at room temperature. The LMS and NB samples for which no pseudo hematoxylin/eosin (H&E) staining was performed, were subsequently fixed in 4% PFA/PBS for 10 min. This additional fixation step served to stabilize tissue integrity. They were washed twice with PBS (1 min each). Excess PBS was removed and the sections were mounted in 45 µl Prolong Gold antifade mounting medium (Invitrogen) using a 24x50 mm cover glass.

Pseudo H&E staining of the LMS tumor followed the same fixation and hybridization procedure as described above. However, DAPI staining was performed after hybridization, followed by five short rinses in PBS and one in distilled water. Subsequently the sample was stained with 300-500 µl of eosin solution for 1 min at room temperature. Eosin solution consisted of 1 volume eosin Y solution (Sigma-Aldrich) mixed with 9 volumes of sterile-filtered Tris-acetic acid buffer (0.45 M, pH 6). The slide was then briefly dipped 15 times in distilled water and further incubated for 15 min and again for 2 min in 2xSSC buffer, before mounting as before. All stained tumor tissue slides were imaged (if possible) immediately after the staining or kept at 4°C in a humidified chamber until microscopy.

2.15 Molecular Cartography (MC) and ALT-FISH on neuroblastoma

Fresh-frozen NB tumor tissue was sectioned at a thickness of 5 µm using an Eprexia CryoStar NX50 cryostat and immobilized onto the specialized MC glass slides. The slides compare to

standard cover glasses in thickness, but have the dimensions of regular microscopy slides. For staining and imaging on the MC instrument, an 8-well imaging chamber is adhered on top of the slide, wherein each well contains one or multiple tissue sections from the same tumor. A description of the MC method is found at www.resolvebioscience.com, including step by step protocols for sample and buffer preparation for the run. The run was conducted as service by the Single-Cell OpenLab at the DKFZ. During the overnight hybridization of the gene-specific probes, which is conducted prior to the run, the Alexa 647 labeled TelC ALT-FISH probe was co-hybridized at a final concentration of 5 nM. After the run had completed, the chambered MC tissue slide was removed from the instrument. The sample was rinsed three times in PBS, stained with 5 μ M DAPI/PBS for 15 min, and again rinsed three times in PBS. Finally, 250 μ l of fresh PBS were added for imaging.

2.16 Confocal microscopy of tissue samples

Tissue sections were imaged on the Andor Dragonfly 505 spinning disc confocal system. Images were recorded with the 100x objective for all NB and LMS samples. The NB tumor sample used for ALT-FISH/MC integration was recorded using the 60x objective. Tile-scans with 10% overlap between adjacent images were acquired from a central region of each tissue section ($\sim 1.3\text{-}2\text{ mm}^2$) as z-stacks of 51 frames (20 μ m with 0.4 μ m step size) or 26 frames (10 μ m, 0.4 μ m step size) for the ALT-FISH/MC sample. Eosin and NCAM immuno-staining (Alexa 488-conjugated secondary antibody) signals were recorded using the GFP filter settings. The same 637 nm laser intensity, camera exposure time and gain settings were used for recording TelC ALT-FISH signals across LMS and NB samples, but varied for the ALT-FISH/MC sample.

2.17 Cell line ALT-FISH image analysis with *Telosegment*

Image stacks from the Andor Dragonfly confocal system were converted into TIF format using the FIJI Bio-Formats importer [254] and then imported into RStudio using the EBImage package [167]. All subsequent analysis was conducted within the *Telosegment* pipeline implemented in R. The pipeline itself and all custom-written associated functions are available at <https://github.com/RippeLab/ALT-FISH>. The main steps of the workflow are outlined in the following paragraph. First, using the *findBestSlices* function, the optimal z-range of each stack was identified by evaluating the intensity gradient in the DAPI channel and range-optimized maximum intensity projections across all channels were generated for quantification. Next, DAPI and ALT-FISH images were processed by a gaussian blurring algorithm ($\sigma = 1$) before segmentation of nuclei and spots, respectively.

The function *makeNucMask* was employed to segment nuclei. *makeNucMask* performs a first pass segmentation using the adaptive thresholding function *thresh* from EBIImage and a user-specified intensity offset. Subsequently, it applies pixel erosion, dilation and hole-filling operators to refine the initial masks. Identification of connected pixel areas (nuclear masks) is then done using the function *bwlabel* from EBIImage. Masks that touch the image borders and those that fall below a user-specified area cutoff are removed. The function *makeCytoMask* was employed to segment the cytoplasm, if a cytoplasm marker was available. It uses a non-linear combination of the DAPI and cytoplasm channel images as input. In a first pass, it generates masks using *thresh* and a user-defined intensity offset. In a second pass, it estimates the background in the surrounding non-mask image regions. The background estimate is then used to refine the mask borders. Pixel erosion and a watershed algorithm are applied to separate adjacent masks. In order to assign nuclear masks to their respective cytoplasm masks, the function *matchCellMasks* was used. In addition, it finds incomplete nuclear/cytoplasmic masks at the image borders and removes those before further downstream analysis.

For spot segmentation, the ALT-FISH channel image regions corresponding to the final nuclear (and cytoplasm) masks were isolated. Then, the function *makeSpotMask* was applied to each region to identify spots. The same approach was applied to the TRF2 immuno-stained samples to identify telomeres. The function first calculates the median pixel intensity in each pixel area as a background estimate. A pixel is then considered as part of a spot when its value exceeds 2.5 times (2 times in case of cytoplasmic spots and TRF2 spots) this value. To obtain smoothed spot masks, erosion and dilation operators are applied. Finally, *bwlabel* assigns connected pixels to the same spot and spots with less than 5 pixels are removed. Spot numbers are then assigned to each cell and area by counting the number of connected pixel groups *bwlabel* calls.

In addition to spot assignment, the *Telosegment* workflow uses all masks to extract intensity features in these image regions across all channels. These include the mean, median, standard deviation, minimum, and maximum values. To extract intensity features in a specific mask, the semantic pixel labels generated by *bwlabel* are used to nullify all non-mask pixels in the image to quantify. Mathematical operators can be applied to the modified image object, which now contains only a selected pixel group values. *Telosegment* also uses the EBIImage function *computeFeatures* on semantic *bwlabel*-maps to extract area, shape features, and position of the individual objects. The shape features of the nuclear objects were used to derive a nuclear shape quality score (NSSQ) defined by: $NSSQ = (r_{\max} \times r_{SD} \times \text{perimeter}) / (r_{\min} \times \text{area})$, where r corresponds to the object radius. Max, min and SD denote the maximum, minimum value and standard deviation of the object radius. The score increases with

increasing deviation from a circular/oval shape. An empirical cutoff of $NSQS > 3$ was used to exclude low quality nuclei.

To calculate the coefficient of variation (CV) for the cell line co-culture ALT-FISH experiments, the standard deviation of nuclear ALT-FISH intensity was divided by the mean nuclear intensity of each nucleus. For measuring co-localization of ALT-FISH spots with signals from immunostaining, the ratio of the mean intensity in the segmented spot mask area to the mean intensity in the corresponding nucleoplasm mask area was used to calculate the fold-enrichment of a signal over background for each spot in each nucleus. Spots with ratios greater than 1.5 were classified as co-localizing with the corresponding signal. This threshold value was determined as the 99th percentile of the distribution of all ratio values for the DAPI signal. This type of co-localization analysis assumes that the DAPI signal is not enriched in the spot areas, and thus serves as a measure for a homogeneous non-enriched signal.

2.18 High-content microplate image analysis with *Telosegment-HT*

Image tile-scans from microplates were converted into TIF format, flat-field corrected, stitched and maximum intensity-projected in FIJI [254] using a custom ImageJ macro developed by Anne Rademacher, a postdoc in our group. Cell nuclei in the DAPI images were segmented using Cellpose [197] (model "cyto", diameter = 40). Spots in the ALT-FISH channel were detected using the FIJI plugin RS-FISH [198] implemented in an ImageJ macro for batch processing. The following RS-FISH parameters were used for spot detection: sigma = 0.77, threshold = 0.003, default RANSAC parameters, no background subtraction, no spot intensity filtering. Pre-processing, and primary analysis (nucleus segmentation and spot detection) was done across multiple plates using the same parameters. Subsequently, secondary analysis was carried out on all wells of a single plate. This was necessary because the well center position in the stitched images differed between plates, and required adjustment specific to each plate. For secondary analysis, the Cellpose nuclear masks and corresponding ALT-FISH and DAPI images were imported into RStudio as EBImage image objects. The output from RS-FISH (spot xy-position, center intensity) was imported and converted into an image object, where each spot was represented by one pixel denoting its center and containing the intensity as pixel value. In a first step, a circular well mask of fixed radius was generated to exclude nuclei and segmentation artifacts that occurred close to or outside the well borders. Its center position in all well images was calibrated by aggregating the Cellpose masks from all wells into a single reference image. After removal of border objects, the final nuclear masks were used to quantify their shape, area and position features using *computeFeatures* from EBImage. A custom function *quantNuclei* was written that assigns spots to nuclei by their center pixel falling into a nuclear mask or not, and furthermore extracts DAPI and ALT-FISH

intensity features for each nucleus. The final ALT-FISH spot count per cell included all spots detected by RS-FISH in the primary analysis independent of their intensity. The cumulative ALT-FISH spot intensity per cell was calculated as the sum of all RS-FISH spot center intensities falling into one cell nucleus area. Quality of the microplate data was assessed on two levels. A manual inspection of the segmentation and spot count data of all wells was conducted. Wells that were found to have considerable segmentation artifacts or stitching problems during pre-processing were eliminated from the analysis. Furthermore, the data was filtered based on a nuclear size range (4000 to 15000 pixels) and an NSQS score smaller than 3 (see above for details). The latter served to removed smaller segmentation artifacts or poorly segmented large groups of nuclei.

2.19 Tissue image analysis

Image tile-scans of LMS and NB tumor tissue sections were stitched using Fusion v2.3 (Andor) on the employed spinning disc confocal system. The tile-scan of LMS tumor L1 was stitched in in FIJI using a custom macro. Maximum intensity z-projections generated in FIJI were used for all subsequent analyses. Images of the NB tumor used for ALT-FISH/MC integration were pre-processed in the same way. Before performing segmentation, the DAPI image from the confocal system was registered to the DAPI image from the MC instrument using the FIJI plugin *bUnwarpJ* (<https://github.com/fiji/bUnwarpJ>). Registration was done by Anne Rademacher, a postdoc in our group.

Segmentation and spot detection was performed as for the microplate data, but with different parameters. Cell nuclei were segmented using Cellpose [197] (model "cyto", LMS diameter = 80, NB diameter = 55). Spots were detected by RS-FISH [198] using the following parameters: sigma = 1.75, threshold = 0.0011, default RANSAC parameters, no background subtraction, minimum spot intensity of 1700 counts. To remove auto-fluorescent tissue artifacts, a semi-automated manual inspection routine was devised in RStudio. It was used to manually mark nuclei overlapping with artifacts and specifically exclude their masks before proceeding with the downstream analyses. Spot assignment to the nuclei and subsequent generation of ALT activity maps for both LMS and NB tumors was conducted using a semi-automated workflow in both FIJI and RStudio. Similar to the approach developed for the microplates, the spot positions were drawn as single pixels to create a binary image (spots = 1, background = 0) with the same dimensions. The spot count per cell was determined by measuring the sum of pixel intensities in each nuclear area on the binary image. To generate the ALT activity maps, the nuclei were categorized into the four ALT activity groups, depending on their spot count: no ALT (0), low ALT (1-2), ALT (3-20), and super-ALT (>20). To isolate nuclear masks that belonged to one category, the corresponding nucleus identifiers were extracted from the

semantic Cellpose masks. The mask objects were then drawn onto a new empty image. Images containing isolated nuclear masks were then treated as channels, combined into a final merged image and re-colored accordingly in FIJI to yield the final ALT activity maps. The same strategy was used to generate the map for the ALT-FISH/MC integration.

For the LMS tumor L1, the primary analysis was conducted as described above. However, spots were assigned to nuclei using a conceptionally similar workflow in FIJI, KNIME [257] and Napari (<https://napari.org>). Nearest neighbor analysis on the ALT activity map was carried out using Giotto [258]. Both of the latter approaches were developed and applied to the data by Stephan Tirier, a former postdoc in our group. The pseudo H&E representation of tumor L1 was generated from the DAPI and eosin images in RStudio using EBImage and utilizing the color transformation method described in ref. [208].

2.20 Cell cycle inference and spatial analysis of microplate data

For cell cycle inference analysis, the raw DNA content per cell was calculated as the sum of pixel intensities in the DAPI channel for a segmented nuclear area. To quantitatively compare DNA content between plates stained and imaged on different days, the raw DNA content values were normalized. Initially, the G01 and G2M peaks were identified by locating the top two maxima in the raw DNA content distribution histogram of each plate. The raw DNA content values DC_{raw} of each cell in that plate were then transformed according to the following formula to yield the normalized DNA content value DC_{norm} of a cell: $[(DC_{raw} - DCG01) / (DCG2M - DCG01)] + 1$, where $DCG01$ and $DCG2M$ represent the raw DNA content value at the center of the G01 and G2M peak, respectively. The value 1 was added so that the normalized G01 and G2M peaks are centered around 1 and 2, respectively. After normalization, data from multiple plates was pooled and further filtered by the quality criteria already described in *section 2.18*. The DNA content distribution of the pooled dataset was then used for cell cycle stage assignment based on specified ranges of DNA content. These were mainly determined by the G01 and G2M peaks. The center values of the two peaks (μ_{G01} , μ_{G2M}) were estimated by fitting two gaussian distributions to the histogram. Cells in G01 and G2M phase were defined as having a DNA content between $\mu \pm 2$ standard deviations (σ) of the respective gaussian fits. Cells with values below the $\mu_{G01} - 2\sigma$ boundary were defined as subG01, and above the $\mu_{G2M} + 2\sigma$ as high ploidy (HP). Cells with values between $\mu_{G01} + 2\sigma$ and $\mu_{G2M} - 2\sigma$ were categorized as S phase cells. Early and late S phase were distinguished by the $\mu_{G01} + 3\sigma$ boundary value.

The analyses to correlate the number of cell neighbors to ALT-FISH signals was conducted on six representative U2OS wells that displayed a broad range of local cell densities. Nearest

neighbor maps were derived from the nucleus positions using the `dnearest` function from the `spdep v1.2-8` R package. To determine neighboring cells, the center of mass of nuclear mask objects was utilized as the reference coordinates. Specifically, cells were deemed neighbors if their nucleus center coordinates fell within a distance range of 15 and 100 pixels (equivalent to 9 and 60 μm). Extraction of microcolonies was done using the `igraph v1.4.1` R package. First, a Euclidean distance matrix was generated, comparing the coordinates of all cells in a well. The distance matrix was then converted to an adjacency graph, where two cells were considered adjacent if the distance between their nuclei centers was less than 100 pixels, and non-adjacent if their distance was greater. The fast greedy clustering algorithm from the `igraph` package was then applied to the adjacency graph to find adjacent cell communities. By applying a size filter of 20 cells to the clusters, the microcolonies were reliably extracted among all identified clusters. Further size filtering was employed to specifically extract datapoints from doublet microcolonies. The following method was used to conduct the ALT-FISH feature similarity analysis between cells in doublets. In a first step, the absolute pair-wise difference in feature value between two cell1 and cell2 in the same doublet was calculated by subtraction. This value was normalized by dividing it by the median of absolute differences between cell1 and all other cells in the dataset but cell2. The resulting pair-wise dissimilarity score indicates more within-colony similarity if below 1 and more between-colony similarity if above 1. In order to establish a background distribution of pairwise similarities among the cells in the dataset, the score was calculated for the same cells after randomly grouping them into doublets.

2.21 Additional methods

Fluorescence correlation spectroscopy (FCS) for SSTR copy number estimation was performed on an LSM710 microscope with a ConfoCor3 extension and equipped with a 40x/1.2 water objective. Alexa 633 dye (Molecular Probes, A33084) and Atto 633 TeIC ALT-FISH probe were diluted in 10 mM Tris-HCl pH 8, 10 mM NaCl and 0.002% Tween20 to a final concentration of 20 nM. Poly-fluorescent 0.1 μm TetraSpeck beads (Thermo Fisher Scientific, T7279) were used undiluted as provided by the manufacturer (300 pM). FCS data were fitted with the ZEN 2008 software. FCS data acquisition and analysis were carried out by Norbert Mücke, a research engineer in our group.

References

1. Spector, D.L., Nuclear domains. *J Cell Sci* **114**, 2891-3 (2001).
2. Cremer, T., et al., The 4D nucleome: Evidence for a dynamic nuclear landscape based on co-aligned active and inactive nuclear compartments. *FEBS Lett* **589**, 2931-43 (2015).
3. van Steensel, B. and E.E.M. Furlong, The role of transcription in shaping the spatial organization of the genome. *Nat Rev Mol Cell Biol* **20**, 327-337 (2019).
4. Matera, A.G., et al., Nuclear bodies: random aggregates of sticky proteins or crucibles of macromolecular assembly? *Dev Cell* **17**, 639-47 (2009).
5. Shaw, P.J. and E.G. Jordan, The nucleolus. *Annu Rev Cell Dev Biol* **11**, 93-121 (1995).
6. Iborra, F.J., et al., Active RNA polymerases are localized within discrete transcription "factories" in human nuclei. *J Cell Sci* **109**, 1427-36 (1996).
7. Rieder, D., Z. Trajanoski, and J.G. McNally, Transcription factories. *Front Genet* **3**, 221 (2012).
8. Lamond, A.I. and D.L. Spector, Nuclear speckles: a model for nuclear organelles. *Nat Rev Mol Cell Biol* **4**, 605-12 (2003).
9. Heun, P., SUMOrganization of the nucleus. *Curr Opin Cell Biol* **19**, 350-5 (2007).
10. Dellaire, G. and D.P. Bazett-Jones, PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. *Bioessays* **26**, 963-77 (2004).
11. Lallemand-Breitenbach, V. and H. de The, PML nuclear bodies. *Cold Spring Harb Perspect Biol* **2**, a000661 (2010).
12. Corpet, A., et al., PML nuclear bodies and chromatin dynamics: catch me if you can! *Nucleic Acids Res* **48**, 11890-11912 (2020).
13. van Steensel, B. and A.S. Belmont, Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* **169**, 780-791 (2017).
14. Briand, N. and P. Collas, Lamina-associated domains: peripheral matters and internal affairs. *Genome Biol* **21**, 85 (2020).
15. Rego, A., et al., The facultative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure. *J Cell Sci* **121**, 1119-27 (2008).
16. Chadwick, B.P. and H.F. Willard, Chromatin of the Barr body: histone and non-histone proteins associated with or excluded from the inactive X chromosome. *Hum Mol Genet* **12**, 2167-78 (2003).
17. Probst, A.V. and G. Almouzni, Pericentric heterochromatin: dynamic organization during early development in mammals. *Differentiation* **76**, 15-23 (2008).
18. Bandaria, J.N., et al., Shelterin Protects Chromosome Ends by Compacting Telomeric Chromatin. *Cell* **164**, 735-746 (2016).
19. Harley, C.B., A.B. Futcher, and C.W. Greider, Telomeres shorten during ageing of human fibroblasts. *Nature* **345**, 458-60 (1990).
20. Guenatri, M., et al., Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol* **166**, 493-505 (2004).
21. Misteli, T., The concept of self-organization in cellular architecture. *J Cell Biol* **155**, 181-5 (2001).
22. Bhat, P., D. Honson, and M. Guttman, Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nature Reviews Molecular Cell Biology* **22**, 653-670 (2021).
23. Verschure, P.J., et al., In vivo HP1 targeting causes large-scale chromatin condensation and enhanced histone lysine methylation. *Mol Cell Biol* **25**, 4552-64 (2005).
24. Velazquez Camacho, O., et al., Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *Elife* **6**, e25293 (2017).

25. Schneider, R. and R. Grosschedl, Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Genes Dev* **21**, 3027-43 (2007).
26. Erdel, F. and K. Rippe, Formation of chromatin subcompartments by phase separation. *Biophys J* **114**, 2262-2270 (2018).
27. Frank, L. and K. Rippe, Repetitive RNAs as Regulators of Chromatin-Associated Subcompartment Formation by Phase Separation. *J Mol Biol* **432**, 4270-4286 (2020).
28. Hemmerich, P., L. Schmiedeberg, and S. Diekmann, Dynamic as well as stable protein interactions contribute to genome function and maintenance. *Chromosome Res* **19**, 131-51 (2011).
29. Rippe, K. and A. Papantonis, Functional organization of RNA polymerase II in nuclear subcompartments. *Curr Opin Cell Biol* **74**, 88-96 (2022).
30. Wang, C., V.N. Uversky, and L. Kurgan, Disordered nucleome: Abundance of intrinsic disorder in the DNA- and RNA-binding proteins in 1121 species from Eukaryota, Bacteria and Archaea. *Proteomics* **16**, 1486-98 (2016).
31. Borgia, A., et al., Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61-66 (2018).
32. Protter, D.S.W., et al., Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep* **22**, 1401-1412 (2018).
33. Brodsky, S., et al., Intrinsically Disordered Regions Direct Transcription Factor In Vivo Binding Specificity. *Mol Cell* **79**, 459-471 e4 (2020).
34. Tompa, P., Intrinsically unstructured proteins evolve by repeat expansion. *Bioessays* **25**, 847-55 (2003).
35. Martin, E.W. and A.S. Holehouse, Intrinsically disordered protein regions and phase separation: sequence determinants of assembly or lack thereof. *Emerg Top Life Sci* **4**, 307-329 (2020).
36. Brangwynne, C.P., et al., Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729-32 (2009).
37. Weber, S.C. and C.P. Brangwynne, Inverse size scaling of the nucleolus by a concentration-dependent phase transition. *Curr Biol* **25**, 641-6 (2015).
38. Larson, A.G., et al., Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* **547**, 236-240 (2017).
39. Kilic, S., et al., Phase separation of 53BP1 determines liquid-like behavior of DNA repair compartments. *EMBO J* **38**, e101379 (2019).
40. Banani, S.F., et al., Compositional Control of Phase-Separated Cellular Bodies. *Cell* **166**, 651-663 (2016).
41. Boija, A., et al., Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **175**, 1842-1855 e16 (2018).
42. Sabari, B.R., et al., Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958-17 (2018).
43. Alberti, S., A. Gladfelter, and T. Mittag, Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419-434 (2019).
44. Choi, J.M., A.S. Holehouse, and R.V. Pappu, Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annu Rev Biophys* **49**, 107-133 (2020).
45. McSwiggen, D.T., et al., Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev* **33**, 1619-1634 (2019).
46. Rippe, K., Liquid-Liquid Phase Separation in Chromatin. *Cold Spring Harb Perspect Biol* **14**, a040683 (2022).
47. Cordaux, R. and M.A. Batzer, The impact of retrotransposons on human genome evolution. *Nat Rev Genet* **10**, 691-703 (2009).
48. Lander, E.S., et al., Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
49. Horz, W. and W. Altenburger, Nucleotide sequence of mouse satellite DNA. *Nucleic Acids Res* **9**, 683-96 (1981).

50. Moyzis, R.K., et al., A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proc Natl Acad Sci U S A* **85**, 6622-6 (1988).
51. Amatori, S., et al., The dark side of histones: genomic organization and role of oncohistones in cancer. *Clin Epigenetics* **13**, 71 (2021).
52. Sakai, K., et al., Human ribosomal RNA gene cluster: identification of the proximal end containing a novel tandem repeat sequence. *Genomics* **26**, 521-6 (1995).
53. van Steensel, B., Chromatin: constructing the big picture. *EMBO J* **30**, 1885-1895 (2011).
54. Ernst, J. and M. Kellis, ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* **9**, 215-6 (2012).
55. Peters, A.H., et al., Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323-37 (2001).
56. Erdel, F., K. Muller-Ott, and K. Rippe, Establishing epigenetic domains via chromatin-bound histone modifiers. *Ann N Y Acad Sci* **1305**, 29-43 (2013).
57. Erdel, F., et al., Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of HP1-Driven Liquid-Liquid Phase Separation. *Mol Cell* **78**, 236-249 e7 (2020).
58. Li, C.H., et al., MeCP2 links heterochromatin condensates and neurodevelopmental disease. *Nature* **586**, 440-444 (2020).
59. Lee, T.I. and R.A. Young, Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* **34**, 77-137 (2000).
60. Zhang, T., S. Cooper, and N. Brockdorff, The interplay of histone modifications - writers that read. *EMBO Rep* **16**, 1467-81 (2015).
61. Ng, H.H., et al., Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11**, 709-19 (2003).
62. Gorisch, S.M., et al., Histone acetylation increases chromatin accessibility. *J Cell Sci* **118**, 5825-34 (2005).
63. Yang, Z., et al., Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* **19**, 535-45 (2005).
64. Dixon, J.R., et al., Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-80 (2012).
65. Rademacher, A., et al., Real-time observation of light-controlled transcription in living cells. *J Cell Sci* **130**, 4213-4224 (2017).
66. Osborne, C.S., et al., Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* **36**, 1065-71 (2004).
67. Dotson, G.A., et al., Deciphering multi-way interactions in the human genome. *Nat Commun* **13**, 5498 (2022).
68. Senecal, A., et al., Transcription factors modulate c-Fos transcriptional bursts. *Cell Rep* **8**, 75-83 (2014).
69. Suter, D.M., et al., Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**, 472-4 (2011).
70. Whyte, W.A., et al., Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-19 (2013).
71. Pott, S. and J.D. Lieb, What are super-enhancers? *Nat Genet* **47**, 8-12 (2015).
72. Boehning, M., et al., RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol Biol* **25**, 833-840 (2018).
73. Kent, S., et al., Phase-Separated Transcriptional Condensates Accelerate Target-Search Process Revealed by Live-Cell Single-Molecule Imaging. *Cell Rep* **33**, 108248 (2020).
74. Chong, S., et al., Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555-17 (2018).
75. Cho, W.K., et al., Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412-415 (2018).

76. Wei, M.T., et al., Nucleated transcriptional condensates amplify gene expression. *Nat Cell Biol* **22**, 1187-1196 (2020).
77. Quintero-Cadena, P., T.L. Lenstra, and P.W. Sternberg, RNA Pol II Length and Disorder Enable Cooperative Scaling of Transcriptional Bursting. *Mol Cell* **79**, 207-220 e8 (2020).
78. Shin, Y., et al., Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* **168**, 159-171 e14 (2017).
79. Patel, A., et al., A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066-77 (2015).
80. Guo, Y.E., et al., Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543-548 (2019).
81. Chong, S., et al., Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. *Mol Cell* **82**, 2084-2097 e5 (2022).
82. Showpnil, I.A., et al., EWS/FLI mediated reprogramming of 3D chromatin promotes an altered transcriptional state in Ewing sarcoma. *Nucleic Acids Res* **50**, 9814-9837 (2022).
83. Klein, I.A., et al., Partitioning of cancer therapeutics in nuclear condensates. *Science* **368**, 1386-1392 (2020).
84. Blackburn, E., Structure and function of telomeres. *Nature* **350**, 569-573 (1991).
85. de Lange, T., Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* **19**, 2100-10 (2005).
86. Makarov, V.L., Y. Hirose, and J.P. Langmore, Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* **88**, 657-66 (1997).
87. de Lange, T., T-loops and the origin of telomeres. *Nat Rev Mol Cell Biol* **5**, 323-9 (2004).
88. Griffith, J.D., et al., Mammalian telomeres end in a large duplex loop. *Cell* **97**, 503-14 (1999).
89. van Steensel, B., A. Smogorzewska, and T. de Lange, TRF2 protects human telomeres from end-to-end fusions. *Cell* **92**, 401-13 (1998).
90. Hockemeyer, D., et al., POT1 protects telomeres from a transient DNA damage response and determines how human chromosomes end. *EMBO J* **24**, 2667-78 (2005).
91. Stansel, R.M., T. de Lange, and J.D. Griffith, T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang. *EMBO J* **20**, 5532-40 (2001).
92. Derheimer, F.A. and M.B. Kastan, Multiple roles of ATM in monitoring and maintaining DNA integrity. *FEBS Lett* **584**, 3675-81 (2010).
93. Cesare, A.J., et al., The telomere deprotection response is functionally distinct from the genomic DNA damage response. *Mol Cell* **51**, 141-55 (2013).
94. Hayflick, L., The Limited in Vitro Lifetime of Human Diploid Cell Strains. *Exp Cell Res* **37**, 614-36 (1965).
95. Levy, M.Z., et al., Telomere end-replication problem and cell aging. *J Mol Biol* **225**, 951-60 (1992).
96. Maciejowski, J., et al., Chromothripsis and Kataegis Induced by Telomere Crisis. *Cell* **163**, 1641-54 (2015).
97. d'Adda di Fagagna, F., et al., A DNA damage checkpoint response in telomere-initiated senescence. *Nature* **426**, 194-8 (2003).
98. Chin, L., et al., p53 deficiency rescues the adverse effects of telomere loss and cooperates with telomere dysfunction to accelerate carcinogenesis. *Cell* **97**, 527-38 (1999).
99. Hanahan, D. and R.A. Weinberg, Hallmarks of cancer: the next generation. *Cell* **144**, 646-74 (2011).
100. Greider, C.W. and E.H. Blackburn, Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* **43**, 405-13 (1985).

101. Shay, J.W. and W.E. Wright, Senescence and immortalization: role of telomeres and telomerase. *Carcinogenesis* **26**, 867-74 (2005).
102. Bryan, T.M., et al., Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J* **14**, 4240-8 (1995).
103. Sobinoff, A.P. and H.A. Pickett, Alternative Lengthening of Telomeres: DNA Repair Pathways Converge. *Trends Genet* **33**, 921-932 (2017).
104. Heaphy, C.M., et al., Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes. *Am J Pathol* **179**, 1608-15 (2011).
105. Lawlor, R.T., et al., Alternative lengthening of telomeres (ALT) influences survival in soft tissue sarcomas: a systematic review with meta-analysis. *BMC Cancer* **19**, 232 (2019).
106. Marinoni, I., et al., Loss of DAXX and ATRX are associated with chromosome instability and reduced survival of patients with pancreatic neuroendocrine tumors. *Gastroenterology* **146**, 453-60 e5 (2014).
107. Chudasama, P., et al., Integrative genomic and transcriptomic analysis of leiomyosarcoma. *Nat Commun* **9**, 144 (2018).
108. Heaphy, C.M., et al., Altered telomeres in tumors with ATRX and DAXX mutations. *Science* **333**, 425 (2011).
109. Hartlieb, S.A., et al., Alternative lengthening of telomeres in childhood neuroblastoma from genome to proteome. *Nat Commun* **12**, 1269 (2021).
110. Zhang, J.M., et al., Alternative Lengthening of Telomeres through Two Distinct Break-Induced Replication Pathways. *Cell Rep* **26**, 955-968 e3 (2019).
111. Min, J., W.E. Wright, and J.W. Shay, Alternative Lengthening of Telomeres Mediated by Mitotic DNA Synthesis Engages Break-Induced Replication Processes. *Mol Cell Biol* **37**, (2017).
112. Dunham, M.A., et al., Telomere maintenance by recombination in human cells. *Nat Genet* **26**, 447-50 (2000).
113. Oganessian, L. and J. Karlseder, Mammalian 5' C-rich telomeric overhangs are a mark of recombination-dependent telomere maintenance. *Mol Cell* **42**, 224-36 (2011).
114. Silva, B., et al., FANCM limits ALT activity by restricting telomeric replication stress induced by deregulated BLM and R-loops. *Nat Commun* **10**, 2253 (2019).
115. Nabetani, A. and F. Ishikawa, Unusual telomeric DNAs in human telomerase-negative immortalized cells. *Mol Cell Biol* **29**, 703-13 (2009).
116. Cesare, A.J. and J.D. Griffith, Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops. *Mol Cell Biol* **24**, 9948-57 (2004).
117. Chen, Y.A., et al., Extrachromosomal telomere repeat DNA is linked to ALT development via cGAS-STING DNA sensing pathway. *Nat Struct Mol Biol* **24**, 1124-1131 (2017).
118. Mazzucco, G., et al., Telomere damage induces internal loops that generate telomeric circles. *Nat Commun* **11**, 5297 (2020).
119. Moreno, S.P., et al., TZAP overexpression induces telomere dysfunction and ALT-like activity in ATRX/DAXX-deficient cells. *iScience* **26**, 106405 (2023).
120. Henson, J.D., et al., DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. *Nat Biotechnol* **27**, 1181-5 (2009).
121. Yeager, T.R., et al., Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res* **59**, 4175-9 (1999).
122. Zhang, H., et al., Liquid condensation drives telomere clustering during ALT. *bioRxiv* 633040 (2019).
123. Zhang, H., et al., Nuclear body phase separation drives telomere clustering in ALT cancer cells. *Mol Biol Cell* mbcE19100589 (2020).
124. Henson, J.D., et al., A robust assay for alternative lengthening of telomeres in tumors shows the significance of alternative lengthening of telomeres in sarcomas and astrocytomas. *Clin Cancer Res* **11**, 217-25 (2005).

125. Dyer, M.A., et al., ATRX and DAXX: Mechanisms and Mutations. *Cold Spring Harb Perspect Med* **7**, (2017).
126. Brosnan-Cashman, J.A., et al., ATRX loss induces multiple hallmarks of the alternative lengthening of telomeres (ALT) phenotype in human glioma cell lines in a cell line-specific manner. *PLoS One* **13**, e0204159 (2018).
127. Lovejoy, C.A., et al., Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLoS Genet* **8**, e1002772 (2012).
128. Turkalo, T.K., et al., A non-genetic switch triggers alternative telomere lengthening and cellular immortalization in ATRX deficient cells. *Nat Commun* **14**, 939 (2023).
129. Cubiles, M.D., et al., Epigenetic features of human telomeres. *Nucleic Acids Res* **46**, 2347-2355 (2018).
130. Azzalin, C.M., et al., Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**, 798-801 (2007).
131. Gauchier, M., et al., SETDB1-dependent heterochromatin stimulates alternative lengthening of telomeres. *Sci Adv* **5**, eaav3673 (2019).
132. Ng, L.J., et al., Telomerase activity is associated with an increase in DNA methylation at the proximal subtelomere and a reduction in telomeric transcription. *Nucleic Acids Res* **37**, 1152-9 (2009).
133. Sampl, S., et al., Expression of Telomeres in Astrocytoma WHO Grade 2 to 4: TERRA Level Correlates with Telomere Length, Telomerase Activity, and Advanced Clinical Grade. *Transl Oncol* **5**, 56-65 (2012).
134. Porro, A., et al., Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. *Mol Cell Biol* **30**, 4808-17 (2010).
135. Jiang, W.Q., et al., Induction of alternative lengthening of telomeres-associated PML bodies by p53/p21 requires HP1 proteins. *J Cell Biol* **185**, 797-810 (2009).
136. Jiang, W.Q., et al., HP1-mediated formation of alternative lengthening of telomeres-associated PML bodies requires HIRA but not ASF1a. *PLoS One* **6**, e17036 (2011).
137. Feretzaki, M., et al., RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. *Nature* **587**, 303-308 (2020).
138. Rippe, K. and B. Luke, TERRA and the state of the telomere. *Nat Struct Mol Biol* **22**, 853-8 (2015).
139. Gocha, A.R., et al., Human sarcomas are mosaic for telomerase-dependent and telomerase-independent telomere maintenance mechanisms: implications for telomere-based therapies. *The American journal of pathology* **182**, 41-8 (2013).
140. Pezzolo, A., et al., Intratumoral diversity of telomere length in individual neuroblastoma tumors. *Oncotarget* **6**, 7493-503 (2015).
141. Barthel, F.P., et al., Systematic analysis of telomere length and somatic alterations in 31 cancer types. *Nature Genetics* **49**, 349-357 (2017).
142. Hakin-Smith, V., et al., Alternative lengthening of telomeres and survival in patients with glioblastoma multiforme. *Lancet* **361**, 836-8 (2003).
143. Hu, J., et al., Antitelomerase therapy provokes ALT and mitochondrial adaptive mechanisms in cancer. *Cell* **148**, 651-63 (2012).
144. Narlikar, G.J., et al., Is transcriptional regulation just going through a phase? *Mol Cell* **81**, 1579-1585 (2021).
145. Frank, L., et al., Transcriptional Activation of Heterochromatin by Recruitment of dCas9 Activators. *Methods Mol Biol* **2351**, 307-320 (2021).
146. Trojanowski, J., et al., Transcription activation is enhanced by multivalent interactions independent of phase separation. *Mol Cell* **82**, 1878-1893 e10 (2022).
147. Kennedy, M.J., et al., Rapid blue-light-mediated induction of protein interactions in living cells. *Nat Methods* **7**, 973-5 (2010).
148. Chen, B., et al., Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479-91 (2013).

149. Zalatan, J.G., et al., Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160**, 339-50 (2015).
150. Polstein, L.R. and C.A. Gersbach, A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat Chem Biol* **11**, 198-200 (2015).
151. Chavez, A., et al., Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**, 326-8 (2015).
152. Tumber, T., G. Sudlow, and A.S. Belmont, Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. *J Cell Biol* **145**, 1341-54 (1999).
153. Schmitz, M.L. and P.A. Baeuerle, The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B. *Embo j* **10**, 3805-17 (1991).
154. Hardwick, J.M., et al., The Epstein-Barr virus R transactivator (Rta) contains a complex, potent activation domain with properties different from those of VP16. *J Virol* **66**, 5500-8 (1992).
155. Almouzni, G. and A.V. Probst, Heterochromatin maintenance and establishment: lessons from the mouse pericentromere. *Nucleus* **2**, 332-8 (2011).
156. Hahn, M., et al., Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes Dev* **27**, 859-72 (2013).
157. Strom, A.R., et al., Phase separation drives heterochromatin domain formation. *Nature* **547**, 241-245 (2017).
158. Anton, T., et al., Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. *Nucleus* **5**, 163-72 (2014).
159. Gerritsen, M.E., et al., CREB-binding protein/p300 are transcriptional coactivators of p65. *Proc Natl Acad Sci U S A* **94**, 2927-32 (1997).
160. R Core Team, R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. URL <https://www.R-project.org/> (2020).
161. Janicki, S.M., et al., From silencing to gene expression: real-time analysis in single cells. *Cell* **116**, 683-98 (2004).
162. Filippakopoulos, P., et al., Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**, 214-31 (2012).
163. Jang, M.K., et al., The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* **19**, 523-34 (2005).
164. Lewis, M., et al., Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**, 1247-54 (1996).
165. Frank, L., et al., ALT-FISH quantifies alternative lengthening of telomeres activity by imaging of single-stranded repeats. *Nucleic Acids Res* **50**, e61 (2022).
166. Rose, A.M., et al., Induction of the alternative lengthening of telomeres pathway by trapping of proteins on DNA. *Nucleic Acids Research* (2023).
167. Pau, G., et al., EBImage - an R package for image processing with applications to cellular phenotypes. *Bioinformatics* **26**, 979-81 (2010).
168. Arora, R., et al., RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun* **5**, 5220 (2014).
169. Tokutake, Y., et al., Extra-chromosomal telomere repeat DNA in telomerase-negative immortalized cell lines. *Biochem Biophys Res Commun* **247**, 765-72 (1998).
170. Baerlocher, G.M., et al., Flow cytometry and FISH to measure the average length of telomeres (flow FISH). *Nat Protoc* **1**, 2365-76 (2006).
171. Rigler, R. and E.S. Elson, Fluorescence correlation spectroscopy: theory and applications. Berlin: Springer(2001).
172. Draskovic, I., et al., Probing PML body function in ALT cells reveals spatiotemporal requirements for telomere recombination. *Proc Natl Acad Sci U S A* **106**, 15726-31 (2009).
173. Grudic, A., et al., Replication protein A prevents accumulation of single-stranded telomeric DNA in cells that use alternative lengthening of telomeres. *Nucleic Acids Res* **35**, 7267-78 (2007).

174. Fasching, C.L., et al., DNA damage induces alternative lengthening of telomeres (ALT) associated promyelocytic leukemia bodies that preferentially associate with linear telomeric DNA. *Cancer Res* **67**, 7072-7 (2007).
175. Lang, M., et al., Three-dimensional organization of promyelocytic leukemia nuclear bodies. *J Cell Sci* **123**, 392-400 (2010).
176. Chung, I., et al., PML body meets telomere: the beginning of an ALTERNATE ending? *Nucleus* **3**, 263-75 (2012).
177. Osterwald, S., et al., A three-dimensional colocalization RNA interference screening platform to elucidate the alternative lengthening of telomeres pathway. *Biotechnol J* **7**, 103-16 (2012).
178. Venturini, L., et al., Prognostic relevance of ALT-associated markers in liposarcoma: a comparative analysis. *BMC Cancer* **10**, 254 (2010).
179. Mason-Osann, E., et al., Identification of a novel gene fusion in ALT positive osteosarcoma. *Oncotarget* **9**, 32868-32880 (2018).
180. Deeg, K.I., et al., Dissecting telomere maintenance mechanisms in pediatric glioblastoma. *bioRxiv* 129106. (2017).
181. O'Sullivan, R.J., et al., Rapid induction of alternative lengthening of telomeres by depletion of the histone chaperone ASF1. *Nat Struct Mol Biol* **21**, 167-74 (2014).
182. Schwartzenuber, J., et al., Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**, 226-31 (2012).
183. Segura-Bayona, S., et al., Toslud-Like Kinases Suppress Innate Immune Signaling Triggered by Alternative Lengthening of Telomeres. *Cell Rep* **32**, 107983 (2020).
184. Mender, I. and J.W. Shay, Telomerase Repeated Amplification Protocol (TRAP). *Bio Protoc* **5**, e1657 (2015).
185. Pampalona, J., et al., Progressive telomere dysfunction causes cytokinesis failure and leads to the accumulation of polyploid cells. *PLoS Genet* **8**, e1002679 (2012).
186. Pickett, H.A., et al., Control of telomere length by a trimming mechanism that involves generation of t-circles. *EMBO J* **28**, 799-809 (2009).
187. Li, J.S., et al., TZAP: A telomere-associated protein involved in telomere length control. *Science* **355**, 638-641 (2017).
188. Chen, X., et al., Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma. *Cell Rep* **7**, 104-12 (2014).
189. Napier, C.E., et al., ATRX represses alternative lengthening of telomeres. *Oncotarget* **6**, 16543-58 (2015).
190. Geiller, H.E.B., et al., ATRX modulates the escape from a telomere crisis. *PLoS Genet* **18**, e1010485 (2022).
191. Clynes, D., et al., Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. *Nat Commun* **6**, 7538 (2015).
192. Deeg, K.I., et al., Cancer Cells with Alternative Lengthening of Telomeres Do Not Display a General Hypersensitivity to ATR Inhibition. *Front Oncol* **6**, 186 (2016).
193. Pladevall-Morera, D., et al., ATRX-Deficient High-Grade Glioma Cells Exhibit Increased Sensitivity to RTK and PDGFR Inhibitors. *Cancers (Basel)* **14**, (2022).
194. Flynn, R.L., et al., Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. *Science* **347**, 273-7 (2015).
195. Koneru, B., et al., ALT neuroblastoma chemoresistance due to telomere dysfunction-induced ATM activation is reversible with ATM inhibitor AZD0156. *Sci Transl Med* **13**, (2021).
196. Laroche-Clary, A., et al., ATR Inhibition Broadly Sensitizes Soft-Tissue Sarcoma Cells to Chemotherapy Independent of Alternative Lengthening Telomere (ALT) Status. *Sci Rep* **10**, 7488 (2020).
197. Stringer, C., et al., Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* **18**, 100-106 (2021).
198. Bahry, E., et al., RS-FISH: Precise, interactive and scalable smFISH spot detection using Radial Symmetry. *bioRxiv* 2021.03.09.434205 (2021).

199. Ben-Shoshan, S.O., et al., Induction of polyploidy by nuclear fusion mechanism upon decreased expression of the nuclear envelope protein LAP2beta in the human osteosarcoma cell line U2OS. *Mol Cytogenet* **7**, 9 (2014).
200. Fidorra, J., et al., Cellular and nuclear volume of human cells during the cell cycle. *Radiat Environ Biophys* **19**, 205-14 (1981).
201. Roukos, V., et al., Cell cycle staging of individual cells by fluorescence microscopy. *Nat Protoc* **10**, 334-48 (2015).
202. Mao, Z., et al., DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* **7**, 2902-6 (2008).
203. Tomaska, L. and J. Nosek, Telomere heterogeneity: taking advantage of stochastic events. *FEBS Lett* **583**, 1067-71 (2009).
204. Wang, S., et al., BRD4 inhibitors block telomere elongation. *Nucleic Acids Res* **45**, 8403-8410 (2017).
205. Lam, F.C., et al., BRD4 prevents the accumulation of R-loops and protects against transcription-replication collision events and DNA damage. *Nat Commun* **11**, 4083 (2020).
206. Akincilar, S.C., et al., Long-Range Chromatin Interactions Drive Mutant TERT Promoter Activation. *Cancer Discov* **6**, 1276-1291 (2016).
207. Zhong, Z.H., et al., Disruption of telomere maintenance by depletion of the MRE11/RAD50/NBS1 complex in cells that use alternative lengthening of telomeres. *J Biol Chem* **282**, 29314-22 (2007).
208. Elfer, K.N., et al., DRAQ5 and Eosin ('D&E') as an Analog to Hematoxylin and Eosin for Rapid Fluorescence Histology of Fresh Tissues. *PLoS One* **11**, e0165530 (2016).
209. Phimister, E., et al., Expression of neural cell adhesion molecule (NCAM) isoforms in neuroblastoma. *J Clin Pathol* **44**, 580-5 (1991).
210. Lipinski, M., et al., Characterization of neural cell adhesion molecules (NCAM) expressed by Ewing and neuroblastoma cell lines. *Int J Cancer* **40**, 81-6 (1987).
211. Moses, L. and L. Pachter, Museum of spatial transcriptomics. *Nat Methods* **19**, 534-546 (2022).
212. Jansky, S., et al., Single-cell transcriptomic analyses provide insights into the developmental origins of neuroblastoma. *Nat Genet* **53**, 683-693 (2021).
213. Kim, J.K. and C.D. Han, Phase Behavior and Phase Transitions in AB- and ABA-type Microphase-Separated Block Copolymers. *Polymer Materials: Block-Copolymers, Nanocomposites, Organic/Inorganic Hybrids, Polymethylenes* **231**, 77-145 (2010).
214. Eberharter, A. and P.B. Becker, Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO Rep* **3**, 224-9 (2002).
215. Taddei, A., et al., Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat Cell Biol* **3**, 114-20 (2001).
216. Muller-Ott, K., et al., Specificity, propagation, and memory of pericentric heterochromatin. *Mol Syst Biol* **10**, 746 (2014).
217. Huang, B., et al., Brd4 coactivates transcriptional activation of NF-kappaB via specific binding to acetylated RelA. *Mol Cell Biol* **29**, 1375-87 (2009).
218. Morin, J.A., et al., Sequence-dependent surface condensation of a pioneer transcription factor on DNA. *Nature Physics* **18**, 271-276 (2022).
219. Muzzopappa, F., et al., Detecting and quantifying liquid-liquid phase separation in living cells by model-free calibrated half-bleaching. *Nat Commun* **13**, 7787 (2022).
220. Irgen-Gioro, S., et al., Fixation can change the appearance of phase separation in living cells. *Elife* **11**, (2022).
221. Schneider, N., et al., Liquid-liquid phase separation of light-inducible transcription factors increases transcription activation in mammalian cells and mice. *Sci Adv* **7**, eabd3568 (2021).
222. Lin, Y., et al., Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* **60**, 208-19 (2015).

223. Lyons, H., et al., Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell* **186**, 327-345 e28 (2023).
224. Popp, A.P., J. Hettich, and J.C.M. Gebhardt, Altering transcription factor binding reveals comprehensive transcriptional kinetics of a basic gene. *Nucleic Acids Res* **49**, 6249-6266 (2021).
225. Shelansky, R. and H. Boeger, Nucleosomal proofreading of activator-promoter interactions. *Proc Natl Acad Sci U S A* **117**, 2456-2461 (2020).
226. Lahnemann, D., et al., Eleven grand challenges in single-cell data science. *Genome Biol* **21**, 31 (2020).
227. Loe, T.K., et al., Telomere length heterogeneity in ALT cells is maintained by PML-dependent localization of the BTR complex to telomeres. *Genes Dev* **34**, 650-662 (2020).
228. Henson, J.D. and R.R. Reddel, Assaying and investigating Alternative Lengthening of Telomeres activity in human cells and cancers. *FEBS Lett* **584**, 3800-11 (2010).
229. Celina, Y.J., et al., Hyperextended telomeres promote C-circle formation in telomerase positive human cells. *bioRxiv* 2023.01.26.525615 (2023).
230. Jeyapalan, J.N., et al., Evidence for alternative lengthening of telomeres in liposarcomas in the absence of ALT-associated PML bodies. *Int J Cancer* **122**, 2414-21 (2008).
231. Benetti, R., et al., Suv4-20h deficiency results in telomere elongation and derepression of telomere recombination. *J Cell Biol* **178**, 925-36 (2007).
232. Slatter, T.L., et al., The alternative lengthening of telomeres pathway may operate in non-neoplastic human cells. *The Journal of pathology* **226**, 509-518 (2012).
233. Jul-Larsen, A., et al., PML-nuclear bodies accumulate DNA in response to polyomavirus BK and simian virus 40 replication. *Exp Cell Res* **298**, 58-73 (2004).
234. Kordon, M.M., et al., PML-like subnuclear bodies, containing XRCC1, juxtaposed to DNA replication-based single-strand breaks. *FASEB J* **33**, 2301-2313 (2019).
235. Boe, S.O., et al., Promyelocytic leukemia nuclear bodies are predetermined processing sites for damaged DNA. *J Cell Sci* **119**, 3284-95 (2006).
236. Azeroglu, B., et al., *Native FISH: A low- and high-throughput assay to analyze the alternative lengthening of telomere (ALT) pathway*, in *Methods in Cell Biology*.(Academic Press (2022).
237. Noer, J.B., et al., Extrachromosomal circular DNA in cancer: history, current knowledge, and methods. *Trends Genet* **38**, 766-781 (2022).
238. Lau, L.M., et al., Detection of alternative lengthening of telomeres by telomere quantitative PCR. *Nucleic Acids Res* **41**, e34 (2013).
239. Chen, Y.Y., et al., The C-Circle Biomarker Is Secreted by Alternative-Lengthening-of-Telomeres Positive Cancer Cells inside Exosomes and Provides a Blood-Based Diagnostic for ALT Activity. *Cancers (Basel)* **13**, (2021).
240. Potts, P.R. and H. Yu, The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. *Nat Struct Mol Biol* **14**, 581-90 (2007).
241. Sieverling, L., et al., Genomic footprints of activated telomere maintenance mechanisms in cancer. *Nat Commun* **11**, 733 (2020).
242. Feuerbach, L., et al., TelomereHunter - in silico estimation of telomere content and composition from cancer genomes. *BMC Bioinformatics* **20**, 272 (2019).
243. Perrem, K., et al., Coexistence of alternative lengthening of telomeres and telomerase in hTERT-transfected GM847 cells. *Mol Cell Biol* **21**, 3862-75 (2001).
244. Ravindranathan, A., et al., Preliminary development of an assay for detection of TERT expression, telomere length, and telomere elongation in single cells. *PLoS One* **13**, e0206525 (2018).
245. Dilley, R.L., et al., Break-induced telomere synthesis underlies alternative telomere maintenance. *Nature* **539**, 54-58 (2016).

246. Özer, Ö., et al., Human cancer cells utilize mitotic DNA synthesis to resist replication stress at telomeres regardless of their telomere maintenance mechanism. *Oncotarget* **9**, 15836-15846 (2018).
247. Grobelny, J.V., A.K. Godwin, and D. Broccoli, ALT-associated PML bodies are present in viable cells and are enriched in cells in the G(2)/M phase of the cell cycle. *J Cell Sci* **113 Pt 24**, 4577-85 (2000).
248. Chen, Y.C., et al., Live cell dynamics of promyelocytic leukemia nuclear bodies upon entry into and exit from mitosis. *Mol Biol Cell* **19**, 3147-62 (2008).
249. Schraivogel, D., et al., High-speed fluorescence image-enabled cell sorting. *Science* **375**, 315-320 (2022).
250. Chen, W., et al., Live-seq enables temporal transcriptomic recording of single cells. *Nature* **608**, 733-740 (2022).
251. Caudron-Herger, M., et al., Alu element-containing RNAs maintain nucleolar structure and function. *EMBO J* **34**, 2758-74 (2015).
252. Toth, K.F., et al., Trichostatin A-induced histone acetylation causes decondensation of interphase chromatin. *J Cell Sci* **117**, 4277-87 (2004).
253. RStudio Team, RStudio: Integrated Development for R. *RStudio, PBC, Boston, MA*. URL <http://www.rstudio.com/> (2020).
254. Schindelin, J., et al., Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-82 (2012).
255. Trojanowski, J., et al., Light-Induced Transcription Activation for Time-Lapse Microscopy Experiments in Living Cells. *Methods Mol Biol* **2038**, 251-270 (2019).
256. Livak, K.J. and T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-8 (2001).
257. Berthold, M.R., et al., *KNIME: The Konstanz Information Miner*, in *Data Analysis, Machine Learning and Applications. Studies in Classification, Data Analysis, and Knowledge Organization*, C. Preisach, et al., Editors.(Springer: Berlin, Heidelberg). p. 319-326 (2008).
258. Dries, R., et al., Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol* **22**, 78 (2021).

Acknowledgements

I would like to express my gratitude to everyone who has supported me along the path of my PhD journey and beyond.

First of all, I would like to thank Karsten Rippe for giving me the opportunity to spend several wonderful years in his group. Thank you, Karsten, for the great scientific discussions, the open environment and the everlasting support. I really appreciate that you always put so much faith in my decisions, ideas and skills and guided them with your experience and advice when it was most needed. I had all the scientific freedom I could ever wish for. This led to various project detours, none of which I would ever want to miss. They taught me a lot about how science works, and which aspects of it can and cannot be controlled. I will undoubtedly carry with me your little credo "*It's not a bug, it's a feature!*" in the years to come. This simple lesson on creative scientific thinking is an important takeaway from my time here.

Further, I would like to thank my TAC members, Prof. Sylvia Erhardt, Prof. Brian Luke and Prof. Darjus Tschaharganeh for the continuous support of my project, the constructive advice and the lively discussions. I would additionally like to thank PD Dr. Carl Hermann and Dr. Sevin Turcan for participation on my defense committee.

I thank the DKFZ Imaging Core facility, the Heidelberg University Advanced Biological Screening Facility and the Single-Cell Open Lab of the DKFZ for their technical support. I would like to especially thank Philipp, Vuong and Kathi for being there, especially in times where instruments were not doing what they were supposed to do. Furthermore, I would like to thank our collaboration partners PD Dr. Frank Westermann, Sabine Stainczyk and Dr. Priya Chudasama for enabling me to work with patient samples.

A huge thank to all the members, alumni and associates of the Rippe Lab: Afzal, Alexandra, Anne, Armin, Caro, Chloe, Delia, Emma, Fabian, Isabelle, Inn, Irene, Jorge, Linda, Lara, Markus, Michele, Nick, Norbert, Robin, Sabrina, Simon, Sofie, Stephan. Also, I would like to thank all the motivated students that have worked with me during my PhD time: Irene, Samira, Maria and Jana.

I would like to thank a few members and alumni in particular: Fabian, who was a great mentor and friend during my first thesis years. You taught me to put trust in my own ideas, look at things from a different angle and sometimes take risks. Delia, you were one of the kindest persons in my time here, which made my start in the group really easy. I hope that some of your organizational talent has passed to me. At least, we already share the excessive coffee consumption. Thank you, Jorge, for the fun team effort on the phase separation project and the times of friendship. You always managed to keep a cool head and this was needed in

times of great stress (paper revisions, broken microscopes, failed experiments). Furthermore, your natural talent for teaching introduced me into the world of image analysis and programming, which was and is now crucial for my scientific future. Thank you, Anne, for your everlasting support. I have learned a lot from the precision and pragmatism you apply to do science and the ways you take initiative to tackle problems on all levels. Thank you, Norbert, for your technical view on things, the lively discussions and your help. You have more than once convinced me to keep on optimizing a method. Furthermore, I would like to thank Robin, Markus, Jorge, Emma, Simon and Isabelle in particular for all the nice times we had together and your advice and help. I enjoyed our coffee breaks, Friday beer hours and numerous culinary excursions (good, bad, and sometimes ugly). I am really glad to have worked with you. Special thanks to Robin. Your humor kept me going through stressful times and I enjoyed our chats on phase separation and on those who preach it from their (flying) ivory towers. Lastly, I would like to thank Caro and Sabrina for all the amazing help throughout the years, for preventing the lab from going into chaos (educating the kids), and for the fun chats and gossip in the early morning hours.

Beyond the lab, I would like to thank all the people who made my time in Heidelberg special. To my dear friends Kai, Elisa, Delia, Philipp, Franzi, David and Katha, with whom I shared wonderful dinners, gaming nights, escape rooms, hiking trips, big and small parties, whisky and holidays in the sun. These moments were much needed to balance the lab work and I could always count on you. Most importantly, I want to thank Janina from the bottom of my heart. You being there for me in all situations lightened my day and we have built a wonderful life together in Heidelberg. Besides sharing the mind of a scientist, you reminded us to take care of ourselves and always found smart solutions for a brighter present and future. Finally, I would like to thank my family, especially my parents. Your constant support, help and understanding gave me the freedom to do what I love and a home where I was always welcome (and more than well-fed of course). Ultimately, I would like to thank my grandparents. Unfortunately, they will not be able to witness my graduation, but they were always excited about me trying to explain to them what I was actually doing.