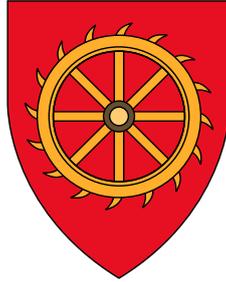


St Catharine's College



**Single-molecule microscopy reveals  
details of heterochromatin  
reorganisation at the onset of  
mouse ES cell differentiation**

**Aleksandra Jartseva**

Department of Biochemistry

University of Cambridge

**Supervisor: Prof Ernest Laue**

*This dissertation is submitted for the degree of*

***Doctor of Philosophy***

December 2021



# Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. It does not exceed the word limit of 60,000 words prescribed for the Faculty of Biology Degree Committee.

Aleksandra Jartseva

December 2021



# Abstract

## **Single-molecule microscopy reveals details of heterochromatin reorganisation at the onset of mouse ES cell differentiation**

Aleksandra Jartseva

In the last decade, spatial organisation of the genome has been recognised as one of the important factors linked to gene expression and cell identity. Although we have come a long way in describing how chromatin folds – from the level of single nucleosomes all the way to compartments and chromosomal territories – it is still unclear what determines this structure and why it changes upon differentiation. Mouse embryonic stem cells provide an excellent system to address these questions, since upon exit from naive pluripotency their genome undergoes a dramatic structural reorganisation.

The development of new fluorophores and imaging technologies has enabled researchers to observe single biological molecules, in both fixed and live cells. These approaches have already proven themselves to be extremely valuable in the chromatin field, for example by revealing the existence and dynamics of chromatin “blobs” (e.g. Ricci et al 2015, Barth et al 2020). This project employs the single-molecule microscopy methods to study chromatin structure and dynamics in ES cells and at the onset of their differentiation.

Research conducted in the Laue lab suggested that heterochromatin might play a key role in the global rearrangement of chromatin as ES cells differentiate. Thus, the first part of the thesis focused on studying the role of heterochromatin protein 1 $\beta$  (HP1 $\beta$ ) in this process. HP1 $\beta$  was found to increase in expression as cells exit from the

ground state, and to change its distribution. Interestingly, this was concomitant with a decrease in both histone H3K9 and H3K27 methylation. 2D and 3D single-particle tracking in live cells was then employed to study HP1 $\beta$  and chromatin dynamics. By applying novel biophysical analysis methods to the data, features of HP1 $\beta$  diffusion not consistent with the current models of heterochromatin compartmentalisation were found, as well as evidence for non-equilibrium processes. Transient slowdown in HP1 $\beta$  diffusion upon differentiation was observed, potentially reflecting a change in the nuclear environment. However, analysis of small-scale chromatin movement showed no change in dynamics between the cell states.

In the second part of this work, PALM- and PAINT-based approaches to image chromatin-binding proteins and DNA in 3D throughout the entire nuclear volume of fixed cells were developed. The microscopy protocol was designed to be compatible with single-cell Hi-C, enabling modeling of the genome structures of the imaged cells (Stevens et al 2017). Correlating the images and the structures will for the first time make it possible to obtain spatial sequence-specific information about the genome fold alongside protein localisation on a single-cell level. I hope that this method will shed light on the molecular functions of chromatin-binding proteins in their native context and help understand the rules of genome organisation.

# Acknowledgements

First and foremost, I would like to thank my lab, who were always there for me: Wei Zhang, Jenny Balmer and David Lando for their advice in the lab and our lovely lunch conversations; Wayne Boucher for his help on the computing front; my wonderful academic siblings Devina Shah, Dominic Hall and Ziwei Zhang for both the laughs and the serious scientific discussions we have had, as well as for sharing with me the constant feeling of confusion that PhD life unavoidably fills one with. I also thank the former members of the lab Srinjan Basu and Aleks Ponjavic for introducing me to the wonderful world of microscopy and giving me all the help I needed whenever I needed it. Finally, I would like to wholeheartedly thank my supervisor Ernest Laue, who is simply one of the kindest, smartest and most wonderful people I have ever met.

I am especially grateful to Aleks, David and Ernest for their comments on this thesis.

My thanks also go to our collaborators: David Klenerman for sharing his infinite microscopy wisdom, as well as David Holcman and Ofir Shukron for development of the data analysis pipelines. I would like to thank the members and ex-members of the Hendrich lab: Brian Hendrich for his always sharp advice, Nicola Reynolds for her expert help with absolutely any lab protocol, Robin Floyd for constructing the mESC lines, and Seun Ogundele and Andria Koulle for the fun times we have spent together. In addition, I am grateful to the members of my graduate thesis panel, Kevin Chalut, Jose Silva and Andy Bannister, for their advice.

I also want to acknowledge Grazvydas Lukinavičius and Jonas Bucevicius for the 5-SiR-Hoechst dye, the Lavis lab for the Janelia Fluor dyes, and the Gribnau, Ferguson-

Smith and Bühler labs, as well as Maike Wiese, for the mESC lines. I thank Joana Cerveira at the Cambridge University Flow Sorting Facility and Martin Lenz at the Cambridge Advanced Imaging Centre for providing equipment, training and help with experiments, as well as the Cambridge Stem Cell Institute Tissue Culture Facility team for reagents and mycoplasma testing.

I am also grateful to the Departments of Biochemistry and Chemistry, Cambridge Stem Cell Institute, my College and the University, as well as the management team behind my Doctoral Training Programme, as well as the UK taxpayers, represented by the Biology and Biotechnology Research Council, for the opportunity to conduct this research.

Last but not least, I would like to thank my friends, my family – my parents Jevgeni and Elena, and my siblings Andrei, Ivan and Maria – as well as my boyfriend Alexander for their endless love and support.

# Contents

<b>Glossary</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Genome organisation . . . . .	2
1.2.1 Beads on a string . . . . .	2
1.2.2 Tools for probing genome organisation . . . . .	5
1.2.3 Higher-order folding of the genome . . . . .	7
1.3 Structural rearrangement of the genome of mESCs upon exit from naive pluripotency . . . . .	8
1.3.1 The pluripotency spectrum . . . . .	8
1.3.2 Chromatin in the formative state . . . . .	12
1.3.3 DNA methylation and the HP1 network are required for the re-organisation . . . . .	15
1.4 HP1 and models of heterochromatin organisation . . . . .	17
1.4.1 Structure and interactions of HP1 . . . . .	17
1.4.2 Activity and dynamics of HP1 on chromatin . . . . .	23
1.4.3 Models of heterochromatin structure . . . . .	29
1.4.4 Functions of HP1 proteins . . . . .	33
1.5 Single-molecule localisation microscopy (SMLM) . . . . .	37
1.5.1 Super-resolution imaging with SMLM . . . . .	37
1.5.2 Live SPT with photoactivatable dyes . . . . .	40

1.5.3	3D single-molecule microscopy with DH-PSF . . . . .	41
1.6	Aims . . . . .	42
<b>2</b>	<b>Investigation of heterochromatin features during exit of mESCs from naive pluripotency</b>	<b>43</b>
2.1	Introduction . . . . .	43
2.2	Characterisation of the HP1 $\beta$ -Halo(N) cell line . . . . .	43
2.2.1	Western blots . . . . .	44
2.2.2	Protein localisation and function . . . . .	45
2.2.3	Cell differentiation . . . . .	46
2.2.4	Separating the populations of HP1 $\beta$ -Halo(N) with differential HP1 $\beta$ expression . . . . .	48
2.3	Differentiation in the absence of HP1 isoforms . . . . .	51
2.4	Levels and distribution of heterochromatic factors during differentiation	54
2.4.1	Heterochromatin decondensation upon exit from naive pluripotency . . . . .	54
2.4.2	HP1 $\beta$ levels and pattern change during differentiation . . . . .	56
2.4.3	No changes are apparent in the HP1 $\alpha$ distribution in the formative state . . . . .	59
2.4.4	H3K9me3 and H3K27me3 staining is not consistent with a replacement hypothesis . . . . .	61
2.5	Discussion . . . . .	64
2.5.1	Changes in the mechanical properties of the nuclei might be linked to chromatin structure . . . . .	65
2.5.2	Reorganisation of chromocentres . . . . .	67
2.5.3	Future work . . . . .	72
<b>3</b>	<b>Studies of HP1<math>\beta</math> dynamics with single-molecule imaging</b>	<b>75</b>
3.1	Introduction . . . . .	75
3.1.1	Single-molecule features of phase separation . . . . .	76
3.1.2	Experimental design . . . . .	79

3.2	Short-exposure sptPALM . . . . .	81
3.2.1	Method optimisation . . . . .	82
3.2.2	Data analysis overview . . . . .	86
3.2.3	Improvements and testing of the 4P algorithm . . . . .	88
3.2.4	HP1 $\beta$ exhibits different behaviour in eu- and heterochromatin . . . . .	95
3.2.5	HP1 $\beta$ at the inter-compartment boundary . . . . .	104
3.2.6	HP1 $\beta$ displays slower dynamics in formative-state mESCs . . . . .	108
3.2.7	Summary . . . . .	108
3.3	Analysis of HP1 $\beta$ mobility with FLIP . . . . .	110
3.3.1	Overview . . . . .	110
3.3.2	Exchange between eu- and heterochromatin is marginally slowed down . . . . .	112
3.4	Diffusion of chromatin-bound molecules and residence time analysis . . . . .	113
3.4.1	Method optimisation . . . . .	114
3.4.2	Dynamics of HP1 $\beta$ -bound chromatin does not change upon dif- ferentiation . . . . .	116
3.4.3	HP1 $\beta$ has at least two modes of chromatin binding . . . . .	118
3.5	Discussion . . . . .	120
3.5.1	Models of heterochromatin organisation . . . . .	122
3.5.2	Differentiation . . . . .	129
3.5.3	Conclusion . . . . .	133
<b>4</b>	<b>Development of whole-nucleus 3D SMLM in combination with single- nucleus Hi-C</b>	<b>137</b>
4.1	Introduction . . . . .	137
4.1.1	General workflow . . . . .	138
4.1.2	Imaging: 3D DH-PSF SMLM . . . . .	139
4.2	Data processing pipeline . . . . .	140
4.2.1	3D localisation detection and determination of coordinates . . . . .	140
4.2.2	Duplicate removal . . . . .	146

4.3	Optimisation of 3D protein SMLM . . . . .	146
4.3.1	Overview . . . . .	146
4.3.2	Sample preparation . . . . .	148
4.3.3	Imaging optimisation . . . . .	150
4.4	Optimisation of 3D DNA PAINT . . . . .	152
4.4.1	Overview . . . . .	152
4.4.2	Sample preparation . . . . .	153
4.4.3	Imaging optimisation . . . . .	156
4.5	Whole-nucleus imaging . . . . .	156
4.5.1	Imaging process . . . . .	156
4.5.2	Accuracy of stitching: beads control . . . . .	160
4.6	Drift correction . . . . .	164
4.6.1	Evaluation of drift . . . . .	164
4.6.2	Drift correction with bright-field image cross-correlation . . . . .	165
4.7	Trial whole-nucleus 3D DH-PSF SMLM imaging of DNA, transcription factors and HP1 . . . . .	168
4.7.1	DNA PAINT . . . . .	168
4.7.2	SOX2 PALM . . . . .	170
4.7.3	NANOG PALM . . . . .	172
4.7.4	HP1 $\beta$ PALM . . . . .	174
4.8	Discussion and future directions . . . . .	176
4.8.1	DNA SMLM . . . . .	177
4.8.2	Accuracy of DH-PSF . . . . .	178
4.8.3	Image completeness and resolution . . . . .	180
4.8.4	Clustering analysis . . . . .	182
4.8.5	Alternatives to 3D DH-PSF SMLM . . . . .	183
4.8.6	Image and snHi-C structure correlation . . . . .	184

**5 Conclusion** **185**

<b>6</b>	<b>Materials and Methods</b>	<b>189</b>
6.1	Materials . . . . .	189
6.1.1	Cell lines . . . . .	189
6.1.2	Tissue culture media and supplements . . . . .	190
6.1.3	Antibodies . . . . .	190
6.1.4	Fluorophores . . . . .	192
6.1.5	Primers and TaqMan probes . . . . .	192
6.2	Characterisation of cell lines . . . . .	193
6.2.1	Cell culture . . . . .	193
6.2.2	Genotyping . . . . .	194
6.2.3	Western blotting . . . . .	194
6.2.4	Alkaline phosphatase assay . . . . .	195
6.2.5	RT-qPCR . . . . .	196
6.2.6	Flow cytometry and cell sorting . . . . .	197
6.3	Confocal imaging . . . . .	199
6.3.1	Sample preparation . . . . .	199
6.3.2	Microscopy . . . . .	200
6.3.3	Image analysis . . . . .	201
6.4	SptPALM of HP1 $\beta$ . . . . .	202
6.4.1	Sample preparation . . . . .	202
6.4.2	Microscopy . . . . .	202
6.4.3	Controls . . . . .	204
6.4.4	Primary image analysis . . . . .	205
6.4.5	Short-exposure SPT data analysis . . . . .	208
6.4.6	Long-exposure SPT data analysis . . . . .	215
6.5	FLIP . . . . .	216
6.5.1	Experimental protocol . . . . .	216
6.6	3D DH-PSF SMLM of fixed cells, combined with snHi-C . . . . .	217
6.6.1	Sample preparation . . . . .	217
6.6.2	Microscopy . . . . .	219

6.6.3	Data analysis . . . . .	220
6.6.4	Control experiments with beads . . . . .	224
<b>References</b>		<b>225</b>

# Glossary

**2i/LIF** two inhibitors + LIF

**5mC** cytosine methylation at position 5

**ac** acetylation

**BIC** Bayesian information criterion

**BSA** bovine serum albumin

**CD** chromodomain

**ChIP-seq** chromatin immunoprecipitation and sequencing

**CSD** chromoshadow domain

**CTCF** CCCTC-binding factor

**CTE** C-terminal extension

**DH-PSF** double-helix point spread function

**DNMT** DNA methyltransferase

**EpiSC** epiblast stem cell

**ESRRB** estrogen-related receptor beta

**FACS** fluorescence-activated cell sorting

**FGF5** fibroblast growth factor 5

**FLIP** fluorescence loss in photobleaching

**fps** frames per second

**FRAP** fluorescence recovery after photobleaching

**FSC** forward scatter

**GMM** Gaussian mixture model

**H3K27me3** histone H3 lysine 27 trimethylation

**H3K9me2/3** histone H3 lysine 9 di- or trimethylation

**Hi-C** high-throughput chromosome conformation capture

**HILO** highly inclined and laminated optical sheet

**HP1** heterochromatin protein 1

**HRP** horseradish peroxidase

**JF<sub>549/646</sub>** Janelia Fluor 549/646

**KAP1** KRAB-associated protein 1

**KLF4** krüppel-like factor 4

**LIF** leukaemia inhibitory factor

**LLPS** liquid-liquid phase separation

**LSPS** liquid-solid phase separation

**MBD** methyl-CpG-binding domain

**me** methylation

**MEA** mercaptoethylamine

**MeCP2** methyl-CpG binding protein 2

**mESC** mouse embryonic stem cell

**MSD** mean squared displacement

**NA** numerical aperture

**NTE** N-terminal extension

**OTX2** orthodenticle homeobox 2

**PA-JF<sub>549/646</sub>** photoactivatable Janelia Fluor 549/646

**PAINT** point accumulation for imaging in nanoscale topography

**PALM** photoactivated localisation microscopy

**PBS** phosphate buffer saline

**pho** phosphorylation

**PRC1/2** polycomb repressive complex 1 or 2

**PSF** point-spread function

**PTM** post-translational modification

**RNAPII** RNA polymerase II

**RT-qPCR** reverse transcription-quantitative PCR

**SMLM** single-molecule localisation microscopy

**snHi-C** single-nucleus Hi-C

**SNP** single-nucleotide polymorphism

**SPT** single-particle tracking

**SSC** side scatter

**TAD** topologically associating domain

**ub** ubiquitination



# Chapter 1

## Introduction

### 1.1 Introduction

Nearly every living cell contains a genome – the information store of protein recipes in the form of an extremely long and thin string (or strings) of DNA. It must be packaged in an efficient way that allows a drastic reduction in its size, but which maintains access to this information. The robust, but adaptable folding of the genome is crucial to maintain genome integrity, to control replication and transcription, to allow cell division, to silence endogenous retroviruses – in short, to survive.

Although generally, every cell of an animal contains the information needed to synthesise the complete set of proteins, which proteins are made at what time determines the type of cell, its morphology, the functions that it can perform, the stage of the cell cycle, its history and its response to the environment. Every step of the process is thus regulated to allow for maximum flexibility. Genome folding is the first layer in this hierarchy. During development, cells undergo drastic changes, and this is reflected in the physical structure of their genome.

In this work, the principles underlying the organisation of heterochromatin, the condensed and less active part of the genome, have been studied in the context of early mammalian development.

## 1.2 Genome organisation

### 1.2.1 Beads on a string

#### The nucleosome

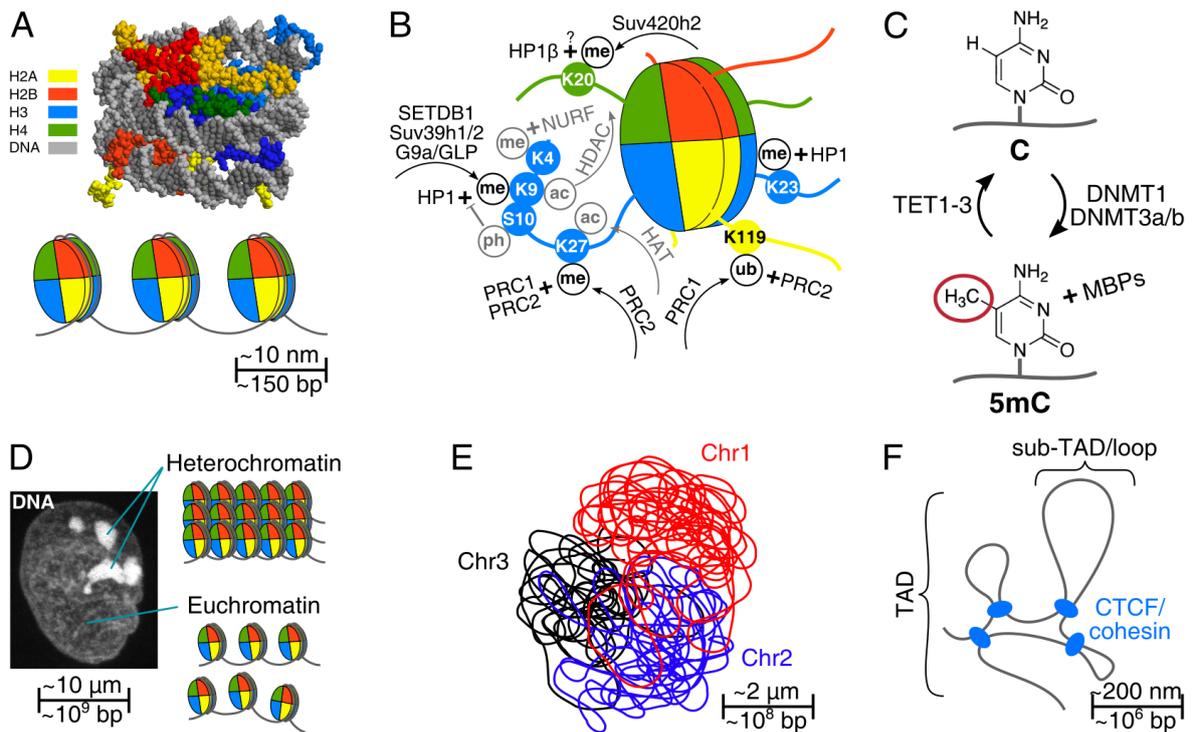
DNA in cells is associated with positively charged histone proteins, forming a substance known as chromatin. The basic unit of chromatin – the nucleosome – contains two copies of each of the four histone proteins H2A, H2B, H3 and H4, and 146 bp of DNA are wrapped twice around the octamer (Figure 1.1A). Since nucleosomes are separated by stretches of linker DNA, this level of chromatin organisation is referred to as “beads on a string”. Due to the strong attraction between histones and DNA, nucleosomes impede most DNA-associated processes such as transcription. Nucleosome positioning can be modulated by chromatin remodellers, which can slide or evict them.

#### The histone code

The terminal unstructured regions of histones, called “histone tails”, protrude away from the nucleosome (Figure 1.1B). They are highly charged and interact with the linker regions between nucleosomes (Hansen et al. 2021). They also harbour numerous sites subject to post-translational modifications (PTMs), such as acetylation (ac), methylation (me), phosphorylation (pho) and ubiquitination (ub). These modifications have remarkably complex and combinatorial effects, and are referred to as “the histone code”.

Some of the PTMs can alter the local chromatin structure directly. This is the case for lysine acetylation, which can occur at a large number of sites. Acetylation weakens the interaction between DNA and histones by neutralising the positive charge of lysines and reducing tail-linker interactions, thus making the DNA more accessible. Consistently, acetylation is mostly found in active regions of the genome – e.g. histone H3 lysine27 acetylation (H3K27ac) is a well-known marker of active enhancers and promoters.

Furthermore, many modifications serve as binding sites for so-called “reader” pro-



**Figure 1.1:** A. Top, structure of a nucleosome, consisting of a histone octamer and 147 bp of DNA wrapped around it twice. The structure was adapted from Wikimedia Commons (user Darekk2), PDB 1EQZ (Harp et al. 2000). Bottom, schematic of the “beads on a string” level of chromatin organisation. B. Illustration of the histone code. Some relevant activating (gray) and repressive (black) chromatin marks are shown, with the respective writer (arrow) and reader (plus sign) proteins. Number of methyl groups is omitted for simplicity. C. Cytosine methylation at position 5 (5mC) is catalysed by DNA methyltransferases and erased by the Tet family demethylases. 5mC is bound by MBD proteins and MeCP2. D. On a large scale, chromatin is organised into loosely-packed and active euchromatin, and tightly-packed and gene-poor heterochromatin. E. Chromosomes typically occupy separate “territories”. F. At intermediate scales, chromatin folds into topologically-associated domains (TADs), which are sometime subdivided into sub-TADs/loops. In A and D-F, the approximate relative scales of genome structure are shown.

teins (Figure 1.1B). Histone H3 lysine 9 di- or trimethylation (H3K9me<sub>2/3</sub>) and histone H3 lysine 27 trimethylation (H3K27me<sub>3</sub>) modifications, markers of inactive chromatin, interact with chromodomain-containing proteins: the former with heterochromatin protein 1 (HP1) (see section 1.4) (Lachner et al. 2001; Bannister et al. 2001), and the latter with components of polycomb repressive complex 1 (PRC1) (Fischle et al. 2003). While both inhibit transcription and modulate chromatin structure, H3K9me<sub>2/3</sub> and HP1 are generally associated with constitutive heterochromatin, and H3K27me<sub>3</sub> and polycomb with facultative heterochromatin. On the other hand, H3K4me<sub>3</sub>, found at active promoters, induces transcription by recruiting the chromatin remodeller NURF (Wysocka et al. 2006).

Histone modifications are deposited by “writer” proteins, which together with reader proteins are often implicated in feedback loops. For example, HP1 recruits H3K9 methyltransferases SETDB1, SUV39h1/2 and G9a through direct and indirect interactions, providing a mechanism for the spreading of the suppressed state. Moreover, there is cross-talk between readers, writers and different histone modifications. In the polycomb network, PRC1 catalyses the ubiquitination of H2AK119, which in its turn recruits polycomb repressive complex 2 (PRC2) that contains a H3K27 methyltransferase and spreads this mark to the neighbouring nucleosomes (Chittock et al. 2017).

## DNA methylation

DNA can also be modified in several ways, by far the most common of which is cytosine methylation at position 5 of the pyrimidine ring (5mC, Figure 1.1C). 5mC is considered a repressive mark, since its readers usually suppress transcription.

DNA methyltransferases come in two flavours. Maintenance DNA methyltransferase DNMT1 ensures that methylation is replicated on the newly synthesised strand upon DNA replication. In contrast, *de novo* DNA methyltransferases DNMT3a and DNMT3b are able to modify genomic loci anew. DNA demethylation, most commonly, proceeds through a step-by-step oxidation of the 5-methyl group by the ten-eleven translocation (TET) family of proteins, followed by excision of the modified cytosine and base excision repair.

DNA methylation is intimately linked with the H3K9me2/3-based repression. 5mC readers methyl-CpG-binding domain 1 (MBD1) and methyl-CpG binding protein 2 (MeCP2), as well as DNMT proteins, interact with H3K9 methyltransferases SUV39h1/2 and SETDB1, and DNMTs also bind HP1. Furthermore, DNMT1 adaptor protein UHRF1 binds H3K9me2/3, which is important for DNMT1 targeting (Rose and Klose 2014).

### 1.2.2 Tools for probing genome organisation

Biochemical, electron microscopy and structural biology studies of purified chromatin have played crucial roles in understanding its organisation on the smallest scale. However, other methods are needed to unravel how nucleosomes are further organised in the context of cellular nuclei.

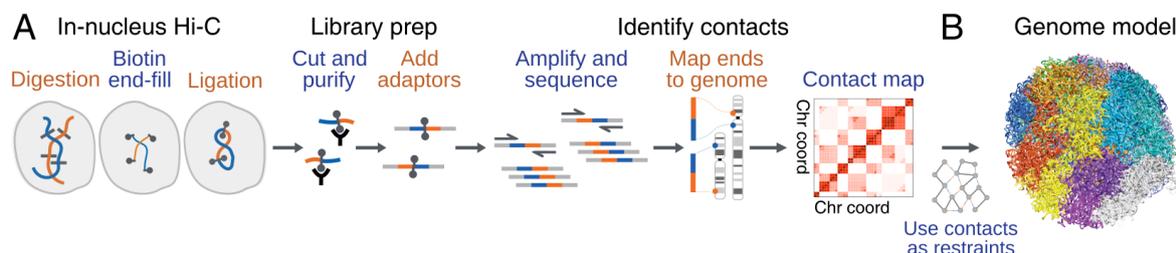
#### Microscopy

Perhaps the oldest method that has been used in this context is light microscopy that was employed to visualise chromosomes. More recently, fluorescence microscopy following labelling of DNA, histones, other chromatin-binding proteins or RNA allowed researchers to deduce some general principles of nuclear organisation, while site-specific labelling methods such as fluorescence *in situ* hybridisation (FISH) provided insights into its functional significance. The advent of super-resolution and single-molecule imaging methods has brought microscopy to a new level, allowing one to probe chromatin structure in native systems on the nanometer scale.

#### Chromatin conformation capture

The limitations of microscopy in studying genome organisation are the lack of sequence-specific information and its low throughput. Complementary sequencing methods collectively known as chromatin conformation capture (3C) have emerged in the last 20 years, which has allowed exploration of the relationship between sequence and structure (Figure 1.2A). In 3C, nuclei are fixed, the DNA is cut into fragments, and the

cross-linked fragments are ligated together. In Hi-C – the genome-wide, all-to-all type of 3C – all the junctions between the sequences that must have been in close spatial proximity in the nucleus are sequenced and mapped to the reference genome.



**Figure 1.2:** A. Hi-C protocol. B. If Hi-C is performed on single cells, contacts can be used as distance restraints in a simulated annealing protocol to build genome models. Figure adapted from Stevens et al. (2017).

A typical Hi-C experiment is performed in bulk on millions of cells, yielding population-averaged information. This allows one to pick up recurrent contacts across many cells, as well as achieve a high resolution of up to 750 bp (Bonev et al. 2017). On the other hand, capturing contacts on a single-cell level can reveal biological variation between cells and link features that co-occur in the same cells (Nagano et al. 2013). Furthermore, using single-nucleus Hi-C (snHi-C) contacts as distance restraints, one can model genome structures of individual cells (Figure 1.2B) (Nagano et al. 2013; Stevens et al. 2017).

## Mapping epigenetic features

On top of the mostly large-scale, spatial information provided by microscopy and Hi-C, several methods for mapping local epigenetic features onto the genome exist. Micrococcal nuclease digestion and sequencing (MNase-seq) and assay for transposase-accessible chromatin using sequencing (ATAC-seq) probe the accessibility of chromatin to MNase or transposase, revealing the positioning of nucleosomes and other protein complexes. Immunoprecipitating the protein or histone mark of interest and sequencing the associated DNA (termed chromatin immunoprecipitation and sequencing (ChIP-seq)) can help determine their distribution on the genome. The most common way to probe the DNA methylation landscape is bisulfite sequencing, which takes advantage of the fact

that methylation protects cytosine conversion to uracil upon bisulfite treatment.

### 1.2.3 Higher-order folding of the genome

#### Compartments

Let us now zoom out about three orders of magnitude from the 10 nm-thick “beads on a string” to the entire nucleus. Staining of the DNA in the majority of eukaryotic cell types allows one to observe brighter and dimmer regions, termed hetero- and euchromatin, corresponding to densely or loosely packed DNA (Figure 1.1D). Heterochromatin mostly contains gene-poor regions and repeats, including telomeres and centromeres, and is associated primarily with H3K9me2/3. It has long been thought that its tight packing makes it inaccessible to other proteins, resulting in transcriptional silencing, but this is currently a matter of debate (see section 1.4). The much more accessible euchromatin, on the other hand, harbours active genes.

Hi-C has also revealed two compartments within the genome, which have more contacts within themselves than between each other. The A compartment is enriched for genes and active transcription marks, while the B compartment is gene-poor. Due to the differences in methodology, it is hard to strictly equate the A and B compartments with eu- and heterochromatin, but they definitely possess very similar features.

Furthermore, chromosomes in most eukaryotic cells tend to occupy separate territories, rather than be entangled and mixed (Figure 1.1E). Escaping from chromosome territories by certain genes has been associated with their activation, although the generality of this phenomenon is debated (Cremer and Cremer 2010).

#### **Nucleosome arrangement: loops, TADs, clutches and blobs**

Hi-C experiments revealed that chromatin is organised into topologically associating domains (TADs) – regions characterised by the loci within them preferentially interacting with each other rather than with adjacent regions (Figure 1.1F) (Dixon et al. 2012). TADs are large loops, whose borders are brought together by CCCTC-binding factor (CTCF) and other proteins, such as cohesin (Pombo and Dillon 2015). Contact density

within TADs is often also non-uniform, harbouring subTADs – smaller regions with a higher density of contacts and looping of their border elements.

From bulk Hi-C experiments, however, it is unclear whether the observed features correspond to physical structures or whether they are merely statistical effects, and if they are found in all cells or only in some. To address these questions, various authors employed super-resolution microscopy to image chromatin in both fixed and live cells (Ricci et al. 2015; Xu et al. 2018; Miron et al. 2020; Shaban et al. 2020; Barth et al. 2020; Mateo et al. 2019), electron microscopy (Ou et al. 2017; Miron et al. 2020), as well as snHi-C (Nagano et al. 2013; Stevens et al. 2017).

These studies report that chromatin is indeed organised into dynamic globules/clutches/blobs, roughly corresponding to subTADs. However, when particular sequences are investigated, the subTADs and TADs found in the Hi-C maps can have very different conformations in different cells (Stevens et al. 2017; Mateo et al. 2019; Miron et al. 2020). Finally, the features of blobs – their size, density, packing and dynamics – vary within the nucleus (e.g. in eu- and heterochromatin) (Barth et al. 2020; Miron et al. 2020), correlate differently with histone marks or presence or histone H1 (Ricci et al. 2015; Xu et al. 2018; Miron et al. 2020), and change upon differentiation (Ricci et al. 2015).

## **1.3 Structural rearrangement of the genome of mESCs upon exit from naive pluripotency**

### **1.3.1 The pluripotency spectrum**

#### **Mouse embryonic stem cells**

Mouse embryos between embryonic days 3.5-4.5 (E3.5-E4.5) adopt a structure known as the pre-implantation blastocyst, consisting of three cell types: the trophoblast, the hypoblast and the epiblast (Figure 1.3, top). During development, the epiblast cells differentiate into all the cell types of the embryo proper, while the tropho- and the

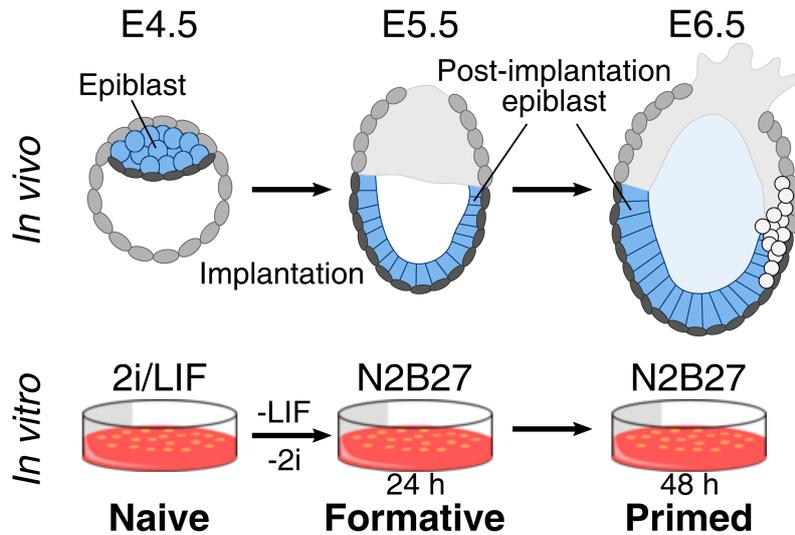
hypoblast give rise to extraembryonic tissues. The term “pluripotency” is used to describe this ability of the epiblast, in contrast to the “multipotency” of adult stem cells with a more restricted potential and to the “totipotency” of the zygote.

Mouse embryonic stem cells (mESCs) are derived from the epiblast of a pre-implantation mouse embryo (Evans and Kaufman 1981; Martin 1981). Unlike other primary cells, under appropriate conditions mESCs can in principle be cultured indefinitely without going into senescence (Miura et al. 2004). Like the epiblast cells in the early embryo, mESCs have an unrestricted developmental potential: upon injection into a recipient blastocyst, they contribute substantially to all the cell lineages of the chimaeric animal (Bradley et al. 1984; Beddington and Robertson 1989). Furthermore, when supplied with appropriate stimuli, they can be differentiated into many cell types in culture (Keller 2005).

Initially, mESCs were cultured using a feeder layer of senescent cells, as well as foetal bovine serum. The critical factor secreted by the feeders keeping mESCs undifferentiated was soon identified to be the leukaemia inhibitory factor (LIF), activating the JAK-STAT3 pathway (Williams et al. 1988; Smith et al. 1988). Finally, a number of serum-free defined media were discovered that support the growth of mESCs, the most widely used of which contains a combination of Chiron and PD03, small-molecule inhibitors of the GSK3 and MEK/ERK signaling pathways, respectively, and LIF (2i/LIF) (Ying et al. 2008).

### **Naive and primed pluripotency**

A different type of pluripotent cells is derived from early post-implantation embryo, E5.5-E8, known as epiblast stem cells (EpiSCs) (Brons et al. 2007; Tesar et al. 2007). EpiSCs can be differentiated into all cell types in culture or in a teratoma assay, retaining the defining features of pluripotency. However, unlike mESCs, when injected into the pre-implantation blastocyst, they contribute little to the tissues of the chimaeric animal. Their molecular identity, termed primed pluripotency, thus corresponds to a later developmental stage, while mESCs are said to be in the ground, or naive pluripotent state.



**Figure 1.3:** The stages of pluripotency. Top: a schematic representation of mouse embryo development around the time of implantation. The epiblast, from which mESCs are derived, is shown in light blue. Bottom: an *in vitro* differentiation protocol, mimicking the transitions occurring in the epiblast *in vivo*.

The molecular hallmarks of naive pluripotency include strong expression of the specific transcription master regulators such as NANOG, estrogen-related receptor beta (ESRRB) and krüppel-like factor 4 (KLF4), a hypomethylated genome, low levels of the H3K27me3 histone mark on the developmental regulators and two active X-chromosomes in female cells (Weinberger et al. 2016). Transcriptionally, mESCs show strong similarity to the E3.75-E4.5 pre-implantation epiblast (Boroviak et al. 2014). Primed cells, on the other hand, express priming factors such as orthodenticle homeobox 2 (OTX2) and fibroblast growth factor 5 (FGF5), have higher global methylation levels due to the activity of DNA methyltransferases, high numbers of bivalent promoters with both H3K27me3 and H3K4me3, and X-chromosome inactivation in females (Weinberger et al. 2016). Their transcriptome resembles the gastrula-stage epiblast, and they also start expressing early lineage-specific markers without yet committing to a particular fate (Kojima et al. 2014).

Ablation of epigenetic repressors, such as DNA methyltransferases or polycomb proteins, also has opposing effects in mESCs and EpiSCs. While the former retain their pluripotent status and even become resistant to differentiation, the latter lose

their pluripotency (Geula et al. 2015). Therefore, the ground-state mESCs can be viewed as “stably” pluripotent, while the primed EpiSCs as poised for differentiation, held back by the epigenetic repressors.

### **The naive-primed transition and formative pluripotency**

While mESCs can be easily differentiated into EpiSCs by culture in the respective medium, when EpiSC are placed into the 2i/LIF conditions they do not revert to the ground state (Guo et al. 2009). Thus, primed pluripotency is a stable downstream state, separated from naive pluripotency by a differentiation roadblock. However, this block can be overcome more easily than in somatic cells – e.g. induced expression of a single Yamanaka factor KLF4 is sufficient for reprogramming (Guo et al. 2009).

Are there any further substates between the naive and primed pluripotency? Firstly, mESCs cultured in different conditions are not exactly the same. In serum/LIF, for example, cells exhibit higher levels of DNA methylation, increased H3K27me3 at regulatory loci and higher transcriptional heterogeneity than those grown in 2i/LIF (Weinberger et al. 2016; Kumar et al. 2014). Therefore, serum/LIF mESCs are considered to be slightly developmentally downstream from the 2i/LIF mESCs. However, the transition between these states is reversible, suggesting that they are not separated by a point of no return.

Recent studies of the naive-primed transition found evidence of a distinct transitional, or formative, pluripotent cell state. *In vivo*, epiblast cells differentiate into the primed state through an ordered sequence of events: first, the naive transcription factor network is extinguished, making cells competent for specification, and only then are early lineage markers upregulated (Smith 2017). This process can be mimicked *in vitro* by first removing LIF, and then triggering the exit from naive pluripotency by withdrawal of the inhibitors (Figure 1.3, bottom) (Kalkan et al. 2017). In both cases, the formative state is defined as cells that have irreversibly lost the naive identity, but have not yet acquired primed characteristics, which occurs *in vivo* around the time of implantation at E5.5, and *in vitro* about 24h after 2i removal (Kalkan et al. 2017). Although formative pluripotency was first defined only as a transient state, there now

exist culture conditions that stabilise it (Kinoshita et al. 2021).

Growing cells without 2i for longer also allows studying the downstream differentiation states (Ying et al. 2003). In this work, cells 48h after the onset of differentiation were used as a proxy for the primed pluripotent state, although they are not exactly equivalent to EpiSCs or gastrula-stage epiblast cells (Betschinger et al. 2013).

### **1.3.2 Chromatin in the formative state**

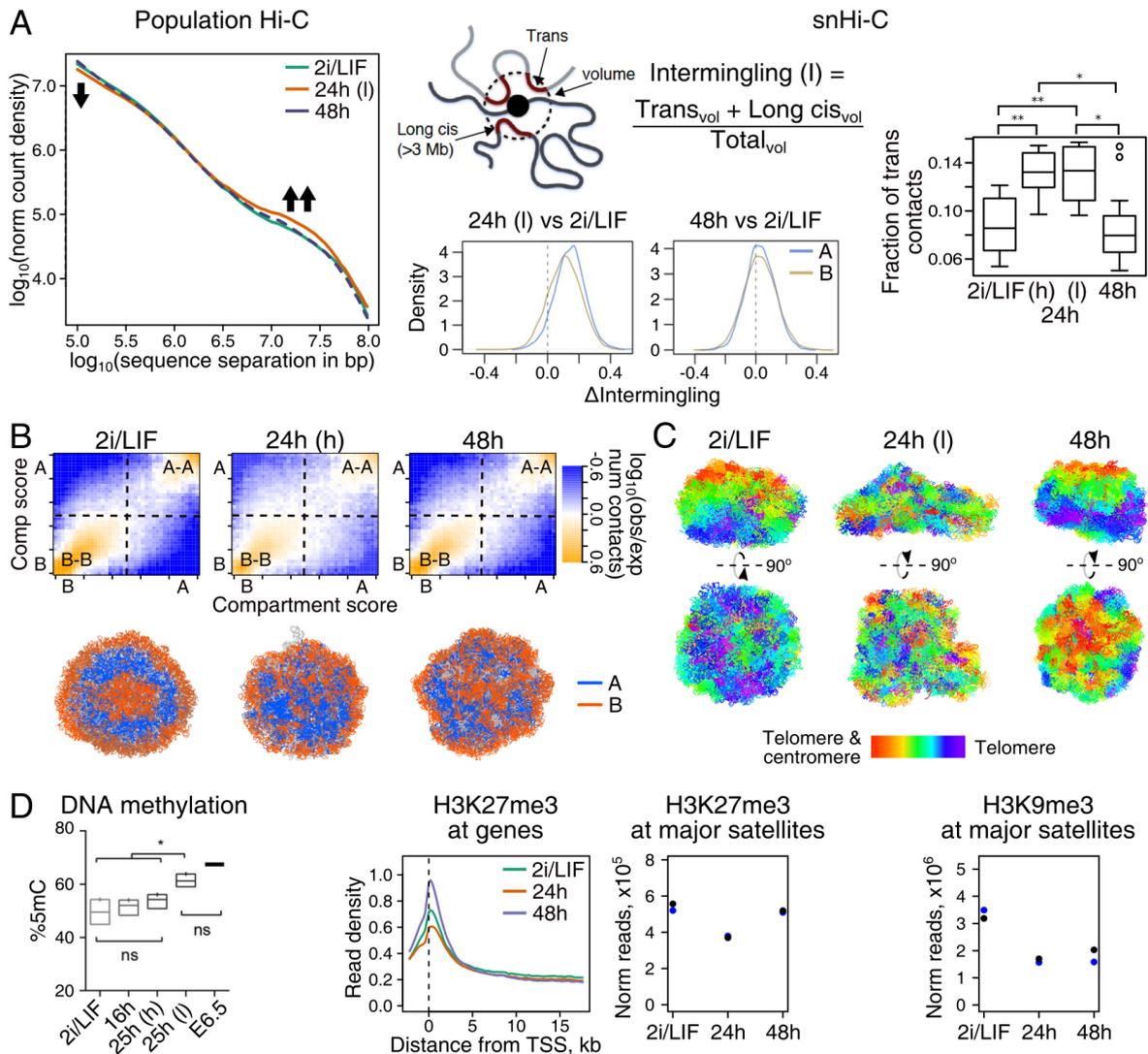
Unexpectedly, investigations into the chromatin conformation during the naive-primed transition, which involves many epigenetic changes, revealed that formative-state mESCs possess a very distinct chromatin structure, whilst the features of the naive and primed state are only slightly different (Laue lab, unpublished).

#### **Intermingling and decondensation**

Firstly, formative-state cells are characterised by a decrease in short-range contacts and a prominent increase in megabase-scale and trans-chromosomal contacts, evident both in bulk and snHi-C measurements (Figure 1.4A). On a single-cell level, this can be described, using the intermingling score – the average proportion of loci (beads) from  $>3$  Mbp away on the same chromosome or on a different chromosome in the vicinity of each locus (Figure 1.4A, middle). The increase in intermingling suggests that a global decondensation of chromatin might take place, leading to every locus being less restricted in the area it is able to sample. Indeed, this has been observed before on the mesoscale (tens-hundreds nm) for formative-state nuclei using electron microscopy (Pagliara et al. 2014).

#### **Weakening of compartment segregation**

Another feature of the formative-state cells is a weaker segregation of the A and B compartments (eu- and heterochromatin), i.e. a higher probability of contacts between them (Figure 1.4B, top). Mapping the compartment identity onto the structures of individual genomes revealed that in mESCs, the compartments are organised into well-



**Figure 1.4:** A. The genome is more intermingled in the formative state. Left, log-log profile of population Hi-C contact lengths for mESCs in 2i/LIF and after 2i removal. Arrows highlight transient changes in the formative state. 24h (h) and (l) refer to cells separated based on the high or low level of Rex1 expression, where 24h (h) can still return to naive pluripotency, while 24h (l) cells are committed to differentiation (Kalkan et al. 2017). Middle, schematic explaining the intermingling score for a single locus in a single cell, and histograms of change in intermingling score for all loci upon differentiation, shown separately for the A and B compartments. Right, fraction of inter-chromosomal contacts in individual cells. B. Compartments are less insulated in the formative state. Top, heatmaps of enrichment of contacts between loci, sorted by compartment score. Bottom, snHi-C genome structures, coloured by compartment. C. SnHi-C genome structures, coloured by position along the chromosome, demonstrating the Rab1 chromosome configuration in 2i/LIF and 48h post 2i removal and lack thereof at 24h. D. Global dynamics of epigenetic marks during exit from naive pluripotency. Left, DNA methylation in mESCs, 16h and 25h after 2i withdrawal, and in embryos at E6.5 (modified from Kalkan et al. (2017)). Middle, levels of H3K27me3 mark on genes and on major satellite repeats (quantitative ChIP-seq). Right, level of H3K9me3 on major satellite repeats (quantitative ChIP-seq). All figures except D, left, are modified from ones made by Dr D. Lando and Dr X. Ma.

defined shells. This organisation is perturbed 24h after 2i removal, leading to fragmentation and mixing of the compartments, and only partially restored 24h later (Figure 1.4B, bottom).

### **Organisation, shape and mechanical properties**

Concomitantly with, or potentially as a result of changes in genome structure, the global organisation of the nucleus, its shape and properties change as well. SnHi-C suggested that the Rab1 configuration of chromosomes, observed in naive mESCs in G1, is significantly perturbed in the formative state, and that the nuclei potentially become flatter (Figure 1.4C). Pagliara et al. (2014) also reported that due to chromatin decondensation, nuclei 24h post differentiation onset exhibit auxeticity.

### **Changes in DNA methylation and histone marks**

As mentioned above, the primed-state cells' genome is substantially more methylated than that of naive cells. This increase in methylation occurs gradually, with formative-state cells having an intermediate level of methylation (Figure 1.4E) (Kalkan et al. 2017). This increase is caused by the expression of all the DNA methyltransferases, and it happens globally, although a small number of regions are resistant to it.

It is generally thought that the H3K9me2/3 mark becomes more abundant as mESCs differentiate, contributing to stronger segregation of constitutive heterochromatin and stable silencing of genes (Nicetto and Zaret 2019). In agreement with this notion, Tosolini et al. (2018) demonstrated that mESCs sometimes harbour H3K27me3-decorated chromocentres, while EpiSCs nearly always have H3K9me3-marked foci. In the early differentiation system described above, H3K27me3 levels decrease transiently and globally 24h after removal of 2i (Figure 1.4E). Interestingly, the spatial clustering of the H3K27me3-bound regions, observed using snHi-C genome structures, was also transiently reduced in the formative state (Laue lab, unpublished). Surprisingly, H3K9me3 levels on major satellite repeats also drop about two-fold at this time point,

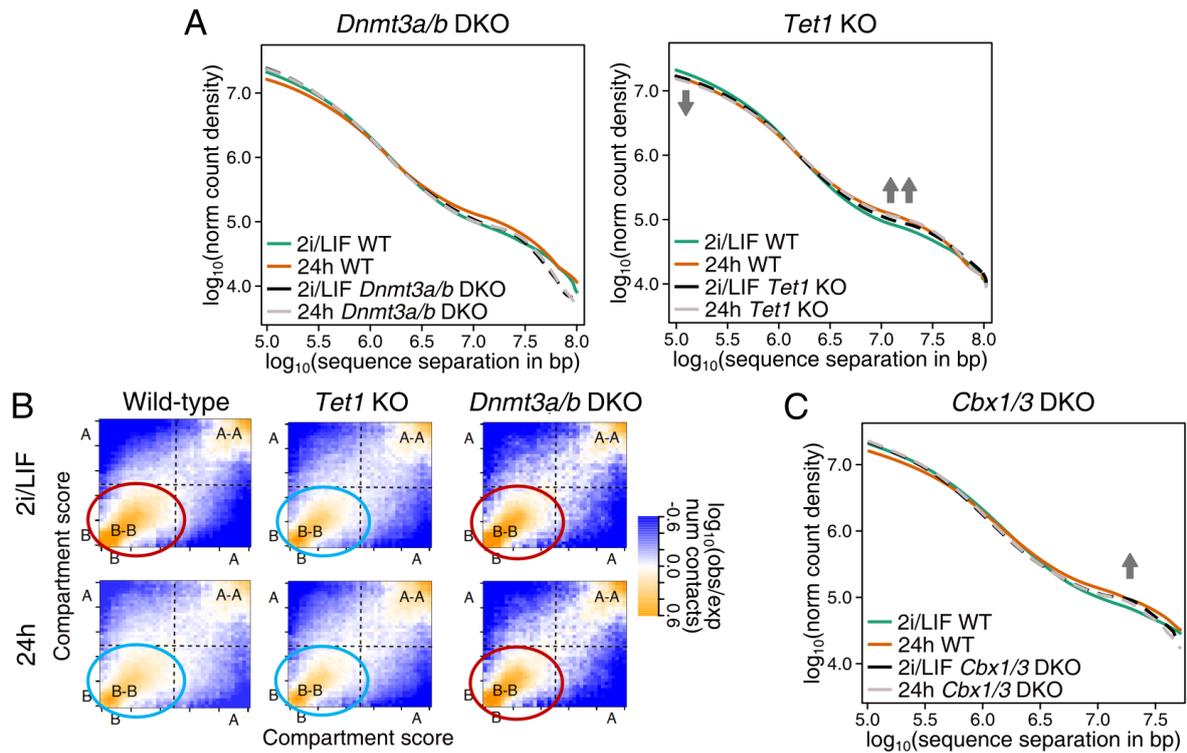
but do not rise again by 48h (Figure 1.4E). Other H3K9me3-decorated genomic regions, such as the endogenous retroviruses or LINE1 elements, generally show a similar trend, although the magnitude of the decrease varies.

### 1.3.3 DNA methylation and the HP1 network are required for the reorganisation

In a search for factors playing key roles in this genome reorganisation, several knockout mESC lines were tested, using the Hi-C contact profiles in 2i/LIF and 24h after 2i withdrawal as readouts.

The most striking results were obtained for the knockouts of genes involved in the DNA methylation pathway. Removal of the *de novo* methyltransferases DNMT3a and DNMT3b resulted in a lack of profile shift upon differentiation (Figure 1.5A). On the other hand, knock-out of the DNA demethylase TET1 lead to an opposite effect: mESCs in the naive state exhibited a profile of contact lengths resembling that of formative-state cells. Furthermore, the A/B compartment segregation was also altered in the knockouts (Figure 1.5B). Together, these observations strongly suggest that DNA methylation is involved in the genome reorganisation upon exit from naive pluripotency. However, the mechanism of this effect remains unclear.

In addition, mESCs devoid of HP1 $\beta$  and HP1 $\gamma$  displayed an altered contact distribution in 2i/LIF, with an increase in long-range contacts almost matching the formative state levels (Figure 1.5C). The profile also did not change upon exit from naive pluripotency. Unlike DNA methylation, the potential role of HP1 in this process can be imagined more easily, since these proteins have an intrinsic ability to change chromatin architecture (see section 1.4). In addition, as mentioned in section 1.2.1, the DNA methylation and the HP1-H3K9me2/3 networks are intimately linked. Therefore, in this study, the features of HP1 in mESCs and upon exit from naive pluripotency were studied in detail, in an attempt to understand its role in chromatin reorganisation.



**Figure 1.5:** A. Log-log contact length profiles in wild-type, *Dnmt3a/b* double knockout (DKO) and *Tet1* knockout mESCs, in 2i/LIF and 24h after 2i withdrawal. *Dnmt3a/b* DKO cells exhibit a profile that is similar to WT 2i/LIF, which does not change upon differentiation (no arrows). *Tet1* KO cells in 2i/LIF already display the changes observed in the WT only 24h after onset of differentiation (gray arrows). B. Heatmaps of contact enrichment within compartments for the same cell lines and conditions as in A. Red circles highlight conditions where the plot is reminiscent of wild-type 2i/LIF, and blue circles indicate heatmaps similar to wild-type 24h. C. Same as A, but for *Cbx1/3* (HP1 $\beta$  and HP1 $\gamma$ ) DKO. Gray arrow indicates the modest increase in long-range contacts in the knockout, and no change is apparent in mutant cells upon differentiation. Figures modified from ones made by Dr D. Lando.

## 1.4 HP1 and models of heterochromatin organisation

HP1 proteins are the hallmark of heterochromatin in all eukaryotes, associated with compaction and segregation of chromatin and repression of transcription and other processes. They are conserved from yeast to plants and mammals, and many organisms have multiple isoforms of HP1, sometimes performing different functions.

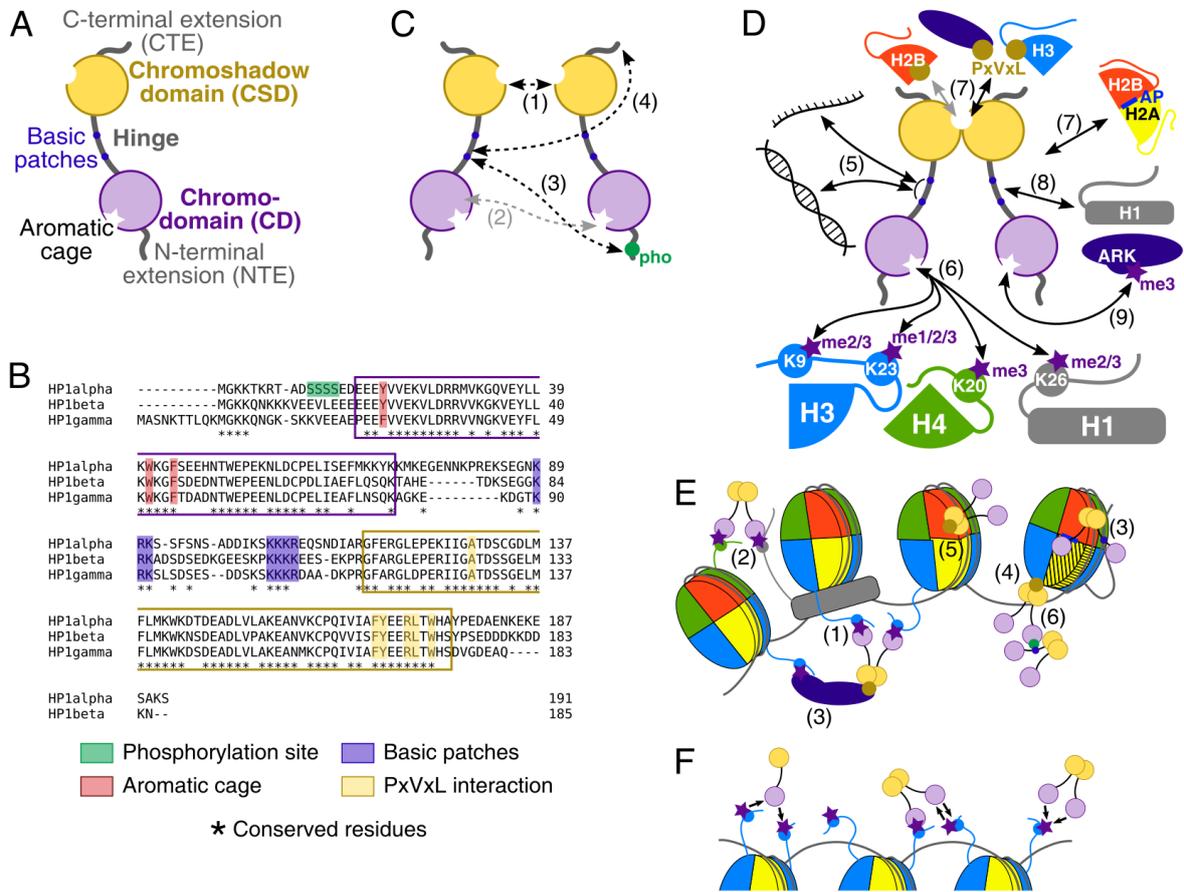
### 1.4.1 Structure and interactions of HP1

#### Structure of mammalian HP1 proteins

All HP1 proteins share the same basic structure with two domains connected by a flexible hinge (Figure 1.6A). In mammals, the 55-residue-long N-terminal chromodomain (CD) interacts with H3K9me2/3 via its aromatic cage. The C-terminal 61-residue-long chromoshadow domain (CSD) homodimerises, creating a landing platform for a variety of proteins. The unstructured hinge region is thought to be involved in low-specificity interactions, especially through its multiple basic patches. Many HP1 proteins also contain disordered regions at either end, termed the N- and C-terminal extensions (NTE and CTE). Overall, HP1 proteins are typically ~180-190 amino acids long and have a mass of ~25 kDa.

Mammals, including mice and humans, have three isoforms of HP1: HP1 $\alpha$  (encoded by the *Cbx5* gene), HP1 $\beta$  (*Cbx1*) and HP1 $\gamma$  (*Cbx3*). The sequence identity between the isoforms is high within the structured domains, while it is significantly lower in the disordered regions (Figure 1.6B).

HP1 proteins are extensively decorated with PTMs, with the diversity and frequency reminiscent of that of the histone proteins (Lomberk et al. 2006; LeRoy et al. 2009). The majority of these PTMs have unknown effects, although a few of them have been shown to significantly alter the activity of HP1 proteins. Among them is, for example, HP1 $\alpha$ -specific serine phosphorylation in the NTE (Figure 1.6B).



**Figure 1.6:** A. Basic structure of HP1 proteins. B. Sequence alignment of the mouse HP1 isoforms. Structured domains are boxed in respective colours, and important residues are highlighted. C. Summary of HP1 self-interactions: dimerisation of the CSD domains (1); putative interaction between ARK loop and CD aromatic cage (2); binding of hinge basic patches to P-NTE (HP1 $\alpha$ ) (3); contacts between basic patches and CTE (HP1 $\alpha$ ) (4). D. Summary of HP1 interactions: interaction between basic patches and DNA/RNA (5); aromatic cage of CD binds H3K9me2/3 and other methylated lysines on histones (6); strong interaction between CSD dimer interface and PxVxL motif-containing proteins, including the  $\alpha$ N helix of histone H3 and  $\alpha$ 1 helix of H2B (7); HP1 interacts with the acidic patch (AP) in the H2A-H2B cleft (8); HP1 hinge binds H1 (9); aromatic cage of CD binds methylated ARK/T/S sequence (10). In C and D, grey arrows indicate interactions not demonstrated for mammalian HP1 proteins. E. In the context of chromatin, HP1 can bridge nucleosomes due to its dimeric nature (1 and 2). Higher valency can be achieved by binding other CD-containing proteins, such as SUV39h1/2, via the CSD (2). Interaction between the hinge and DNA or RNA can also occur, and it might be stabilised by the extended acidic patch of the H2A.Z histone variant (3). HP1 CSD dimer can bind histone H3  $\alpha$ N helix, which might require partial unwrapping of DNA (4), or potentially H2B  $\alpha$ 1 helix, affecting the nucleosome conformation (5). Finally, some conformations of HP1 can be oligomerisation-competent, e.g. through P-NTE-hinge interaction (HP1 $\alpha$ ) when the CD is not engaged (5). F. Avidity of HP1 to H3K9me2/3-modified chromatin is increased by microassociation and -dissociation events (left), which are further enhanced for the HP1 dimer (centre), and higher dimer association probability (right).

## Interactions of HP1

HP1 proteins have a vast interactome, including both proteins and nucleic acids (Figures 1.6C,D; numbers match the headings below). Moreover, it varies between the isoforms, and even for common partners the affinities can be very different. In the cell, these interactions can be context-dependent, modulated either by the PTMs or by other interactions, and the complex effects of this are still poorly understood.

### Self-interactions

**CSD dimerisation (1)** While all HP1 proteins dimerise through the CSD, the affinity of this interaction varies between the HP1 variants. For example, the  $K_d$  has been reported to be <150 nM for mouse HP1 $\beta$  (Brasher et al. 2000), but only ~1  $\mu$ M for human HP1 $\alpha$  (Larson et al. 2017). Furthermore, this interaction might be modulated by the CTE, which for example reduces the dimerisation constant of HP1 $\beta$  to ~500 nM (Mendez et al. 2011).

Moreover, dimerisation creates a surface to which many other proteins bind (see below), and this is thought to promote dimer formation in cells (Mendez et al. 2011). It is thus generally considered that in physiological contexts HP1 exists as a dimer. HP1 proteins can form heterodimers between different isoforms (Ye et al. 1997; Nielsen et al. 2001; Müller-Ott et al. 2014), but it has not been directly demonstrated that this occurs through the CSD (Kumar and Kono 2020).

**CD dimerisation (2)** Other interactions between HP1 proteins are also known. In the fission yeast HP1 homologue Swi6, a sequence in the CD resembling the H3 tail can bind to the aromatic cage of the other CD (Canzio et al. 2013). This can result in either a closed Swi6 dimer conformation, or higher-order oligomerisation. It also diminishes the affinity of Swi6 for H3K9me2/3. This interaction, however, does not appear to be preserved in mammalian HP1 (Brasher et al. 2000; Munari et al. 2012; Munari et al. 2013; Hiragami-Hamada et al. 2016).

**Pho-NTE + hinge (3)** Furthermore, mammalian HP1 $\alpha$  possesses another two types of self-association. Contacts between the CTE and the hinge (4) stabilise an autoinhibited closed state of the dimer. Phosphorylation of HP1 $\alpha$  NTE (pho-NTE), on the other hand, promotes an open state, which is capable of higher-order oligomerisation through interactions between the pho-NTE and the basic patch on the hinge (Larson et al. 2017).

### **DNA and RNA (5)**

HP1 proteins bind both RNA and DNA, predominantly through the basic patches in their hinge regions. Due to the heterogeneous effects of the nucleic acid length and sequence on the interaction, it is hard to pinpoint a single  $K_d$  value for it. For ~100-200 bp DNA fragments, it was suggested to be in the order of 5-20  $\mu$ M (Sugimoto et al. 1996; Zhao et al. 2000; Nishibuchi et al. 2014). The affinity of HP1 $\alpha$  to unmodified nucleosome with linker DNA was measured to be 0.6  $\mu$ M, and was primarily driven by interaction with DNA (Ryan and Tremethick 2018). Interestingly, HP1 $\alpha$  seems to bind DNA about two- to three-fold stronger than either of the two other mammalian isoforms, but the affinity drops upon NTE phosphorylation (Nishibuchi et al. 2014). Moreover, HP1 $\alpha$  appears to interact with RNA even stronger than with DNA (Muchardt et al. 2002).

### **Histone proteins**

**CD + H3K9me2/3, H3K23me1/2/3, H1.4K26me2, H4K20me3 (6)** The CD of HP1 binds methylated lysines on histone tails. The most well-characterised of these interactions is with H3K9me2/3, for which it is known that the methyl groups are engaged in the aromatic cage of HP1 (Nielsen et al. 2002; Jacobs and Khorasanizadeh 2002). Other residues in the CD make contacts with the rest of the H3 tail, providing the specificity towards K9 methylation, and the complex is further stabilised by the HP1 NTE. The variation in the NTE sequence is likely one of the reasons for a differential affinity of HP1 isoforms to H3K9me2/3, with  $K_D$  estimated at 13-35, 8, 2-6 and 7-8  $\mu$ M for HP1 $\alpha$ , pho-HP1 $\alpha$ , HP1 $\beta$  and HP1 $\gamma$ , respectively (Nielsen et al. 2002; Fischle

et al. 2005; Hiragami-Hamada et al. 2011). Importantly, the phosphorylation of H3 serine 10 (H3S10pho) disturbs the CD-H3K9me<sub>2/3</sub> interaction, and is responsible for HP1 eviction from chromatin during mitosis (Fischle et al. 2005).

The HP1 CD also binds H1.4K26me<sub>2/3</sub> and H3K23me<sub>2/3</sub> with affinities similar or slightly lower than for H3K9me<sub>2/3</sub> (Daujat et al. 2005; Ruan et al. 2012; Liu et al. 2010; Shanle et al. 2017). The mechanism of binding to H1.4K26me<sub>2/3</sub> is similar to H3K9me<sub>2/3</sub> (Ruan et al. 2012), and it is likely to also be so for H3K23me<sub>2/3</sub> (Pokorná et al. 2021). The interaction with H1.4K26me<sub>2</sub> is important for HP1 localisation to chromatin in cells (Daujat et al. 2005). On the other hand, while H3K23me<sub>1</sub>, whose affinity for HP1 CD is weaker, co-localises with HP1 $\beta$  in NIH 3T3 cells, H3K23me<sub>2</sub> does not (Liu et al. 2010). In addition, HP1 $\beta$  specifically was recently found to bind H4K20me<sub>3</sub>, although the affinity was weaker than for H3K9me<sub>3</sub> (Bosch-Presegué et al. 2017).

**CSD + H3  $\alpha$ N helix and H2B  $\alpha$ 1 helix (7)** The  $\alpha$ N helix of the globular domain of histone H3 contains a motif similar to the consensus PxVxL sequence (see below), and binds the HP1 CSD dimer interface (Lavigne et al. 2009; Richart et al. 2012). The strength of this interaction varies between 4 and 60  $\mu$ M, and is higher for full-length HP1 than for the CSD alone (Richart et al. 2012; Liu et al. 2017). The H3  $\alpha$ N helix is positioned at the DNA entry/exit site on the nucleosome, suggesting that DNA might need to partially unwrap to allow binding of HP1 (Liu et al. 2017).

Furthermore, yeast HP1 homologue Swi6 potentially binds to the PxVxL-like motif on  $\alpha$ 1 helix of histone H2B, stabilising an alternative conformation of the nucleosome core (Sanulli et al. 2019).

**H2A-H2B acidic patch (8)** HP1 proteins might interact with the acidic cleft on the H2A-H2B interface, since the extended acidic patch of the histone variant H2A.Z strengthens the binding of HP1 to nucleosomes (Fan et al. 2004). The affinity is enhanced to such an extent that it can functionally substitute for H3K9me<sub>2/3</sub> modification, potentially making this an important targeting mechanism for HP1 $\alpha$  (Ryan

and Tremethick 2018).

**Hinge + H1 (9)** In addition to the CD-H1.4K26me<sub>2/3</sub> interaction, the HP1 $\alpha$  hinge region was also found to bind a mixture of H1 variants (Nielsen et al. 2001).

### **Other proteins**

**PxVxL + CSD dimer interface (7)** A variety of proteins bind to the HP1 CSD dimer interface via a peptide with a consensus sequence of PxVxL (or, more accurately,  $\Phi X(V/P)X\Phi$  (Liu et al. 2017)). The strength of the interaction is determined by the residues flanking this sequence, which contact the  $\beta$ -sheets of both HP1 monomers (Thiru et al. 2004; Mendez et al. 2013). This leads to a large variation in the  $K_D$  values, which can be as low as 60 nM (Lechner et al. 2000). Among the interactors of the CSD is KRAB-associated protein 1 (KAP1), which connects HP1 to the H3K9 methyltransferase SETDB1 (Schultz et al. 2002; Lechner et al. 2000). Another H3K9 methyltransferase, SUV39h1/2, binds HP1 CSD dimer interface directly through its N-terminal unstructured region (Schotta 2002; Yamamoto and Sonoda 2003), although the precise motif has not been identified. Similarly, both DNMT1 and DNMT3a/b also interact with CSD, although they do not possess the canonical PxVxL motifs, and the mode of binding is yet to be determined (Fuks 2003).

**Methyllysine + CD interactions (10)** The CD of HP1 can recognise other methylated ARKS/T sequences, such as ARKme<sub>3</sub>T on G9a, the euchromatic histone H3K9 methyltransferase, with affinity similar to H3K9me<sub>3</sub> (Chin et al. 2007). Notably, G9a can methylate itself at this site, and the levels of methylated G9a in cells were found to be significant (Chin et al. 2007). A number of other methylated proteins bind HP1 in a similar way, among them the kinase DNA-PK, involved in DNA repair, and the spliceosome component eIF4A3 (Liu et al. 2013).

## Nucleosomes

From the self-interactions, interactions with histone proteins, DNA and RNA, a complex picture of potential modes of binding of HP1 to nucleosomes emerges (Figure 1.6E). Firstly, the HP1 dimer can bridge nucleosomes in the “classical” manner, by binding to two H3 tails modified at K9, or to other histone tail PTMs. Interestingly, while human HP1 $\beta$  has a high specificity towards H3K9me2/3-containing chromatin, human HP1 $\alpha$  is more promiscuous (Hiragami-Hamada et al. 2016).

In addition, HP1 – especially HP1 $\alpha$  – might associate with chromatin via interactions with DNA. Indeed, HP1 $\alpha$  has a substantial affinity for unmodified nucleosomes, dependent on the linker DNA, which was enhanced by incorporation of H2A.Z and inhibited by H1 (Ryan and Tremethick 2018). Alternatively, HP1 might bind to the octamer core via the histone fold of H3 (or potentially H2B), altering the nucleosome conformation and affecting properties of the chromatin fiber (Richart et al. 2012; Liu et al. 2017; Sanulli et al. 2019). Another way for HP1 to be tethered to the nucleosome could be via the hinge-H1 interaction (Nielsen et al. 2001).

When the HP1 CD is not engaged, HP1 has limited capacity to cross-link chromatin, but more of its interaction sites might be available to form higher-order oligomers (Canzio et al. 2014; Kumar and Kono 2020).

### 1.4.2 Activity and dynamics of HP1 on chromatin

#### HP1 interaction with chromatin is multivalent and dynamic

Although the interaction between a single CD domain and H3K9me2/3 mark is not strong, HP1 is retained on H3K9me2/3-modified nucleosome arrays due to multiple factors (Hiragami-Hamada et al. 2016; Kilic et al. 2015) (Figure 1.6F). Firstly, the avidity of HP1 interaction with chromatin containing many binding sites in close spatial proximity is increased via rapid reassociation events (Kilic et al. 2015). Interestingly, dwell time decreases at higher HP1 concentrations, suggesting that competition between HP1 molecules can prevent local reassociation (Kilic et al. 2015). Furthermore, divalency of the HP1 dimer also contributes to the slowed apparent off-rate, since reassociation of

one chromodomain is facilitated by tethering of the dimer via the other one (Hiragami-Hamada et al. 2016; Kilic et al. 2015). In agreement with both ideas, HP1 residence time is prolonged significantly by an increased density of H3K9me3 tails (Hiragami-Hamada et al. 2016; Kilic et al. 2015; Bryan et al. 2017). The interactions between HP1 and DNA, although weak, also contribute to increased avidity (Bryan et al. 2017). In addition to affecting dissociation kinetics, the dimeric nature of HP1 also increases the on-rate (Kilic et al. 2015). Overall, the apparent  $K_D$  for the interaction between the constitutively dimeric HP1 $\alpha$  and chromatin is at least 60-fold higher than that for HP1 $\alpha$  and H3K9me3 peptide (Kilic et al. 2015).

Importantly, as a result, chromatin conformation might enhance or reduce HP1 binding. For example, compacting nucleosome arrays using high  $Mg^{2+}$  concentration increases HP1 $\alpha$  affinity for chromatin (Fan et al. 2004).

### **HP1 compacts DNA and nucleosome arrays**

As expected from their nucleosome bridging activity, HP1 proteins compact nucleosome arrays by cross-linking them, and at high concentrations can even bridge nucleosomes from separate arrays (Fan et al. 2004; Hiragami-Hamada et al. 2016). On top of bridging, other mechanisms also contribute to efficient chromatin compaction by HP1, such as the contacts between the H4 tail and the H2A-H2B acidic patch (Fan et al. 2004). The combination of HP1's ability to compact chromatin and its higher affinity for tightly packed nucleosomes suggests a mechanism for cooperativity, observed in multiple studies (Fan et al. 2004; Larson et al. 2017).

### **HP1 promotes liquid-liquid phase separation**

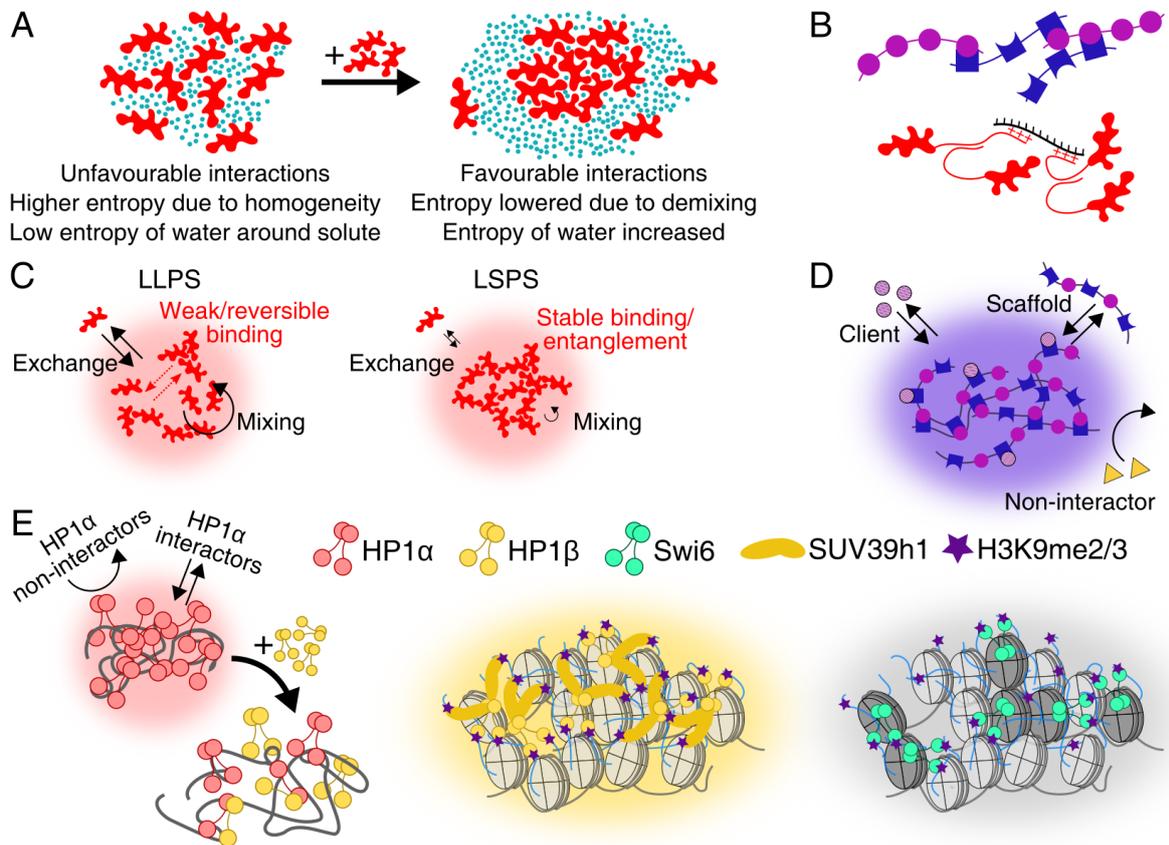
Phase separation is a widespread physical phenomenon. It occurs under poor solvent conditions, when the weak non-specific interactions between the solute molecules (proteins and nucleic acids in biology) are stronger than those between the solute and the solvent (water) (Banani et al. 2017). In this case, the free energy of the solution has two (or more) minima at different solute concentrations, and upon reaching a certain

total concentration threshold, the solution spontaneously demixes into a dilute and a dense phase (Figure 1.7A). Within each phase, the favourable solvent-solvent and solute-solute interactions prevail, and the entropic cost of demixing is balanced by releasing the ordered water molecules that would otherwise surround the solute particles. Equal numbers of molecules transition into and out of the dense phase, and the system is thermodynamically stable. Within cells, this mechanism is proposed to allow the formation of membraneless organelles with increased concentrations of certain components without the need for a constant input of energy (Banani et al. 2017; Alberti 2017).

Phase separation is promoted by additional specific multivalent homo- or heterotypic interactions due to a lower entropic cost of confining a complex of tethered solute molecules into the condensed phase (Banani et al. 2017) (Figure 1.7B). Thus, proteins with modular domains that can form specific multivalent interactions tend to phase separate more easily. Furthermore, many types of forces – charged, polar, cation- $\pi$  and  $\pi$ -stacking, or hydrophobic – are responsible for both specific and non-specific interactions that contribute to phase separation of internally-disordered proteins.

The condensed phase might display different material properties, defined by the dynamics of its components (Figure 1.7C). Gels and precipitates are examples of solid-like dense phases, formed by liquid-solid phase separation (LSPS). In contrast, drops of oil in water possess liquid-like characteristics: the oil molecules can rearrange internally and exchange with the environment. The latter behaviour is termed liquid-liquid phase separation (LLPS), and it has received much attention recently due to its unique potential to form stable, yet dynamic, biomolecular condensates. Whether the dense phase displays solid-like or liquid-like behaviour depends on the strength and abundance of strong specific interactions and the entanglement of long polymers (Banani et al. 2017).

Phase-separated domains can include and exclude various molecules, although the exact rules for this selectivity are unclear (Alberti 2017). The components partitioning into the dense phase might be either multivalent “scaffolds”, driving the LLPS, or monovalent or low-valency “clients”, which are admitted into the droplets through



**Figure 1.7:** A. Upon increasing solute (red shapes) concentration, phase separation occurs when the solute-solute interactions are stronger than the solute-solvent (blue circles) interactions. B. Phase separation is promoted by multivalency. In the context of proteins, common themes are specific interactions between modular domains (top) or less specific interactions between intrinsically disordered regions, sometimes also involving DNA or RNA (bottom). C. Liquid-liquid phase separation (LLPS) occurs when molecules within the dense phase can mix and change their interaction partners (left). Liquid-solid phase separation (LSPS), whereby molecules are crosslinked and cannot diffuse, is a consequence of vitrification or entanglement (right). D. A condensed phase formed by scaffold proteins can incorporate the proteins that interact with it and exclude the ones that do not. E. *In vitro* LLPS, observed for HP1 proteins. HP1 $\alpha$  forms condensates with DNA, while HP1 $\beta$  is able to dissolve them (left). All HP1 isoforms induce phase separation when mixed with H3K9me3-modified chromatin in the presence of an interactor which also contains an H3K9me3-interacting chromodomain, e.g. SUV39h1 (middle). Chromatin has an inherent capability to phase separate, which might be enhanced by HP1 binding, e.g. through altering nucleosome conformation (right).

binding to the scaffold. The latter can stabilise or disrupt LLPS, but are themselves neither necessary nor sufficient for condensation (Banani et al. 2017) (Figure 1.7D). Thus, in cells, LLPS might create selectively permeable boundaries, admitting molecules into the condensate only if they bind one of its components.

HP1 has been found to phase separate in different contexts (Figure 1.7E). Firstly, HP1 $\alpha$  displays LLPS at moderate concentrations when mixed with DNA, RNA or on its own when phosphorylated in the NTE (Larson et al. 2017; Strom et al. 2017; Novo et al. 2020). This is facilitated by the electrostatic interactions of its hinge region either with negatively charged nucleic acids or phosphate groups. Interestingly, neither HP1 $\beta$  nor HP1 $\gamma$  possess this property due to sequence divergence of their hinge domains, but they can localise to and destabilise HP1 $\alpha$ -DNA droplets (Larson et al. 2017; Keenen et al. 2021) (Figure 1.6B). The pho-HP1 $\alpha$  condensates incorporate H3K9me3-labelled nucleosomal arrays and HP1 interactors like Aurora B, but exclude other molecules such as bacterial Hsp90 (Larson et al. 2017).

Secondly, all HP1 isoforms were demonstrated to promote LLPS in a more complex system, containing H3K9me2/3-modified chromatin (Wang et al. 2019). The mechanism behind this phenomenon involves multivalent CD-H3K9me2/3 interactions: while HP1 on its own is unable to form chromatin condensates, droplets appear upon addition of SUV39h1 or KAP1, which increases the number of linked chromodomains from 2 to 3-6. The SUV39h1-HP1 $\beta$  droplets exclude the general transcription factor TFIIB and are resistant to DNase I treatment.

Finally, chromatin is able to phase separate on its own under physiological salt conditions (Gibson et al. 2019). Since HP1 can change the properties of the chromatin fiber, it may promote or inhibit this process. Indeed, yeast homologue of HP1 Swi6 was found to promote chromatin LLPS, potentially by altering the conformation of nucleosomes (Sanulli et al. 2019). Importantly, histone H1 also promotes chromatin phase separation, increases chromatin density within droplets, and makes them more solid-like (Gibson et al. 2019). This could occur through cross-linking of DNA (Turner et al. 2018), alteration of nucleosome conformation (Bednar et al. 2017), and/or relieving electrostatic repulsion by shielding of the negative charge on the linker DNA

(Hansen et al. 2021). Thus, HP1 and H1 might be expected to have opposing or synergistic effects on chromatin LLPS, depending on how they interact with chromatin and with each other.

Overall, the above studies show that HP1 is able to promote LLPS through multiple mechanisms in reconstituted *in vitro* systems. Their cumulative action might result in different biophysical characteristics of chromatin condensates, such as their material properties and stability. The effect will likely depend on the relative amounts of the HP1 isoforms, their PTMs, the features of the chromatin fiber and its PTMs, and the other molecules present in the system. However, how exactly this manifests in cells and whether HP1 and/or chromatin LLPS is relevant in physiological cellular environment is currently unclear.

### **HP1 proteins in cells are dynamic**

Since heterochromatic foci are rather static and immobile, HP1 was initially thought to form stable cross-linked macromolecular complexes with nucleosomes. This model was challenged by measurements of HP1 mobility in multiple cell types, which indicated that all mammalian HP1 isoforms are very dynamic, exchanging on much shorter timescales ( $t_{1/2}$  of seconds to tens of seconds) compared to chromatin movement (Cheutin et al. 2003; Festenstein et al. 2003; Schmiedeberg et al. 2004; Müller et al. 2009). HP1 in euchromatin was found to be more mobile than in heterochromatin (Cheutin et al. 2003; Festenstein et al. 2003). Fast and slow dynamic populations of HP1 were observed in both eu- and heterochromatin, while an additional immobile, stably-bound fraction was present predominantly in heterochromatin (Festenstein et al. 2003; Schmiedeberg et al. 2004; Müller et al. 2009).

Measurements in mESCs revealed that mobility of heterochromatic HP1 correlates with pluripotency. Fluorescence recovery after photobleaching (FRAP) recovery times were significantly shorter in mESCs than in neural progenitor cells or even in ESCs 24h after LIF withdrawal (Meshorer et al. 2006). Notably, this feature extended to other architectural proteins, such as histones.

Most of the chromatin-HP1 interactions described above affect HP1 dynamics in

live cells: an increase in mobility is observed upon deletion of either the CD or the CSD, as well as disruption of H3K9me2/3 binding, dimerisation, PxVxL protein binding or DNA binding with point mutations (Cheutin et al. 2003; Kilic et al. 2015; Bryan et al. 2017).

### **1.4.3 Models of heterochromatin structure**

Heterochromatin differs from euchromatin in multiple ways, but the physical and chemical basis for this is still not understood. In this section, the distinctive features of heterochromatin and the potential (non-exclusive) mechanisms behind them are reviewed.

#### **High density and slow mixing of chromatin**

Heterochromatin obviously contains more DNA and histones than euchromatin, and heterochromatic domains could represent a different phase state of chromatin. Purified chromatin falls out of solution in the presence of divalent salts at nearly physiological concentrations (Hansen et al. 2021). This LSPS process occurs due to self-interaction of chromatin fibres – resulting especially from interactions between the DNA and the histone tails – upon shielding of the charge repulsion between the DNA phosphate backbones. The packing of nucleosomes in the precipitate is very similar to cellular heterochromatin (Hansen et al. 2021). In addition, heterochromatic domains behave as gel-like solids in at least some cell types, since the mixing of chromatin, as measured by FRAP, is extremely slow (recovery half-time >several hours) (Lever et al. 2000; Phair and Misteli 2000; Kimura and Cook 2001; Strickfaden et al. 2020).

On the other hand, under some conditions, purified chromatin can form liquid droplets (Gibson et al. 2019). This LLPS process might be more relevant in some cell types, e.g. mESCs, where chromatin mixing has been shown to be much faster than in differentiated cell types (recovery of 20-40% in 100 s) (Meshorer et al. 2006; Novo et al. 2020).

Chromatin condensation could also be induced by exogenous factors, such as HP1, which add internucleosomal links, or H1, that shield the negative charge of the linker

DNA (Hansen et al. 2021). Crosslinking by HP1, for example, has been demonstrated to increase the stiffness of chromatin (Strom et al. 2021). However, the HP1-H3K9me3 interaction was shown to often be dispensable for chromocentre formation (Peters et al. 2001; Maison et al. 2002; Mateos-Langerak et al. 2007; Erdel et al. 2020; Strom et al. 2021), although in some cases it is important for proper compaction (Bosch-Presegué et al. 2017; Wang et al. 2019). Thus, it appears that chromatin LSPS in physiological conditions could either be an inherent feature of (appropriately modified) chromatin, and/or be caused by different chromatin cross-linkers.

A conceptually different mechanism could be preferential inclusion of (appropriately modified) chromatin into LLPS foci formed by other factors, such as HP1. Indeed, phospho-HP1 $\alpha$  condensates incorporate nucleosomal arrays at high density (Larson et al. 2017). Similarly to the cellular heterochromatic domains, HP1 $\alpha$ -DNA foci display little mixing of DNA, and are able to resist forces (Keenen et al. 2021).

Overall, four mechanisms might play a role in heterochromatin compaction and solid-like behaviour: chromatin LSPS, potentially enhanced or caused by cross-linking of chromatin by exogenous factors, chromatin LLPS, or recruitment of (appropriately modified) chromatin into liquid droplets of other proteins.

### **Enrichment and high mobility of heterochromatic factors**

Certain factors – HP1, SUV39h1/2, and others – are present in heterochromatic foci at increased concentrations, which can be explained by two models: enrichment of binding sites or LLPS. Chromatin condensed into a globule should contain a higher density of binding sites for proteins, and by extension higher concentrations of these proteins, which might be further enhanced by specific histone modifications. Alternatively, or in addition to simple binding, a condensate of heterochromatic proteins might form through LLPS. Several scenarios are then possible: chromatin condenses first, and then nucleates the formation of a protein droplet; heterochromatic factors form a droplet, which then incorporates and condenses chromatin; or a combination of the two.

Although LLPS is an attractive model, it has been argued that the protein:chromatin ratios used in many *in vitro* experiments are significantly higher than

physiological (Hansen et al. 2021). For HP1 $\alpha$ , 50  $\mu$ M protein is required to phase separate with DNA, while concentrations in cells are typically only a few  $\mu$ M (Larson et al. 2017; Erdel et al. 2020). However, other experiments shows that HP1 can condense long fragments of DNA at much lower concentrations (Keenen et al. 2021). Furthermore, in cells, increased crowding, the presence of higher-valency HP1 complexes with other proteins (Wang et al. 2019), and potentially high local concentrations may promote LLPS (Larson and Narlikar 2018). Finally, evidence for HP1a LLPS in the cells of live *Drosophila* embryos, such as special behaviour of molecules at the boundary of heterochromatic domains (see chapter 3), has been presented (Strom et al. 2017). Thus, whether LLPS of heterochromatic proteins is plausible in cells remains an open question.

As discussed above, many proteins within heterochromatin move much faster than chromatin itself. This rapid exchange would be expected for LLPS, and has been observed *in vitro* (Larson et al. 2017; Keenen et al. 2021). However, it could also be compatible with the model of protein binding to a condensed chromatin globule, if the interaction between proteins and chromatin is not very stable, and if the binding sites on chromatin are accessible (Heltberg et al. 2021). Indeed, HP1 binding to chromatin is dynamic in nature (Kilic et al. 2015). In addition, condensed chromatin *in vitro* is accessible to some extent to MNase, restriction enzymes and transcriptional machinery, arguing that a certain level of dynamics could still be possible even within solid-like heterochromatin (Maeshima et al. 2016; Mishra and Hayes 2018; Zhou et al. 2007; Tse et al. 1998).

### **Exclusion of non-heterochromatic proteins, repressive state and crowding**

It is generally believed that particular factors – e.g. transcriptional machinery – are excluded from heterochromatin (Elgin and Reuter 2013; Larson and Narlikar 2018). This claim is only partially true: e.g. in some cell types, heterochromatin is transcribed (Novo et al. 2020). Moreover, in human fibroblasts and CHO cells, RNA polymerase II (RNAPII) is absent from some, but not all, heterochromatic foci (Verschure et al. 2003). However, transcription does preferentially occur outside condensed regions

of chromatin, on the border with the interchromatin space (Fakan and Driel 2007). Furthermore, ChIP-seq experiments show a clear partitioning between the heterochromatic factors and RNAPII (Elgin and Reuter 2013). How is this achieved?

Perhaps the most long-standing theory explains this by crowding caused by condensed chromatin. The effects of crowding include size-dependent volume exclusion and slower diffusion, which could strongly affect the bulky transcriptional machinery (Bancaud et al. 2009). Indeed, inert probes are partially excluded from heterochromatin in a size-dependent manner (Verschure et al. 2003; Bancaud et al. 2009), and their diffusion is slower in DNA-dense regions (Bancaud et al. 2009; Xiang et al. 2020). However, even 90-nm-diameter particles are still able to access the foci to some extent. Moreover, the decrease in the diffusion rate is mild, and the effect is not size-dependent, suggesting that simple crowding is not sufficient to explain the repressive environment (Bancaud et al. 2009). Intriguingly, however, there is an inverse correlation between the size of the enzymes associated with various classes of chromatin domains and the density of these domains (Miron et al. 2020).

An alternative model for protein exclusion could be the insulation of heterochromatin by a phase boundary. It is plausible that such a boundary could be largely impermeable to proteins unable to interact with HP1, other heterochromatic proteins or the chromatin fiber, although the exact molecular grammar of this is not known. Consistently, exclusion of factors from HP1-induced droplets has been observed *in vitro* (Larson et al. 2017; Wang et al. 2019). In addition, diffusion of an inert YFP probe, which is depleted from heterochromatin in *Drosophila* S2 cells, is altered at the boundary of HP1a domains, arguing for an LLPS-mediated mechanism for exclusion (Strom et al. 2017). In contrast, in mouse cells, GFP depletion from chromocentres was found to be independent from HP1 (Erdel et al. 2020).

Another possibility is that there is no specific exclusion. If DNA in heterochromatin is somehow locally shielded from RNAPII binding, the lower concentration of RNAPII and/or transcriptional silencing might simply result from the absence of bound fraction, as has been recently suggested for the inactive X chromosome territory (Collombet et al. 2021). In other words, freely diffusing RNAPII could be perfectly able to access

heterochromatin, but unable to bind and therefore stay there for long. Such local repression could also be mediated by HP1, although the mechanism for this is unclear (Erdel et al. 2020).

In summary, in some cases inert probes and cellular proteins are excluded from heterochromatin. This partitioning could be caused by steric effects – indeed, simulations suggest that crowding is sufficient to explain the dynamic distribution of free GFP (Bancaud et al. 2009). However, crowding does not appear to be able to explain transcriptional repression in heterochromatin. On the other hand, LLPS (of HP1 and/or other factors) is an elegant model that could provide a mechanism for the selective enrichment or depletion of different proteins in heterochromatin, but the evidence for it remains controversial.

Overall, chromatin in heterochromatic domains generally possesses characteristics of a solid, but might exist as a liquid in certain cell types (Hansen et al. 2021; Novo et al. 2020). These states might be intrinsic to the chromatin fiber with particular features, such as histone modifications, in a particular environment and at a certain concentration. Architectural proteins, such as HP1, could cause or facilitate these phase transitions, and/or change the material properties of heterochromatin. Alternatively, heterochromatic proteins might condense chromatin through their own ability to phase separate (Larson and Narlikar 2018). Heterochromatin is clearly a crowded environment, which impedes, but does not prevent, diffusion of proteins within it (Bancaud et al. 2009). Whether any specific exclusion or inclusion of factors takes place in heterochromatin, and how transcriptional repression is achieved, remains an open question.

#### **1.4.4 Functions of HP1 proteins**

Although HP1 proteins are generally associated with heterochromatin condensation and repression, their cellular functions are much more varied and not well understood. Different HP1 isoforms perform overlapping, but distinct tasks, targeted and modulated by their vast pool of interactors.

## Localisation of HP1 isoforms

HP1 isoforms are distributed differently within the nucleus. The earliest immunofluorescence experiments, performed on immortalised mouse 3T3 and P19 cells, indicated that HP1 $\alpha$  displays a clustered pattern overlapping with chromocentres, HP1 $\gamma$  exhibits a diffuse pan-nuclear staining, while HP1 $\beta$  is enriched in chromocentres, but is also present elsewhere (Minc et al. 1999; Nielsen et al. 2001). However, HP1 localisation was found to vary drastically between cell types.

Even mESCs display different HP1 patterns, depending on the precise culture conditions and genetic background. In mESCs grown in serum/LIF, HP1 $\beta$  was found to form foci in nearly all cells, while HP1 $\alpha$  and HP1 $\gamma$  did so only in a fraction of them. Upon removal of LIF, both HP1 $\alpha$  and HP1 $\beta$  foci became smaller and more numerous (Dialynas et al. 2007). In contrast, another publication suggested that HP1 $\beta$  and HP1 $\gamma$  in mESCs are not enriched at pericentromeric heterochromatin, and are predominantly soluble rather than associated with chromatin (Mattout et al. 2015). Yet another study found HP1 $\beta$  enriched in chromocentres, but also displaying a significant diffuse staining (Hiragami-Hamada et al. 2016). In human ESCs, all HP1 proteins had a diffuse localisation, and upon induction of differentiation using retinoic acid HP1 $\alpha$  concentrated in chromocentres, while HP1 $\beta$  formed foci distinct from chromocentres (Bártová et al. 2008). Other studies also showed that large and diffuse HP1 $\alpha$ -marked chromocentres are a hallmark of pluripotent mESCs, while upon differentiation they become more condensed and numerous (Meshorer et al. 2006; Novo et al. 2016). Therefore, overall it is not well understood what happens to HP1 isoforms, especially HP1 $\beta$  and HP1 $\gamma$ , upon differentiation, particularly in the very first steps of exit from naive pluripotency.

HP1 distribution, determined either by ChIP-seq or by co-staining, generally overlaps with H3K9me2/3, and HP1 is displaced from chromocentres in cells lacking this modification (Lachner et al. 2001; Maison et al. 2002; Hiragami-Hamada et al. 2016). The CSD domain of HP1 is also important for its targeting, likely due to its function in dimer formation (Cheutin et al. 2003; Hiragami-Hamada et al. 2016). Furthermore,

HP1 localisation is defined by additional factors, including RNA (Muchardt et al. 2002; Maison et al. 2002; Quinodoz et al. 2021) and histone H1 (Daujat et al. 2005), while interaction with at least some of the PxVxL motif proteins appears to stabilise chromatin binding (Thiru et al. 2004).

### **Compaction, segregation and silencing**

As discussed above, although the precise role of HP1 in heterochromatin formation is not understood, it is undeniably important for its function. HP1 proteins were initially identified in *Drosophila* as mediators of position-effect variegation, i.e. transcriptional silencing of heterochromatic regions (James and Elgin 1986; Eissenberg et al. 1990). At least in some cell types, knockout of HP1 affects chromocentre compaction (Bosch-Presegué et al. 2017; Wang et al. 2019). Formation of the constitutive heterochromatin environment at centromeres by HP1 appears to be important for genome integrity, with HP1 or SUV39H knockouts resulting in chromosome segregation defects during cell division and genome instability (Ekwall et al. 1995; Peters et al. 2001). HP1 is also involved in telomere capping (Fanti et al. 1998).

HP1 also can, and does, silence genes in euchromatin. Artificial recruitment of HP1 to euchromatic loci results in gene repression, although it may sometimes be dependent on their chromatin environment (Ayyanathan et al. 2003; Hiragami and Festenstein 2005). HP1 was also shown to silence particular euchromatic genes, such as cyclin E, in the endogenous context. It is usually recruited to euchromatic targets by adapter proteins, such as Rb or KAP1 (Hediger and Gasser 2006). Interestingly, KAP1 is involved in promoter-proximal pausing of RNAPII (Bunch et al. 2014), while HP1a accumulates at transcription start sites with stalled RNAPII in *Drosophila* (Yin et al. 2011). Furthermore, HP1 forms a complex with ADNP and CHD4, which has been suggested to inhibit CTCF binding and repress transcription at SINE B2 elements genome-wide (Ostapcuk et al. 2018; Kaaij et al. 2019). The mechanism by which HP1 represses euchromatic genes – transcriptional arrest, local effect on epigenetic marks, formation of a local heterochromatin-like environment, or physical relocation of genes into heterochromatin and/or nuclear periphery – and how it varies depending

on context is poorly understood (Hiragami and Festenstein 2005).

### **Transcription activation**

HP1 has also been suggested to participate in active transcription. For example, HP1 is recruited to heat shock-induced genes in *Drosophila*, where it has been proposed to interact with the nascent transcripts (Piacentini et al. 2003). Furthermore, HP1 $\gamma$  associates with elongating RNAPII (Vakoc et al. 2005). An HP1 $\beta$ -HP1 $\gamma$  switch has been identified upon release of paused RNAPII into processive elongation of the HIV long terminal repeat, underscoring the potential interplay between HP1 isoforms in transcription regulation (Mateescu et al. 2008). Finally, HP1 $\beta$  and DNA methylation were found to globally regulate alternative splicing (Yearim et al. 2015). It is likely that the different functions of HP1 proteins depend on their sequence variation, PTMs and binding partners.

### **Differentiation and development**

Since HP1 proteins have roles in genome architecture, genome stability and gene expression, it is not surprising that they are crucial for cell differentiation and development. HP1 $\beta$  (but not HP1 $\alpha$ ) knockout mESCs, cultured in serum/LIF, differentiate spontaneously and exhibit severely reduced rates of self-renewal (Mattout et al. 2015). They are also biased towards the neuroectodermal fate in differentiation assays. This is thought to be the result of misregulation of a large number of genes upon HP1 $\beta$  knock-out, suggesting that it plays an important regulatory role in mESCs. In addition, the interaction of KAP1 with HP1 $\beta$  and HP1 $\gamma$  is required for differentiation of embryonic carcinoma F9 into parietal endoderm-like cells (Cammass et al. 2004; Cammass et al. 2007).

HP1 $\beta$ , but not HP1 $\alpha$ , has also been demonstrated to be crucial for mammalian development. Mice homozygous for the HP1 $\beta$  knockout die of respiratory failure at birth, and the embryos display defective neuromuscular and cerebral cortex development (Aucott et al. 2008). Interestingly, severe reduction of binding of all HP1 isoforms

to chromocentres in SUV39H1/2 knockout mice is not embryonic-lethal, even though the mice do have severely impaired viability and an increased tumour risk due to chromosome instabilities (Peters et al. 2001). Thus, the essential role of HP1 $\beta$  in mouse development may be related not only to pericentromeric regions.

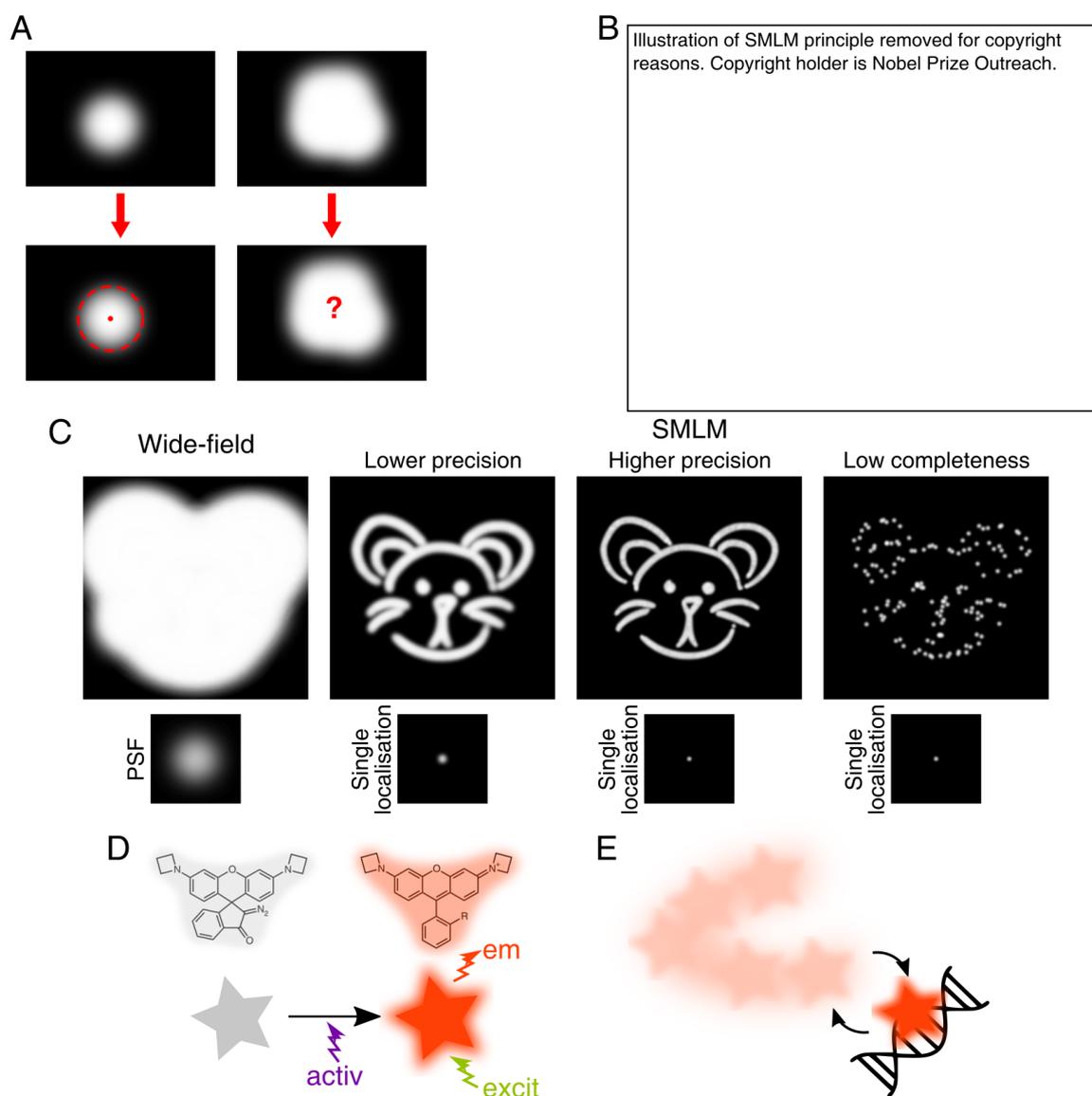
## **1.5 Single-molecule localisation microscopy (SMLM)**

Light microscopy is the most non-invasive type of experiment that is available for understanding the inner workings of cells. Fluorescence microscopy, especially in combination with genetic engineering, has allowed scientists to label and observe specific structures and molecules of interest. However, the resolution of light microscopy is inherently limited to a few hundred nanometers. In the last 20 years, several ingenious solutions were proposed to this seemingly unbreakable barrier, such that nowadays under certain circumstances a resolution of 1-2 nm can be achieved (Balzarotti et al. 2016). This has allowed the observation of processes at the molecular level in intact fixed and live cells.

### **1.5.1 Super-resolution imaging with SMLM**

The umbrella term single-molecule localisation microscopy (SMLM) describes a number of super-resolution imaging techniques with a similar principle. Although the point-spread function (PSF) of a point source has a diameter of a few hundred nanometers, knowing the properties of light and having enough photons allows one to estimate its position rather accurately. The problem occurs when multiple sources are situated close to each other, so that their PSFs overlap, making it impossible to distinguish them (Figure 1.8A). Instead of separating individual molecules in space, SMLM solves this problem by detecting them at different times.

To produce a super-resolved reconstruction, a series of images is taken, each containing only a few PSFs from blinking fluorophores (Figure 1.8B). In each frame, these



**Figure 1.8:** A. Left, a molecule's position can be determined with high precision from its point-spread function (PSF). Right, overlapping PSFs make localisation of individual molecules impossible. B. Principle of SMLM. Adapted from the Nobel Prize website (<https://www.nobelprize.org/prizes/chemistry/2014/press-release/>). C. Precision and completeness determine the quality of an SMLM image. D. Photoactivatable dyes switch from a non-fluorescent to a fluorescent state upon activation with a photon of UV light. Photoactivatable Janelia Fluor 549 dye is shown as an example. E. PAINT approach relies on motion blurring to render freely diffusing molecules invisible. Furthermore, fluorogenicity – the increase in fluorescence upon binding to the specific target – is often utilised to reach better signal-to-noise ratio.

fluorophores are then localised with a high precision, typically by fitting the photon counts to a 2D Gaussian distribution. The resulting data is a list of “localisations”, which can then be rendered into an image.

Two parameters are of key importance to reach high image quality: localisation precision and completeness (Figure 1.8C). Localisation precision refers to the error associated with fitting the Gaussian distribution to the PSF. The higher the photon count, the smoother the experimental distribution is, and the more confidence one can have in the fitting. Thus, bright fluorophores and high laser intensities are normally used for SMLM. In addition, localisation precision is affected by the density of PSFs in each frame, with overlaps leading to ambiguity and reduction in quality.

Even with very high localisation precision, the reconstruction may still not represent the reality faithfully if the image does not have a high sampling rate, or completeness (Figure 1.8C). Thus, assessing the resolution of the final image is not a simple matter.

The different types of SMLM are mainly distinguished by the strategy that is employed to make fluorophores blink. Below, the two approaches used in this work for protein and DNA imaging are described.

### **PALM with photoactivatable dyes**

Photoactivated localisation microscopy (PALM) is based on the phenomenon of photoactivation, whereby a fluorophore transitions from a non-fluorescent into a fluorescent form upon encountering a photon of light at a particular frequency. Although PALM was first developed, using fluorescent proteins (such as photoactivatable GFP), nowadays a wide variety of probes exists. In this study, organic photoactivatable fluorophores were used, which undergo a structural change when “activated” by UV light (Figure 1.8D).

To achieve sparsity, an appropriate activation laser intensity is used. Once the molecules are imaged, they are permanently bleached by the high-intensity excitation laser. Therefore, accurate counting can be achieved, especially when the fluorophore-to-target ratio is one-to-one. On the other hand, there is only one chance to detect a molecule before it is bleached, and thus completeness might become an issue.

## **PAINT with reversibly binding dyes**

Point accumulation for imaging in nanoscale topography (PAINT) takes a different approach to separating molecules in time, based on the principle of motion blurring: molecules that diffuse very fast with respect to the exposure time are not visible, only contributing to background fluorescence, while fluorophores that are moving slower produce a sharp image (Figure 1.8E). When diffusing freely, small-molecule dyes move very quickly, and thus with typical frame rates used in microscopy, they only become visible once they bind to their target macromolecule. Using reversible binders and very low concentrations ensures that binding events are sparse and dynamic, allowing timelapse single-molecule detection. The signal-to-noise ratio can be further enhanced by using fluorogenic probes that increase in fluorescence upon binding.

PAINT has opposite features, compared to PALM: although a high sampling rate is more easily reached due to multiple binding events, accurate counting and cluster analysis is harder to perform. Thus, PAINT is often used to image “structures”, such as the membrane or DNA, rather than individual proteins.

### **1.5.2 Live SPT with photoactivatable dyes**

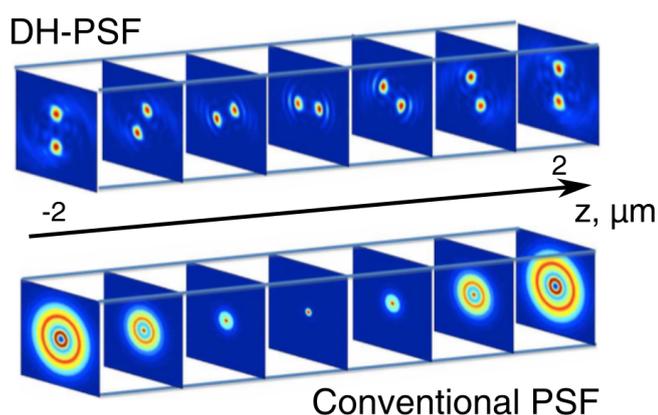
Obtaining a high-resolution image is not the only purpose for which SMLM can be used. Single-particle tracking (SPT) is a method which can be used to understand the diffusive behaviour of molecules. Instead of detecting and then bleaching individual fluorophores, molecules are observed for longer periods of time, which enables one to track and analyse their movement. This can be done either in reconstituted systems *in vitro*, or in live cells. In cells, high frame rates can be used to detect and track both the diffusing and the chromatin-tethered proteins, while long exposure time can be used to motion blur the moving molecules and specifically monitor those bound to chromatin. Furthermore, by measuring the period for which the molecules remain visible in these conditions, the dwell time on chromatin can be determined (see chapter 3).

Since proteins are normally present in cells in many copies, sparsity can either be achieved by under-labelling with low concentrations of a dye, or using photoactivatable

fluorophores, which enables more data to be collected from the same cell.

### 1.5.3 3D single-molecule microscopy with DH-PSF

To obtain three-dimensional SMLM reconstructions or to track molecules diffusing in 3D, several methods can be used. In this work, the double-helix point spread function (DH-PSF) approach was employed (Pavani et al. 2009) (Figure 1.9). A phase mask, introduced into the emission path of the microscope, transforms the canonical double cone-shaped PSF into a double helix, which in a perpendicular plane appears as a double peak. The pitch of the helix corresponds to the position of the emitter with respect to the focus of the system, allowing one to determine its z-coordinate, while the midpoint between the lobes is its position in x and y. To enable mapping of the z-coordinate to the DH-PSF angle, a calibration series is taken before every imaging session by recording the PSF of point emitters at different positions in z. Due to the inherent symmetry, the range of the DH-PSF is  $\pi$ , which corresponds to about 4  $\mu\text{m}$ . The mask is designed in such a way that beyond this, the PSF gradually becomes defocused.



**Figure 1.9:** DH-PSF microscopy enables mapping of the emitter's z-coordinate by transforming the conventional double-cone 3D PSF (bottom) into a double helix (top). Modified from Pavani and Piestun (2008).

## 1.6 Aims

The general aim of this study was to characterise heterochromatin structure and features in ground-state mESCs and upon global reorganisation of the genome at the onset of differentiation. Studying heterochromatin in this context served a dual role: firstly, unraveling the mechanism behind genome reorganisation, and secondly, better understanding heterochromatin organisation through relating physical parameters to dynamic changes in the cell.

In particular, HP1 $\beta$  was chosen as the focus of the project, due to its localisation and functions both in eu- and heterochromatin, as well as the Hi-C contact profile phenotype of the *Cbx1,3* DKO mESCs. Three goals were set: i) characterise the distribution of HP1 $\beta$  and other heterochromatic factors and marks in naive, formative and primed mESCs (chapter 2); ii) investigate the dynamics of HP1 $\beta$  in mESCs in the context of different models of heterochromatin organisation (chapter 3); and iii) relate changes in genome structure to redistribution of HP1 and other proteins. To achieve the latter two aims, novel methods were applied or developed. In chapter 3, various analysis methods of single-particle tracking data for HP1 $\beta$  in eu- and heterochromatin are presented, alongside a critical review of the physical features of heterochromatin they may reveal. In chapter 4, the development of 3D DH-PSF SMLM super-resolution imaging of entire mammalian nuclei in conjunction with snHi-C is described.

## Chapter 2

# Investigation of heterochromatin features during exit of mESCs from naive pluripotency

*The cell lines with HaloTagged HP1 used in this project were produced by Dr Maike Wiese. The RT-qPCR experiments were conducted in Brian Hendrich's lab at the Cambridge Stem Cell Institute, with help from Dr Nicola Reynolds.*

### 2.1 Introduction

### 2.2 Characterisation of the HP1 $\beta$ -Halo(N) cell line

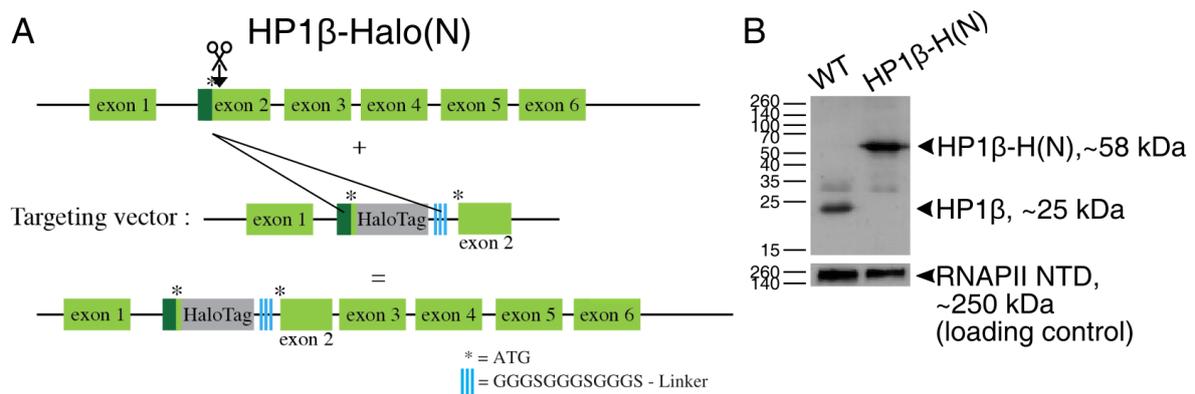
To enable antibody-free super-resolution and live cell imaging of HP1 $\beta$  (chapter 3), mESCs with HP1 $\beta$  tagged with a HaloTag were constructed by Dr Maike Wiese. With a view to potentially using the same cell line for single-nucleus Hi-C experiments (chapter 4), targeting was performed in a female F1 hybrid line between *Mus musculus domesticus* 129/Ola and *Mus musculus castaneus*. The endogenous *Cbx1* gene was genetically engineered using the CRISPR-Cas9 technology to have the HaloTag at either the N- or the C-terminus of the HP1 $\beta$  protein. The resultant mESC lines will be

referred to in this work as HP1 $\beta$ -Halo(N) and HP1 $\beta$ -Halo(C), respectively. However, characterisation of the lines suggested that HP1 $\beta$  function could have been slightly perturbed by the tag in HP1 $\beta$ -Halo(C) cells, and thus all experiments were performed with HP1 $\beta$ -Halo(N), while HP1 $\beta$ -Halo(C) mESCs were only used in the initial single-particle tracking experiment trials (chapter 3).

## 2.2.1 Western blots

An initial characterisation of the HP1 $\beta$ -Halo(N) line was performed to check whether the editing of the *Cbx1* gene had been successful, whether the resultant clones were homo- or heterozygous, and if the protein was expressed.

After targeting the N-terminus of *Cbx1* (Figure 2.1A), the positive clone was identified by fluorescent labelling and fluorescence-activated cell sorting (FACS) and confirmed by microscopy. The genotype of the cells was probed, using PCR, but the experiment was initially unsuccessful. Thus, the presence and levels of the tagged protein were confirmed by western blot, which indicated that the targeting was successful, and the clone was potentially homozygous (Figures 2.1B). However, it could not be excluded at that point that the line could have been hemizygous with one of the alleles knocked out during targeting.



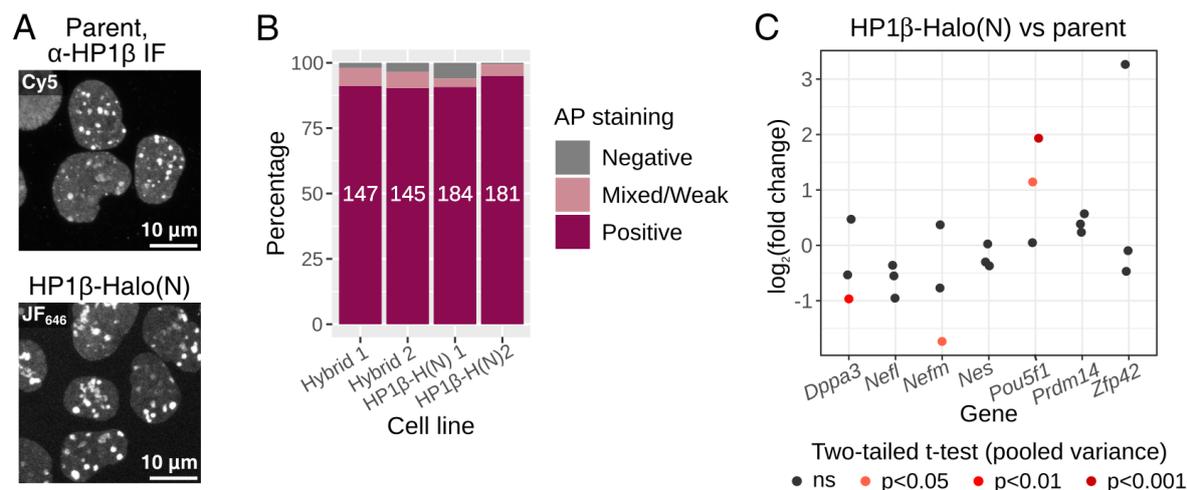
**Figure 2.1:** A. CRISPR targeting strategy for insertion of an N-terminal HaloTag to the *Cbx1* open reading frame. Figure by Dr M. Wiese. B. Anti-HP1 $\beta$  western blot for the wild-type parent (WT) and the HP1 $\beta$ -Halo(N) lines.

## 2.2.2 Protein localisation and function

Following the initial characterisation, the tagged protein's correct localisation and function were tested in the new cell line.

### Microscopy

First, the cells were checked for the expected nuclear pattern of HP1 $\beta$  by microscopy. The distribution of HP1 $\beta$  in the new cell line was very similar to that in the parent wild-type line, suggesting the protein localisation was not affected by the HaloTag (Figure 2.2A).



**Figure 2.2:** A. Confocal images of HP1 $\beta$ , labelled in the HP1 $\beta$ -Halo(N) cell line with HaloTag ligand-JF<sub>646</sub> to confirm the presence of the tag, as well as the localisation of the protein. Immunofluorescence using an anti-HP1 $\beta$  antibody in the parent cell line is shown for comparison. Images are maximum projections of z-stacks, with brightness adjusted for clarity. Scale bar, 10  $\mu$ m. B. Alkaline phosphatase assay of the parent F1 hybrid and the HP1 $\beta$ -Halo(N) cells, grown in serum/LIF conditions for three passages. Each bar corresponds to a biological replicate (two per cell line), with the number of colonies counted shown in the middle. C. RT-qPCR analysis of RNA levels of genes, misregulated upon HP $\beta$  knockout (Mattout et al. 2015): *Dppa3*, *Nefl*, *Nefm*, *Nes* and *Prdm14*, and additional pluripotency markers *Pou5f1* and *Zfp42*, in HP1 $\beta$ -Halo(N) cells, compared to the parent hybrid cell line. *Gapdh* was used as a reference gene. Each point represents a biological replicate, and the colour shows the statistical significance of the fold-change, calculated using pooled variance from all samples within the replicate (see section 6.2.5). None of the genes displayed significantly different RNA levels, when testing using all biological replicates (two-tailed t-test).

## Alkaline phosphatase assays

Mattout et al. (2015) suggest that in serum/LIF conditions, knockout of HP1 $\beta$  leads to an increased propensity to differentiate. Therefore, an alkaline phosphatase assay was performed on the control parent and HP1 $\beta$ -Halo(N) cells, grown in serum/LIF for three passages (Figure 2.2B). The results showed that the N-terminal tagging of HP1 $\beta$  did not lead to an increase in spontaneous differentiation, suggesting that the function of HP1 $\beta$  was intact.

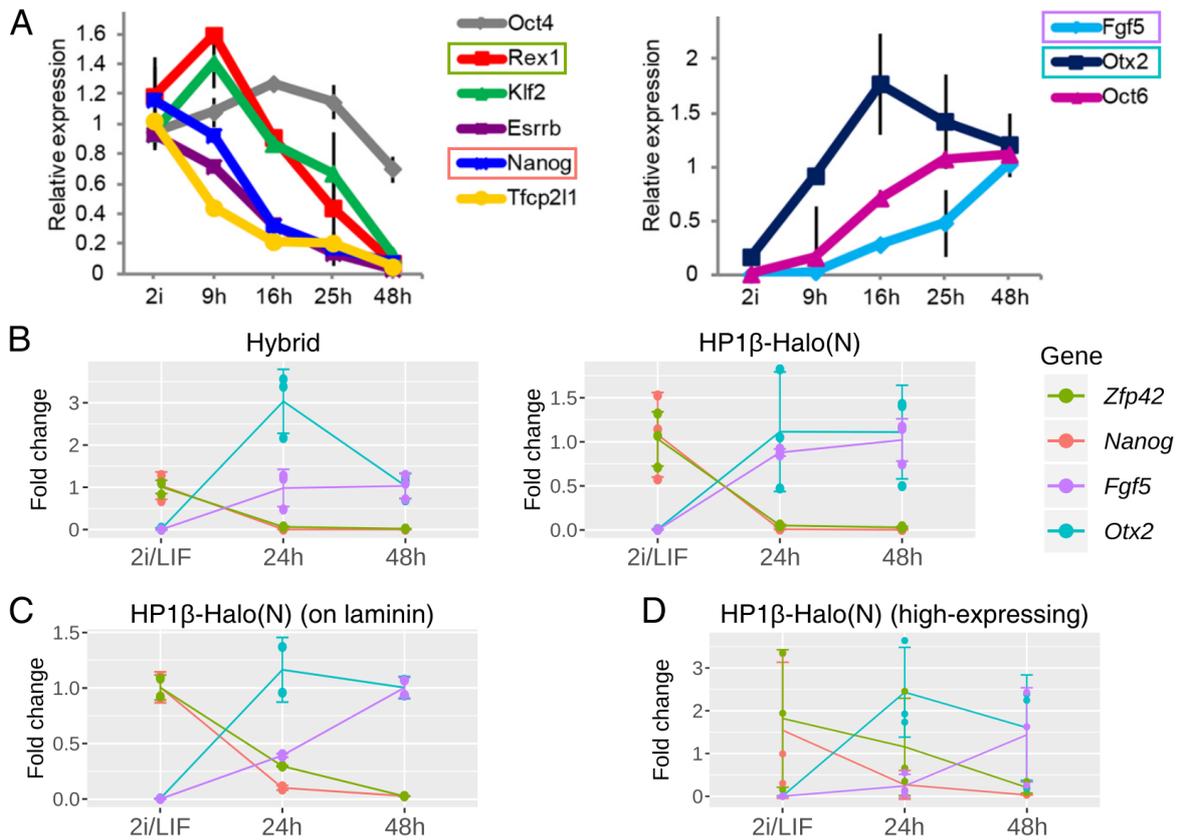
## RNA levels of HP1 $\beta$ -controlled genes with RT-qPCR

Next, the RNA levels of genes known to be misregulated upon HP1 $\beta$  knockout (Mattout et al. 2015), as well as several classical pluripotency markers, were measured by reverse transcription-quantitative PCR (RT-qPCR) (Figure 2.2C). Most genes did not show significant changes in any of the biological replicates, and some genes showed a difference only in one out of four. Furthermore, applying statistical tests to the biological replicates revealed that no genes were significantly up- or downregulated in the HP1 $\beta$ -Halo(N) line. All in all, there was no evidence to suggest that the function of HP1 $\beta$  was perturbed in the HP1 $\beta$ -Halo(N) cell line. In contrast, the HP1 $\beta$ -Halo(C) cells displayed mild, but significant misregulation of *Nes* and *Pou5f1* (not shown).

### 2.2.3 Cell differentiation

To be able to investigate the formative and primed states defined in Kalkan et al. (2017), it was important to check that the new cell line differentiated with the same kinetics as the mESCs used in their study (Figure 2.3A). Therefore, RT-qPCR was used to analyse changes in RNA levels of the pluripotency markers *Zfp42* (Rex1) and *Nanog*, and the differentiation markers *Otx2* and *Fgf5* (Figure 2.3B).

The hybrid cells showed differentiation kinetics almost identical to that published, indicating that 24h and 48h post-2i withdrawal were suitable time points for investigation. The HP1 $\beta$ -Halo(N) cells' behaviour was very similar to the parent line, while the HP1 $\beta$ -Halo(C) cells displayed somewhat slower differentiation (not shown).



**Figure 2.3:** RNA levels of pluripotency and differentiation markers upon exit of cells from naive pluripotency. A. Published data for RGd2 cells (E14 background) (Kalkan et al. 2017). Genes investigated in this work are boxed with corresponding colours. B. RT-qPCR data for parent F1 hybrid and HP1 $\beta$ -Halo(N) mESCs. C. Data for HP1 $\beta$ -Halo(N) cells, differentiated on poly-L-ornithine- and laminin-coated glass. D. Data for HP1 $\beta$ -Halo(N) cells, flow-sorted to purify cells with wild-type levels of HP1 $\beta$  expression. The colour scheme is consistent in B, C and D. In B, C and D, expression levels are shown as fold change relative to 2i/LIF for *Nanog* and *Zfp42*, and to 48h samples for *Fgf5* and *Otx2*; each point is a biological replicate, and error bars indicate standard deviation. The reference gene used in A was *Gapdh*, in B and C *Ppia*, in D *Atp5a1*.

## Differentiation on laminin-coated glass

The surface on which cells grow affects their morphology and in some cases even their identity and differentiation capacity (Chowdhury et al. 2010; Ozdil et al. 2019). Since many of the microscopy experiments in this study were carried out on glass dishes, coated with poly-L-ornithine and laminin for better adhesion, the RNA levels of the pluripotency and differentiation markers in HP1 $\beta$ -Halo(N) cells grown on this substrate were assessed upon withdrawal of 2i (Figure 2.3C). The results were very similar to cells grown on gelatin.

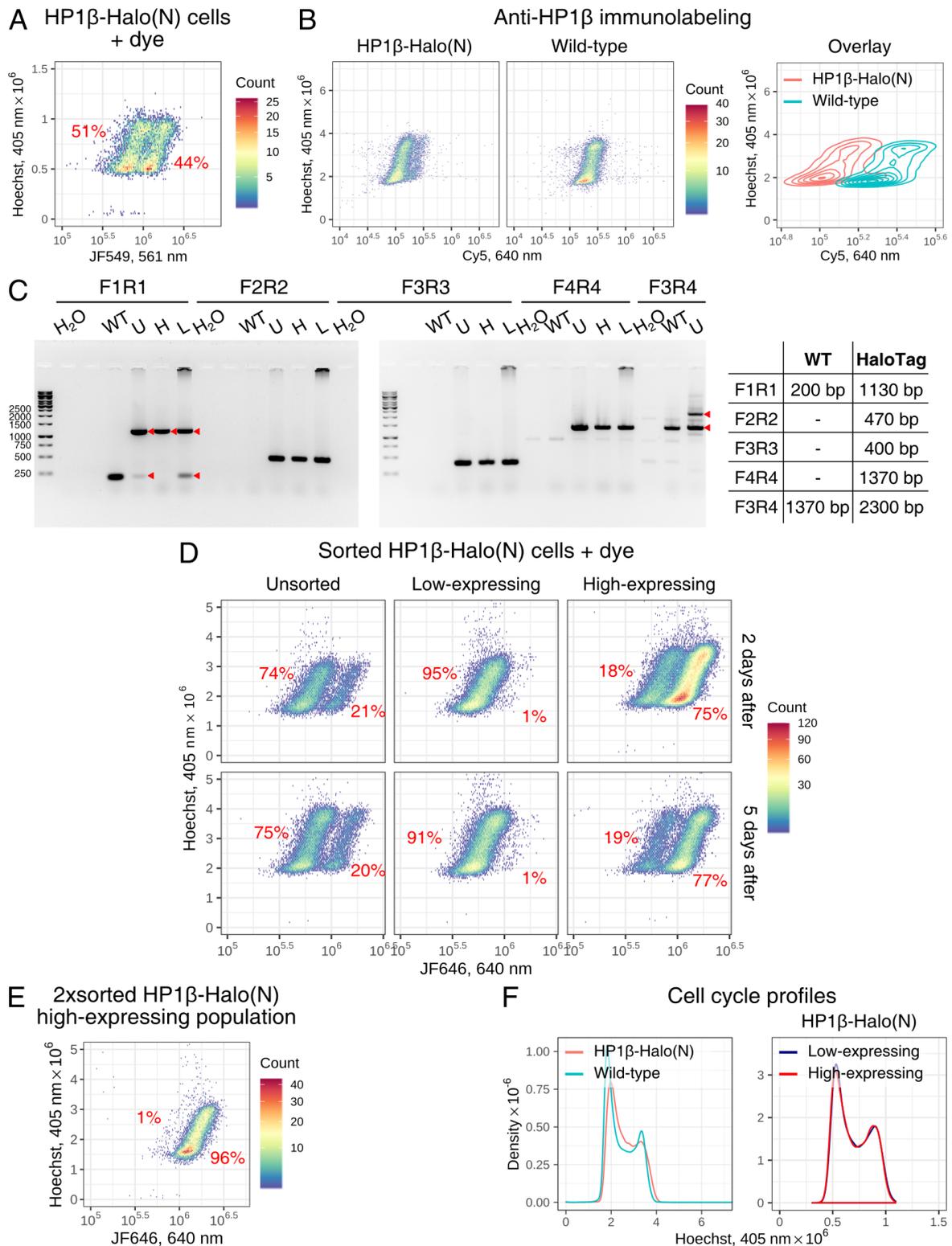
Overall, mESCs lines were made that expressed HP1 $\beta$ , fused with HaloTag either on the N- or the C-terminus, from the endogenous locus. The HP1 $\beta$ -Halo(N) cells did not show any signs of the tag interfering with the protein function, while the HP1 $\beta$ -Halo(C) cells displayed a delayed differentiation timecourse, potentially indicating that the HP1 $\beta$  function was perturbed in these cells.

### 2.2.4 Separating the populations of HP1 $\beta$ -Halo(N) with differential HP1 $\beta$ expression

When conducting initial experiments to assess the HP1 $\beta$  localisation during differentiation (section 2.4.2), extreme variability in staining intensity of HP1 $\beta$ -Halo(N) cells was observed. Flow cytometry analysis of live mESCs upon labelling of HP1 $\beta$  with Janelia Fluor 549 (JF<sub>549</sub>) revealed the presence of two populations of cells with about a two-fold difference in fluorescence intensity levels (Figure 2.4A). Further experiments were carried out to understand the source of this variation.

#### Immunostaining

Firstly, immunolabelling with anti-HP1 $\beta$  antibody was used to confirm that the JF<sub>549</sub> staining levels reflected differences in HP1 $\beta$  protein expression (Figure 2.4B). An analogous experiment in the parent mESC line indicated that the HP1 $\beta$  levels in the wild-type cells matched these of the higher-expressing population of HP1 $\beta$ -Halo(N) cells



**Figure 2.4:** A. 2D histogram from flow cytometry analysis of HP1 $\beta$ -Halo(N) cells, stained with Hoechst and HaloTag ligand-JF<sub>549</sub>. The approximate proportions of the HP1 $\beta$ -high and -low expressing populations are shown. Performed on 24/02/2020. B. Data from flow cytometry analysis of HP1 $\beta$ -Halo(N) and wild-type parent mESCs, labelled with Hoechst and anti-HP1 $\beta$  antibody. High- and low-expressing populations are present in the HP1 $\beta$ -Halo(N), but not in the parent line.

*Legend continued on the next page.*

(Figure 2.4B). Thus, it appeared that the HP1 $\beta$ -Halo(N) cells were a mix between two populations, homo- and hemizygous for tagged *Cbx1*, which was not detected in western blots (section 2.2.1), since all cells expressed exclusively the tagged protein.

### Genotyping PCR

The two populations of the HP1 $\beta$ -Halo(N) cells were separated using FACS, based on the intensity of HP1 $\beta$  staining with JF<sub>646</sub>. To understand the genetic basis of the differences in HP1 $\beta$  levels, genotyping PCR for the targeted locus was conducted on the unsorted cells, as well as each of the populations separately (Figure 2.4C). The unsorted cells and the low-expressing population had bands corresponding to both successful HaloTag targeting and to the untargeted wild-type gene, while for the highly-expressing population the latter was not visible. This confirmed that monoallelic expression in hemizygously targeted cells was the source of the lower level of HP1 $\beta$ .

### Stability of expression

Finally, it was noticed that the relative proportions of the two populations varied between two experiments conducted five months apart (Figures 2.4A, D). Since the cells used in both experiments were taken from the same frozen batch, but cultured for a different amount of time, it was possible that the homo- and hemizygous populations may have had different growth rates.

Flow cytometry analysis of mESCs FACS-sorted based on HP1 $\beta$  expression levels indicated that the relative proportions of the populations were stable for several

**Figure 2.4:** The right panel shows an overlay, indicating that the high-expressing population of HP1 $\beta$ -Halo(N) corresponds to wild-type cells. C. Genotyping PCR results for wild-type parent line (WT), and unsorted (U), sorted and highly-expressing (H) or sorted and lowly-expressing (L) HP1 $\beta$ -Halo(N) mESCs, with negative control reactions (H<sub>2</sub>O). The expected fragment sizes are summarised in a table on the right. The bands indicated with red arrowheads were sent for sequencing, but it did not reveal the nature of the knockout mutation. D. Flow cytometry analysis of HP1 $\beta$ -Halo(N) cells, FACS-sorted into high- and low-expressing populations, 2 and 5 days after sorting. Performed on 22-27/07/2020. E. Analysis of HP1 $\beta$  expression in HP1 $\beta$ -Halo(N) cells, FACS-sorted twice to enrich for the highly-expressing population. F. Hoechst profiles of the unsorted HP1 $\beta$ -Halo(N) mESCs and the parent line (left; same data as shown in B) high- and low-expressing HP1 $\beta$ -Halo(N) populations (right; same data as shown in the top left panel in D).

passages (Figure 2.4D). Thus, at least on short timescales, it appeared that neither of the populations had a growth advantage. It remained unclear why their relative proportions were different between measurements.

For subsequent experiments, only the highly-expressing population of HP1 $\beta$ -Halo(N) was used. Sorting was performed twice to ensure a high level of enrichment for this population (Figure 2.4E), and the expression levels were periodically checked by flow cytometry. The differentiation progression of the sorted cells was similar to the wild-type (Figure 2.3D). The low-expressing hemizygous population was also kept for potential future experiments.

### **Cell cycle progression is not affected by the HaloTag**

The cell cycle progression between the wild-type and tagged mESCs was also compared, using flow cytometry (Figure 2.4F). There was no significant difference neither between the homo- or hemizygously tagged HP1 $\beta$ -Halo(N) cells, nor between the HP1 $\beta$ -Halo(N) and the parent line. This indicated that neither fusion of HP1 $\beta$  with HaloTag nor even a 50% reduction in its expression induced any effects leading to checkpoint activation.

## **2.3 Differentiation in the absence of HP1 isoforms**

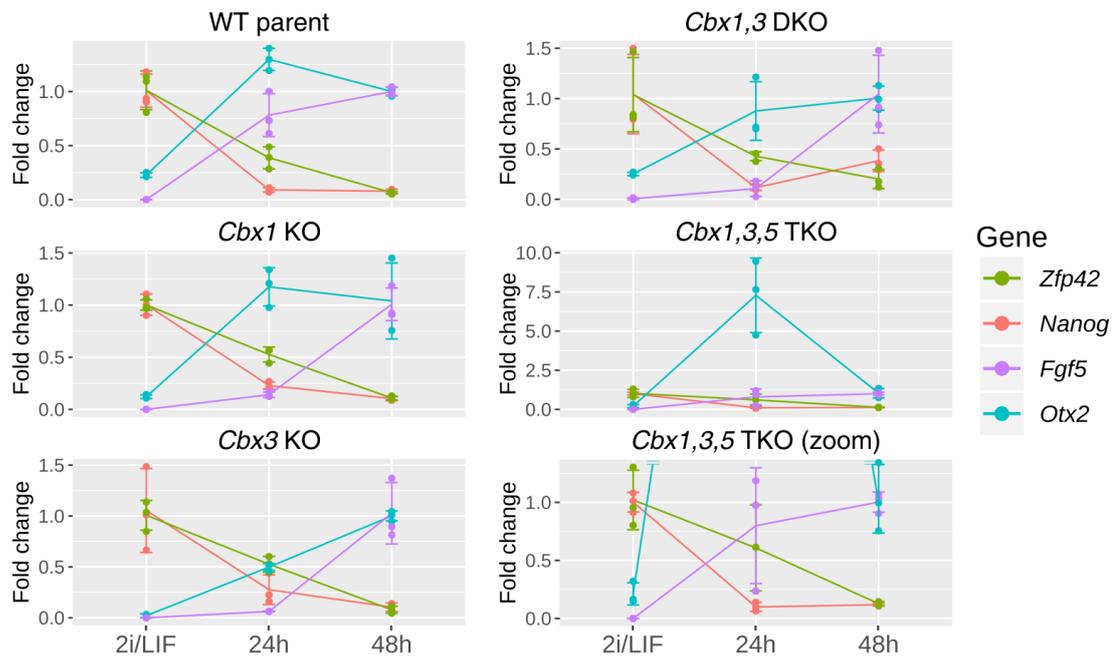
A double knockout of HP1 $\beta$  and HP1 $\gamma$  leads to subtle changes in the Hi-C contact length distribution in naive mESCs. Furthermore, this profile does not change upon removal of 2i and LIF, suggesting that the genome structure of these cells may not be undergoing the same global rearrangement as in the wild-type (see section 1.3.3).

These observations prompted me to ask whether differentiation of mESCs was affected by single, double or triple knockouts of the HP1 isoforms. If the kinetics of differentiation in the HP1 $\beta$  and HP1 $\gamma$  double knockout was the same as in the wild-type, it would imply that genome reorganisation is not immediately necessary for exit from naive pluripotency, and that the knockout affects one without severely affecting the other. If the progression through differentiation was different, the lack of chromatin reorganisation could be slowing down or even preventing exit from naive pluripotency.

Alternatively, the double knockout itself could be impeding differentiation, resulting in the lack of changes in the Hi-C contact profile.

To address this question, the differentiation kinetics of HP1 knockout cell lines (Ostapcuk et al. 2018) was characterised using RT-qPCR for the naive state and differentiation markers (Figure 2.5). The strongest effect was observed in the *Cbx3* (encoding HP1 $\gamma$ ) knockout, where changes in levels of *Otx2* and *Fgf5* were noticeably slower than in the wild-type. MESCs lacking *Cbx1* (encoding HP1 $\beta$ ) or both *Cbx1* and *Cbx3* had milder deviations in *Fgf5* levels. Triple knockout of all HP1 isoforms lead to an interesting phenotype of very high *Otx2* induction 24h after 2i withdrawal, but the general trends were unaltered.

Overall, despite the fact that double- and triple-knockout cells exhibited poor growth, it appeared that lack of HP1 proteins did not dramatically affect exit from naive pluripotency. *Zfp42* and *Nanog* downregulation proceeded in a timely manner in all of the lines, unlike for example in *Dnmt3a/b* double knockout (Li et al. 2017). Neither *Otx2* nor *Fgf5* upregulation failed completely for any of the knockout lines, although *Cbx3* knockout in particular delayed it. Thus, firstly, although the genome structure is altered in the *Cbx1,3* double-knockout cells in a manner similar to the formative-state wild-type cells, this does not facilitate exit from naive pluripotency, and may delay it slightly. In addition, it appears likely that the failure to change the 3D structure of the nucleus in the absence of HP1 $\beta$  and HP1 $\gamma$  is not directly linked to a failure to differentiate. Testing the expression of other formative state markers could help determine whether the establishment of the formative state is perturbed by knockout of HP1 proteins. Furthermore, it might be interesting to probe the Hi-C contact distribution profile of the knockout cells 48h after removal of 2i.



**Figure 2.5:** RNA levels of pluripotency and differentiation markers upon exit of cells, deficient of one or several HP1 isoforms, from naive pluripotency. The mESCs used in this experiment are constitutive and conditional single, double or triple knockouts (KO, DKO and TKO, respectively), kindly provided by Marc Bühler (Ostapcuk et al. 2018). Expression levels are shown as fold change relative to 2i/LIF for *Nanog* and *Zfp42*, and to 48h samples for *Fgf5* and *Otx2*; each point is a biological replicate, and error bars indicate standard deviation. The results for *Cbx1,3,5* TKO are shown twice with different scales. The reference gene used was *Gapdh*.

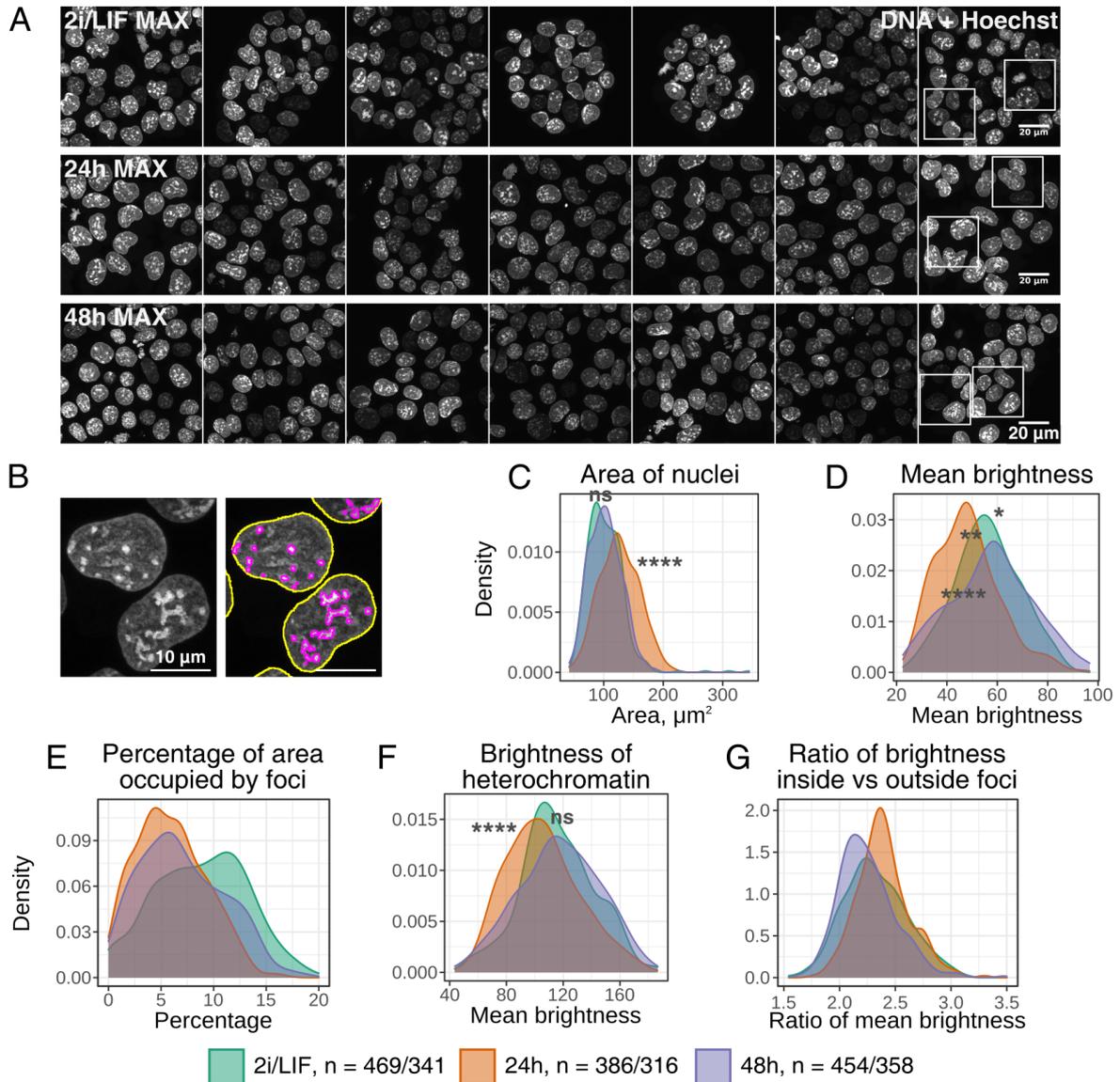
## 2.4 Levels and distribution of heterochromatic factors during differentiation

### 2.4.1 Heterochromatin decondensation upon exit from naive pluripotency

It has been previously demonstrated that at least on a fine scale, chromatin decondenses upon exit from naive pluripotency (Chalut et al. 2012; Pagliara et al. 2014). I was interested to see if it could be extended to the large-scale partitioning of the genome into eu- and heterochromatin.

To this end, mESCs in 2i/LIF and upon exit from naive pluripotency were stained with Hoechst and imaged with a confocal microscope (Figure 2.6A). The area of the nuclei was visibly larger in the formative state, confirmed by segmentation and quantification (Figures 2.6B,C). Interestingly, the mean staining intensity was significantly reduced in the formative state (Figure 2.6D). However, the variability in Hoechst staining could have affected the measurements substantially, and the results obtained from projections are also hard to interpret. Thus, this finding must be confirmed, either using an optimised staining procedure or an orthogonal technique, ideally combined with volumetric analysis.

Although there was no obviously visible change in the DNA density pattern, quantitative analysis of segmented foci within each nucleus suggested that the area occupied by heterochromatin decreased upon differentiation (Figures 2.6B,E). The DNA density inside heterochromatin also decreased transiently in the formative state (Figure 2.6F). However, the density of heterochromatic foci relative to euchromatin appeared to slightly increase in the formative state, and decrease in primed mESCs (Figure 2.6G). Thus, although in general chromatin appeared to decondense, this process occurred to a stronger extent in euchromatin than in heterochromatin.

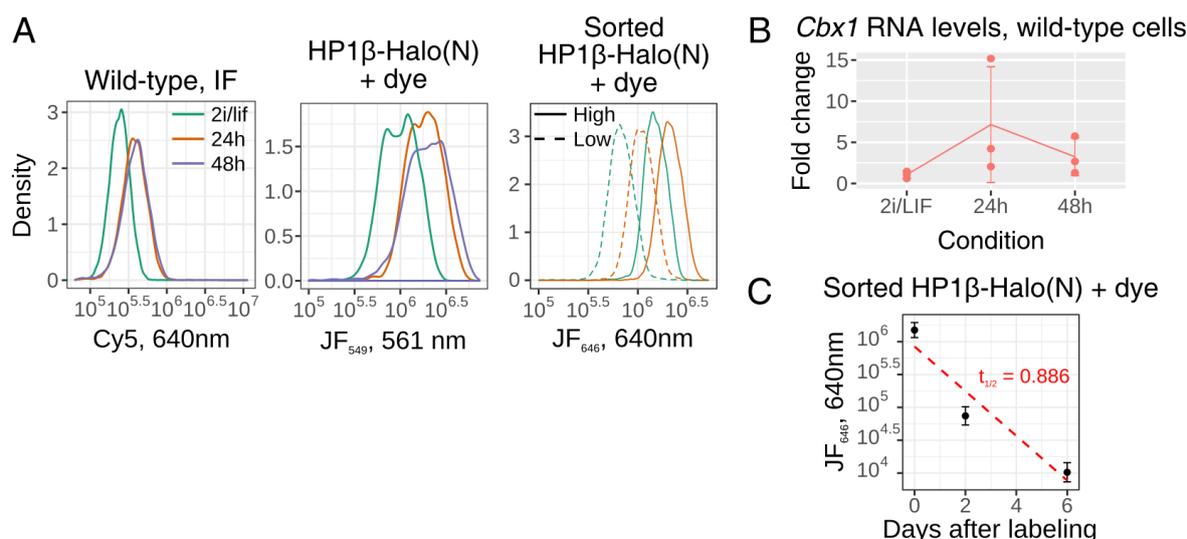


**Figure 2.6:** A. Representative scanning confocal microscopy images of DNA, labelled with Hoechst, in HP1 $\beta$ -Halo(N) mESCs in naive, formative and primed pluripotent states. Shown are maximum projections of 10  $\mu$ m-thick z-stacks. The intensity has been adjusted consistently between all images. The white boxes indicate the regions shown in Figure 2.9. Scale bar, 20  $\mu$ m. B. An example of segmentation of nuclei (yellow) and heterochromatic foci (pink). Scale bar, 10  $\mu$ m. C-G. Quantification of the confocal microscopy data. 15 images were analysed per timepoint, and the parameters were calculated on a per-nucleus basis. Poorly stained nuclei were manually excluded from the calculations shown in D-G, thus  $n$  shows the entire sample size/number of bright nuclei analysed. One-way ANOVA with Tukey post-hoc testing was applied: ns, non-significant, \*,  $p < 0.05$ , \*\*,  $p < 0.01$ , \*\*\*,  $p < 0.001$ , \*\*\*\*,  $p < 0.0001$ ; all samples in D and F were significantly different.

## 2.4.2 HP1 $\beta$ levels and pattern change during differentiation

### HP1 $\beta$ expression increases in formative phase

Proteomics data from Kalkan et al. (2017) indicated that protein levels of HP1 $\beta$ , but not HP1 $\alpha$  or HP1 $\gamma$ , increase in the formative state approximately 1.5-2-fold. To characterise HP1 $\beta$  levels in our system, the expression in differentiating HP1 $\beta$ -Halo(N) or wild-type cells was measured using flow cytometry (Figure 2.7A). Indeed, these assays showed a 1.6-2-fold increase in the abundance of HP1 $\beta$  between the cells in 2i/LIF and 24h after 2i withdrawal, and the levels stayed high after 48h. The rise in the RNA levels of *Cbx1* in the formative state suggested that transcriptional upregulation could contribute to the increase in HP1 $\beta$  protein levels, although the data was quite variable (Figure 2.7B).



**Figure 2.7:** A. Density plots of HP1 $\beta$  levels in single mESCs at the three differentiation timepoints, measured using flow cytometry. In the left panel, data for wild-type cells immunolabelled using an anti-HP1 $\beta$  antibody is shown. In the middle and right panels, fluorescence intensity measurements in live HP1 $\beta$ -Halo(N) mESCs stained with Janelia Fluor dyes are plotted. In the right panel, HP1 $\beta$ -Halo(N) cells were sorted into high- and low-expressing populations. B. The RNA levels of *Cbx1* were measured with RT-qPCR in mESCs upon exit from naive pluripotency. The data is shown as fold change with respect to the levels in 2i/LIF, and *Atp5a1* was used for normalisation. Individual points are biological replicates, and the error bars indicate the standard deviation. C. Mean  $\pm$  SD fluorescence levels in single high-expressing HP1 $\beta$ -Halo(N) cells, labelled with JF<sub>646</sub> and left to grow for 2 or 6 days. Linear fit to the log-transformed data and the corresponding decay half-time in days are shown.

The rate of fluorescence decay after labelling HP1 $\beta$  with JF<sub>646</sub> was similar to the mESC growth rate, indicating slow turnover of HP1 $\beta$  (Figure 2.7C). Thus, it was

reasonable to expect a delay between the decrease in *Cbx1* transcription and the drop in protein levels.

### **HP1 $\beta$ concentrates in foci in the primed state**

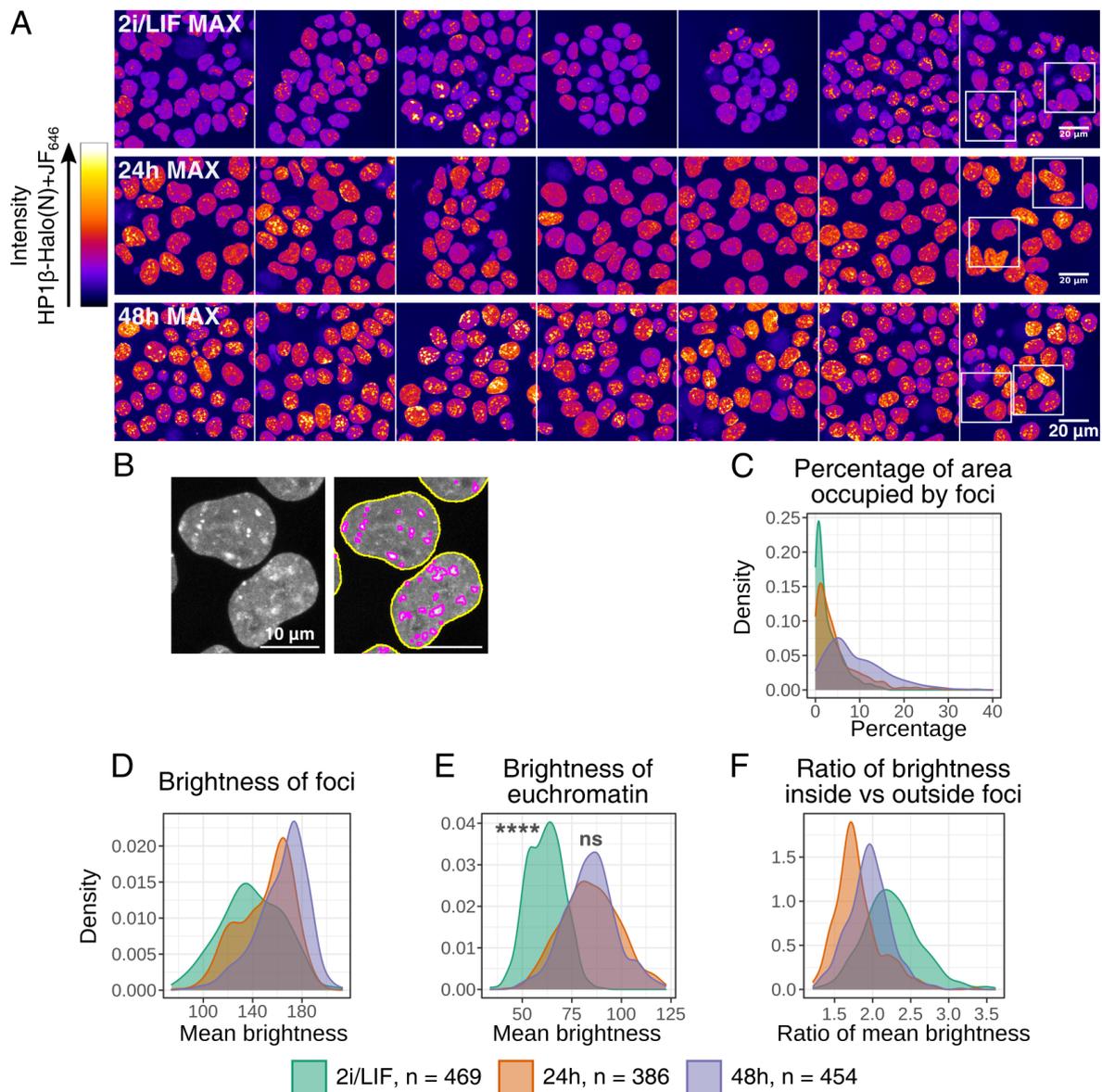
Since the genomic compartment segregation weakens in the formative state, while the levels of HP1 $\beta$  increase, I asked whether the distribution of HP1 $\beta$  in the nucleus changes upon exit from naive pluripotency.

To address this question, confocal images of HP1 $\beta$  were taken at the three stages of pluripotency (Figure 2.8A). The brightness of the signal clearly increased in more differentiated cells compared to cells in 2i/LIF, consistent with previous results. Interestingly, the pattern of HP1 $\beta$  in the nucleus appeared to be visibly different in formative- and primed-state cells: the foci appeared to be more pronounced (i.e. brighter and perhaps larger or more numerous) in the latter.

These observations were confirmed by segmentation and quantitative analysis (Figure 2.8B). The proportion of the area covered by heterochromatin increased most between the formative and the primed states (Figure 2.8C), and the mean brightness (i.e. HP1 $\beta$  concentration) of heterochromatin increased gradually throughout the protocol (Figure 2.8D). Interestingly, the distributions of the mean brightness of foci appeared multimodal, perhaps with the relative proportions of the different populations changing upon differentiation. In contrast, the mean brightness of euchromatin (area outside the foci) increased dramatically 24h after 2i withdrawal, but did not change afterwards (Figure 2.8E). Consistently, the relative difference in HP1 $\beta$  concentration between eu- and heterochromatin was largest in naive mESCs, then dropped at 24h and increased somewhat by 48h post 2i withdrawal (Figure 2.8F). Repeating this analysis for single confocal slices through the middle of the cell produced similar results.

### **HP1 $\beta$ moves into DNA-dense regions in the primed state**

Curiously, the HP1 $\beta$  and DNA distributions in cells looked somewhat different (Figure 2.9A). Although large chromocentres appeared in both channels, small foci of HP1 $\beta$



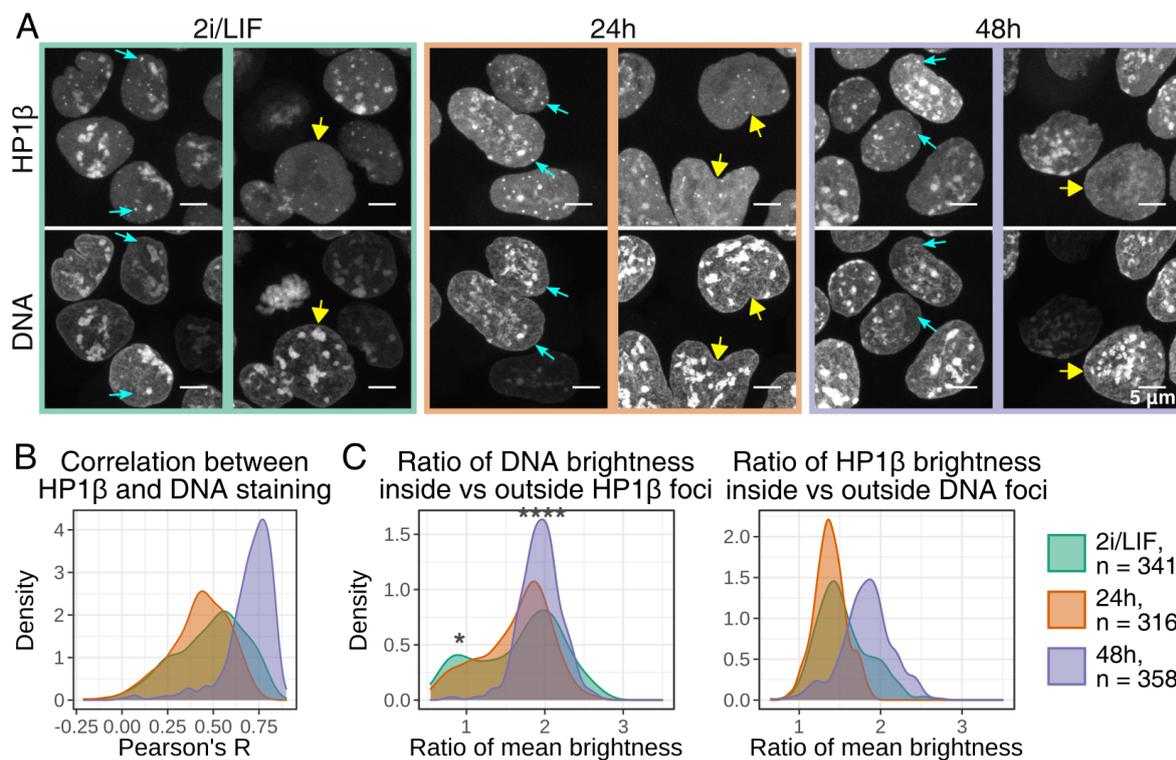
**Figure 2.8:** A. Representative scanning confocal microscopy images of HP1 $\beta$ , labelled with JF<sub>646</sub>, in HP1 $\beta$ -Halo(N) mESCs in naive, formative and primed pluripotent states. The same sample as in Figure 2.6 is shown here. Images are maximum projections of 10  $\mu$ m-thick z-stacks, colour denotes the true unadjusted intensity. Scale bar, 20  $\mu$ m. B. An example of segmentation of nuclei (yellow) and HP1 $\beta$  foci (pink). Scale bar, 10  $\mu$ m. C-F. Quantification of the confocal microscopy data. 15 images were analysed per timepoint, and the parameters were calculated on a per-nucleus basis, with  $n$  denoting the sample size. Since 20, 5 and 8 nuclei in the 2i/LIF, 24h and 48h conditions, respectively, did not have any foci, they were not included in panels D and F. Significance of differences was tested with one-way ANOVA with Tukey post-hoc (ns, non-significant; \*\*\*\*,  $p < 0.0001$ ); all samples in panels C, D and F are significantly different.

were present that did not correspond to a DNA-dense area. It appeared that such small HP1 $\beta$  domains were more prevalent in naive and formative states than in the primed state. On the other hand, in some cells HP1 $\beta$  did not appear to localise to the DNA-dense chromocentres. Performing a colocalisation analysis revealed that the correlation between the DNA and HP1 $\beta$  distributions increased substantially in the primed state (Figure 2.9B).

The enrichment in DNA density within HP1 $\beta$ -marked foci displayed a bimodal distribution for naive- and formative-state cells, which confirmed the observation that in some cells, HP1 $\beta$  does not localise to chromocentres (Figure 2.9C, left). The fact that this tended to occur on a cell-by-cell, rather than focus-by-focus basis, suggested that this might be regulated on a global cellular level, e.g. in a cell cycle- or cell state-dependent manner. Interestingly, the distribution of HP1 $\beta$  enrichment in the DNA-dense domains was more unimodal (Figure 2.9C, right). This can probably be explained by the fact that in cells where HP1 $\beta$  does not concentrate in chromocentres, it only forms small foci in euchromatin. Thus, the ratio of HP1 $\beta$  density inside and outside the DNA foci in such cells will be close to 1, while the ratio of DNA concentration within and outside the HP1 $\beta$  foci will be substantially  $<1$ , i.e. much smaller than in cells where HP1 $\beta$  and DNA colocalise better. Both of these parameters increased to about 2 in the primed-state mESCs, consistent with the colocalisation analysis.

### **2.4.3 No changes are apparent in the HP1 $\alpha$ distribution in the formative state**

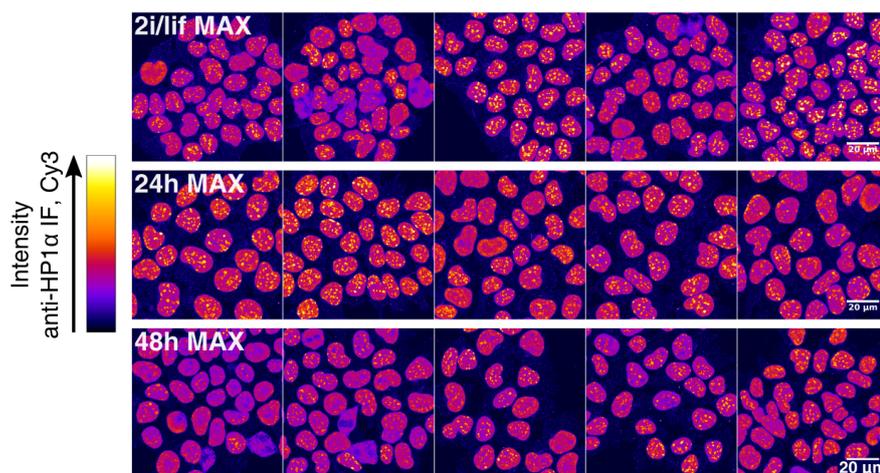
Since HP1 $\beta$  localises largely to chromocentres in our mESC lines, and HP1 $\beta$  is known to affect the stability of HP1 $\alpha$ -induced phase-separated droplets *in vitro* (Keenen et al. 2021), the distribution of HP1 $\alpha$  in the three pluripotent states was investigated using immunofluorescence (Figure 2.10). The pattern resembled that of HP1 $\beta$  in mESCs, but no change in the distribution of the protein was apparent in the formative state. However, while the area and brightness of HP1 $\beta$  foci increased in the primed state, those of HP1 $\alpha$  seemed to decrease at this time point, potentially suggesting that HP1 $\alpha$



**Figure 2.9:** A. Zoom-in into areas denoted by white boxes in Figures 2.6A and 2.8A. Cyan arrows show HP1 $\beta$  foci absent in the DNA channel, yellow triangle arrows show cells where HP1 $\beta$  does not localise to DNA-dense regions. Scale bar, 5  $\mu$ m. B. Colocalisation analysis, performed by calculating Pearson's correlation coefficient between pixel intensities in the two channels on a per-nucleus basis. All samples are highly significantly different (one-way ANOVA with Tukey's post-hoc). C. Quantification of DNA staining brightness ratio within and outside of HP1 $\beta$  foci, and vice versa, calculated on a per-nucleus basis. Statistical testing with one-way ANOVA and Tukey's post-hoc: \*,  $p < 0.05$ ; \*\*\*\*,  $p < 0.0001$ ; all samples in the right panel are highly significantly different. B and C show only Hoechst-bright nuclei.

in the chromocentres might be replaced by HP1 $\beta$  at this stage.

Because immunofluorescence is less quantitative than labelling with a small-molecule fluorophore, especially for adherent cells growing in colonies, it would be beneficial to confirm this finding using a more quantitative labelling approach. Although the samples in different conditions were prepared in parallel, accessibility of the antibody to its antigen may still vary between them, affecting the results.



**Figure 2.10:** Scanning confocal microscopy images of anti-HP1 $\alpha$  immunofluorescence in wild-type mESCs in naive, formative and primed pluripotent states. Samples were prepared in parallel and imaged together. Shown are maximum projections of 10  $\mu$ m-thick z-stacks, colour denotes the true unadjusted intensity. Scale bar, 20  $\mu$ m.

#### 2.4.4 H3K9me3 and H3K27me3 staining is not consistent with a replacement hypothesis

Tosolini et al. (2018) observed that in naive mESCs, chromocentres are sometimes marked with H3K27me3 modification alone or in combination with H3K9me2/3, rather than solely H3K9me2/3. However, it was no longer the case in EpiSCs, where H3K27me3 staining was not enriched in DNA foci, and an intermediate phenotype was observed in mESCs in serum/LIF. Thus, the authors proposed that upon exit from naive pluripotency, H3K27me3 is replaced by H3K9me2/3 in constitutive heterochromatin.

The quantitative ChIP-seq data obtained by my colleagues for mESCs in 2i/LIF

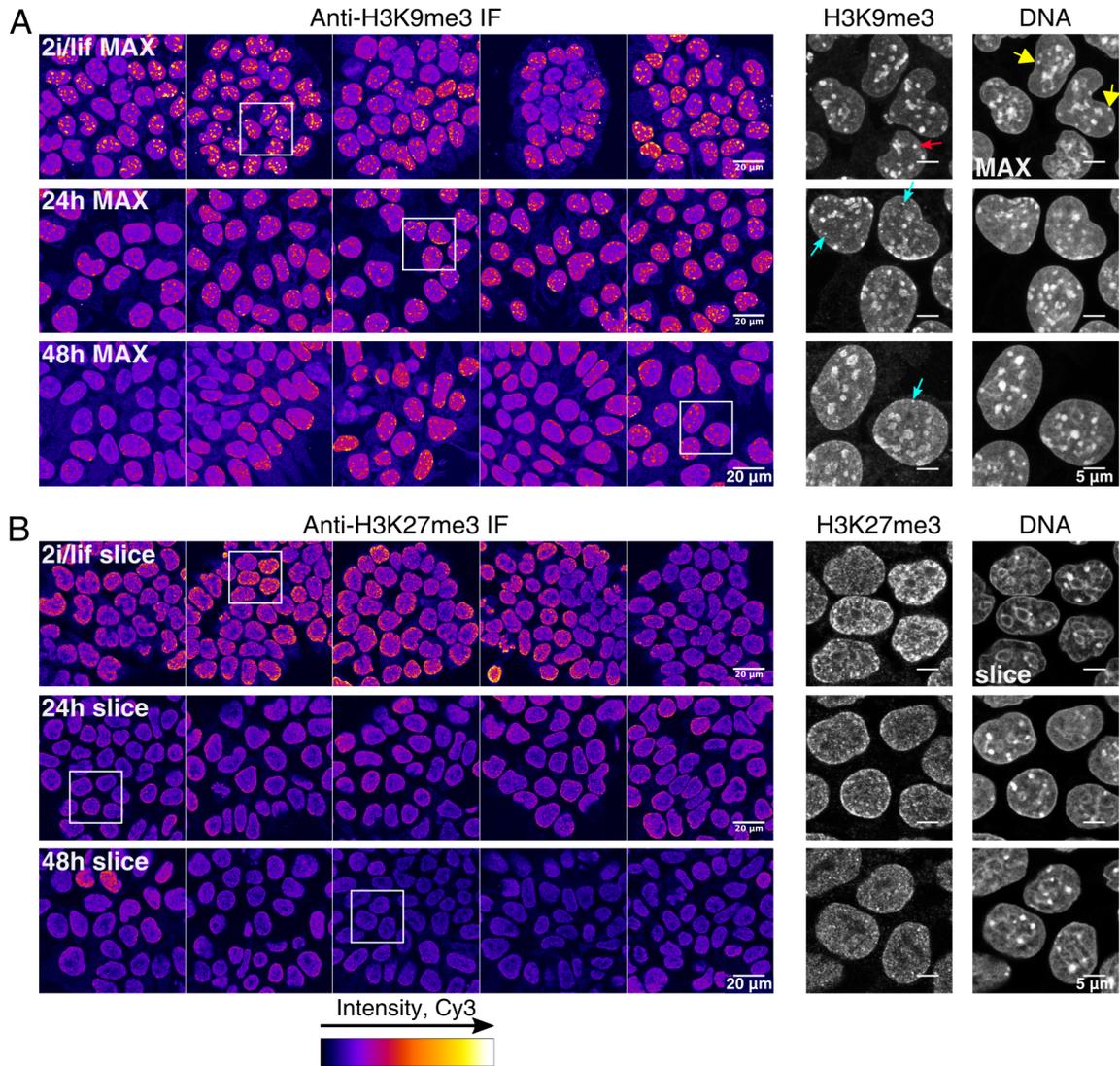
and upon withdrawal of 2i did not agree with this idea: the H3K9me3 levels on major satellite repeats were about two-fold lower 24h and 48h after the onset of differentiation than in the naive mESCs (see section 1.3.2). The levels of H3K27me3 at the same sites also dropped in the formative state, but were then restored at the 48h timepoint.

To investigate whether there were any changes in the spatial distribution of these modifications, immunofluorescence was performed. H3K9me3 staining agreed with the ChIP-seq data, and disagreed with findings of Tosolini et al. (2018): the focal pattern was apparent in the vast majority of naive mESCs (Figure 2.11A). 24h after the onset of differentiation, however, the chromocentre H3K9me3 enrichment became weaker, and it dropped further in the primed state. Interestingly, small puncta that did not correspond to DNA-dense regions and that were similar to the ones seen for HP1 $\beta$  (Figure 2.9A) could be observed, especially in formative-state cells. However, they were less pronounced than the HP1 $\beta$  speckles. Furthermore, large foci of DNA or H3K9me3 also did not always coincide with each other, but unlike for HP1 $\beta$ , this varied within individual cells.

H3K27me3 displayed a fine granular pattern in all conditions (Figure 2.11B). This pattern appeared to be coarser in cells in 2i/LIF, with occasional 0.5-1  $\mu$ m-size domains. However, these domains were not as large as chromocentres, and they almost never coincided DNA-dense regions. The staining was also often brighter in the nuclear periphery, although this might have been an artifact of antibody penetration. In agreement with quantitative ChIP-seq, the H3K27me3 staining intensity decreased in the formative state, but in contrast to the ChIP-seq, it was not restored in the primed state. The change in pattern also agreed with H3K27me3 peak mapping onto snHi-C structures indicating lower H3K27me3 spatial clustering in the formative state.

As with the HP1 $\alpha$  results, the immunofluorescence data for histone marks needs to be treated with caution due to potential differences in antibody penetration. A decrease in staining in the primed state observed in all experiments might indicate lower cell or chromatin accessibility in these conditions, and thus the absolute intensity quantifications may not be meaningful.

However, even bearing that in mind, the observations made here do not agree with



**Figure 2.11:** Scanning confocal microscopy images of anti-H3K9me3 (A) or anti-H3K27me3 (B) immunofluorescence in wild-type mESCs in 2i/LIF or 24h and 48h after 2i withdrawal. Samples in A were prepared in parallel and imaged together. 24h and 48h samples in B were prepared and imaged together, while the 2i/LIF sample was prepared and imaged separately. Left, maximum projections of 10  $\mu\text{m}$ -thick z-stacks (A) or single slices through the middle of the cells (B), colour denotes the true unadjusted intensity. Middle, close-ups of the boxed regions, with intensity adjusted to showcase the histone mark patterns. Right, DNA staining with Hoechst for the same regions, intensity adjusted. In A, the cyan arrows indicate small puncta of H3K9me3, reminiscent of those seen for HP1 $\beta$  in Figure 2.9; the red narrow arrows and the yellow triangular arrows point to larger foci of H3K9me3 that show little enrichment for DNA, or vice versa. Scale bars, 20  $\mu\text{m}$  or 5  $\mu\text{m}$  on the close-ups.

the published hypothesis (Tosolini et al. 2018). There is no clear H3K27me3 staining on chromocentres in naive cells, and H3K9me3 enrichment does not appear to increase upon differentiation. On the contrary, in the formative state, H3K9me3 foci appear to become less pronounced, mirroring a similar trend in HP1 $\beta$  staining (Figure 2.8). This contradiction is likely the result of the differences in the differentiation protocols: EpiSCs are “more primed”, compared to cells grown in N2B27 for 48h, while mESCs in serum/LIF are considered to be a mixture of cells in different pluripotent substates rather than being in the formative state (see section 1.3.1). Furthermore, the different results obtained for cells grown in 2i/LIF might also be due to the different mESC lines intrinsically corresponding to slightly different states on the pluripotency spectrum (Mattout et al. 2015). In addition, the anti-H3K27me3 antibody used in this study was different from that used by Tosolini et al. (2018), which may also have contributed to the inconsistencies.

## 2.5 Discussion

In this chapter, the critical control experiments were conducted to confirm that genetic modification of the *Cbx1* locus to create a HaloTag fusion did not disrupt the function of the HP1 $\beta$  protein. The line was then used to characterise HP1 $\beta$  during exit of mESCs from naive pluripotency. *Cbx1* RNA levels were found to increase transiently 24h after the inhibitor removal, resulting in higher HP1 $\beta$  protein expression in both formative and primed states. Furthermore, HP1 $\beta$  concentration increased preferentially in euchromatin at first, resulting in a lower contrast between eu- and heterochromatin.

In parallel, heterochromatic domains in mESCs in different pluripotent states were also characterised using Hoechst staining. The nuclei of cells in the formative state were found to have larger area. Furthermore, the DNA density appeared to decrease at this stage. In addition, small puncta of HP1 $\beta$  were observed in euchromatin, outside of DNA-dense regions. On the other hand, there appeared to be a dichotomy between cells which did or did not show HP1 $\beta$  enrichment within chromocentres. The correspondence between DNA and HP1 $\beta$  foci increased significantly in primed-state cells.

Other characteristic heterochromatic markers were studied using immunofluorescence. The redistribution of HP1 $\beta$  in the formative state was not mirrored by HP1 $\alpha$ . However, the H3K9me3 modification – the classic binding site for HP1 proteins on chromatin – was reduced within heterochromatic foci at that stage, in agreement with the decrease in its ChIP-seq signal at pericentromeric repeats observed previously. The spatial pattern of H3K27me3 also changed upon differentiation, although it was not clear whether this was linked with changes in HP1 and H3K9me3.

The elevated levels of HP1 $\beta$  upon differentiation, alongside the effect of the *Cbx1*, *3* double knockout on genome reorganisation, prompted me to investigate the differentiation progression of mESCs lacking one or several of the HP1 proteins. However, no striking misregulation of gene activation or suppression during exit from naive pluripotency could be observed even in the *Cbx1*, *3*, *5* triple knockout, despite a notable decrease in cell fitness. Some knockout cell lines showed a mild delay in transitioning to the formative state, which could be studied further by analysing expression of a larger panel of genes.

### **2.5.1 Changes in the mechanical properties of the nuclei might be linked to chromatin structure**

It is well-established that differentiation-induced changes in cellular morphology are often linked to changes in their mechanical properties. In particular, mESCs in 2i/LIF are compact and near-spherical, while release from the naive state leads to flattening and emergence of protrusions, somewhat akin to the epithelial-to-mesenchymal transition. This process has been found to be crucial for differentiation (Chowdhury et al. 2010; Belly et al. 2021; Bergert et al. 2021).

Although the cell membrane and cytoskeleton have so far been identified as the components playing functional roles in this process, the mechanical properties of the nucleus also change at the onset of mESC differentiation. In serum/LIF, mESCs displaying low levels of the naive transcription factor NANOG were found to have softer nuclei compared to the more naive cells with high levels of NANOG (Chalut et al. 2012).

Moreover, nuclei of mESCs 24h after removal of 2i/LIF display auxeticity – a special mechanical property, whereby upon stretching an object expands in the direction perpendicular to the applied force – also linked to increased softness (Pagliara et al. 2014). Interestingly, this transient softening is at odds with the general trend for stiffening that occurs during mESC differentiation (Pajerowski et al. 2007; Pillarisetti et al. 2011).

The flat appearance of the genome structures of formative-state cells (Figure 1.4C), as well as the increased area of their nuclei observed in this work (Figure 2.6C) could therefore be caused by the softer nuclei being more easily deformed by gravity. Interestingly, this phenotype is reverted 48h after removal of 2i, even though cells preserve a fibroblast-like appearance at this stage. Thus, the stiffness of the nucleus and the cell’s outer membrane tension may not always change in unison.

The changes in mechanical features were proposed to be caused by the lower degree of chromatin condensation in these states, judged either by the more diffuse localisation of HP1 $\alpha$  and H3K9me3 (Chalut et al. 2012) or by direct estimation of chromatin compaction using electron microscopy (Pagliara et al. 2014). On the other hand, crosslinking by HP1 $\alpha$  has recently been demonstrated to increase chromatin stiffness (Strom et al. 2021). Thus, the drop in H3K9me3 levels (Figure 1.4D) and the relocalisation of HP1 $\beta$  (Figure 2.8) might contribute to the nuclear softening in the formative-state mESCs. Indeed, HP1 $\beta$  in euchromatin, where the density of modified H3 tails is very low, will have a lower capacity for crosslinking even when it is bound to chromatin.

Consistently, the mean DNA density of nuclei decreased transiently in the formative state, and so did the percentage of the area occupied by heterochromatic foci and the density of these foci (Figure 2.6). However, decondensation within chromocentres appeared to be somewhat milder than in euchromatin, potentially due to the presence of HP1 $\alpha$ .

## 2.5.2 Reorganisation of chromocentres

### HP1 $\beta$ is recruited to chromocentres upon differentiation

The confocal images of HP1 $\beta$  and DNA revealed two interesting features of HP1 $\beta$  localisation during exit from naive pluripotency (Table 2.1).

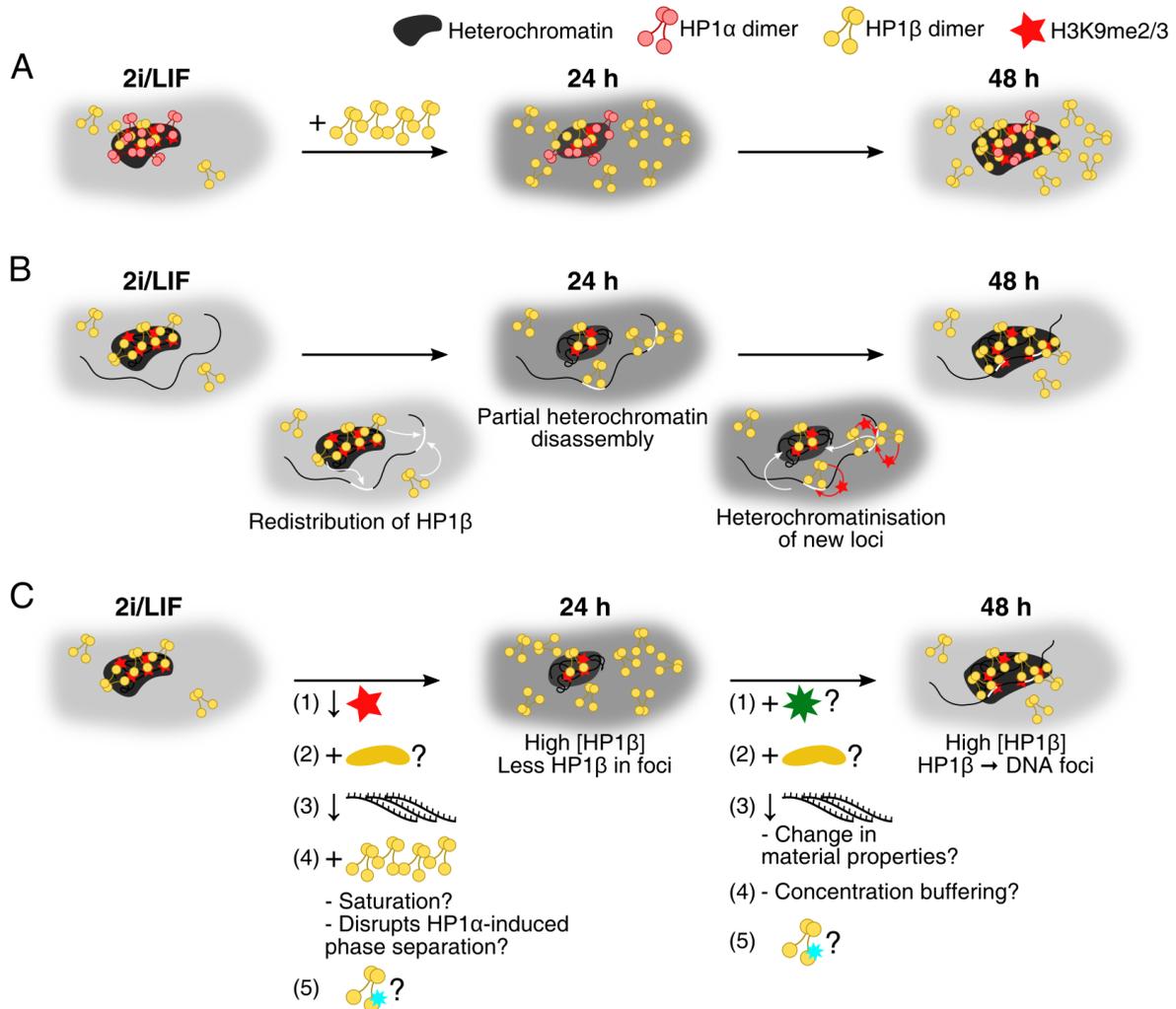
1. Although HP1 $\beta$ -dense domains generally colocalise with DNA foci, HP1 $\beta$  forms small puncta in euchromatin, while in some cells chromocentres lack HP1 $\beta$ . The correspondence between HP1 $\beta$  and DNA density increases in the primed state.
2. While relative DNA density within chromocentres changes only slightly throughout differentiation, HP1 $\beta$  “escapes” the foci in the formative state. In other words, although the levels of HP1 $\beta$  rise globally, it becomes less enriched in heterochromatin.

**Table 2.1:** Enrichment in DNA or HP1 $\beta$  staining intensity in foci of DNA or HP1 $\beta$  compared to outside of foci, and the Pearson’s correlation coefficient between those images. The indicated values are mean $\pm$ SD of individual cells’ measurements.

	Enrichm. in DNA foci		Enrichm. in HP1 $\beta$ foci		Correlation DNA & HP1 $\beta$
	DNA	HP1 $\beta$	DNA	HP1 $\beta$	
<b>2i/LIF</b>	2.33 $\pm$ 0.29	1.54 $\pm$ 0.32	1.68 $\pm$ 0.54	2.26 $\pm$ 0.37	0.49 $\pm$ 0.19
<b>24h</b>	2.39 $\pm$ 0.23	1.38 $\pm$ 0.19	1.63 $\pm$ 0.44	1.81 $\pm$ 0.30	0.42 $\pm$ 0.16
<b>48h</b>	2.25 $\pm$ 0.25	1.83 $\pm$ 0.31	1.95 $\pm$ 0.26	1.96 $\pm$ 0.28	0.68 $\pm$ 0.16

These observations suggest that in the formative state more HP1 $\beta$  is produced, that is then distributed into constitutive heterochromatin later in the primed state. As a result, the balance of HP1 isoforms at centromeres must shift, which could have an effect on their biophysical properties, especially if LLPS phenomena do play a role in heterochromatin segregation (Figure 2.12A). Recent *in vitro* findings suggest that an increase in HP1 $\beta$  might decrease the stability of the HP1 $\alpha$ -mediated droplets (Keenen et al. 2021), in contrast to the general idea of gradual gelatinisation of heterochromatin upon differentiation. However, effects in cells might be different, especially given that HP1 $\beta$  also promotes chromatin LLPS (Wang et al. 2019).

Another potential role for HP1 $\beta$  in this context could be bringing of new genomic loci into heterochromatin to be permanently silenced (Figure 2.12B). HP1 $\beta$  might first



**Figure 2.12:** A. The increase in HP1 $\beta$  levels and its redistribution might change the balance between the HP1 isoforms within heterochromatic foci, potentially affecting their properties, e.g. the level of compaction or stiffness. B. HP1 $\beta$  might help “shuttle” the genes that must be silenced during differentiation (white) into heterochromatin. C. HP1 $\beta$  is distributed differently in naive, formative and primed mESCs: 24h after the onset of differentiation, the level of HP1 $\beta$  increases, but its enrichment in foci becomes weaker; at 48h timepoint, HP1 $\beta$  foci become more prominent and colocalise better with DNA-dense regions. These changes could be mediated by the following non-exclusive mechanisms: decrease in H3K9me2/3 and subsequent increase of an unknown histone modification (1); through interactions with protein partners (2); due to potential decrease in major satellite repeat transcripts, and the resultant change in the material properties of chromocentres (3); due to the increase in HP1 $\beta$  concentration, which could saturate the binding sites in heterochromatin, lead to concentration buffering and thus increase in the area of heterochromatic foci, or disrupt HP1 $\alpha$ -mediated LLPS (4); or via PTM on HP1 $\beta$ .

be attracted to targets in euchromatin or facultative heterochromatin, either due to them being marked with H3K9me2/3, via KAP1, or in some other way. By engaging histone methyltransferases, it could then create a local repressed environment – perhaps manifested as the small non-chromocentre HP1 $\beta$  puncta visible in the images. Either spontaneously over time, or stimulated by a trigger, these puncta could then fuse with larger heterochromatic domains. Since many genes might need to be turned off at the exit from naive pluripotency to create a block to cell fate reprogramming, a rise in HP1 $\beta$  levels might facilitate this process.

However, since HP1 $\beta$  does not appear to be essential for exit from naive pluripotency, its role might be auxiliary or redundant. Alternatively, the observed process might be important for later stages of differentiation. Finally, although cells do exit from naive pluripotency in the absence of HP1 $\beta$ , their resultant identity might be different from that of the wild-type cells due to gene misregulation.

### **Potential mechanisms for HP1 $\beta$ relocalisation**

Another point of interest is the mechanism of HP1 $\beta$  redistribution into euchromatin in the formative state, and into DNA-dense foci in the primed state. Given the fast dynamics of HP1 proteins in cells (Cheutin et al. 2003; Festenstein et al. 2003), it is surprising that although the levels of HP1 $\beta$  are already high 24h after 2i withdrawal, it is not distributed into foci until later. Therefore, two events potentially occur in succession: first, HP1 $\beta$  is released from the existing foci, and then recruited into chromocentres. Potentially, the former event could be a necessary step to enable the latter. Alternatively, the time lag between the two triggers might be needed to allow HP1 $\beta$  find its gene targets to bring back with it into heterochromatin. Several possible models are suggested below and in Figure 2.12C, and the potential ways to address them are discussed in the next section.

Firstly, release of HP1 $\beta$  from and its recruitment to heterochromatic foci could be mediated by histone modifications. The decrease in H3K9me3 levels on major satellite repeats in the formative state might contribute to the lower enrichment of HP1 $\beta$  in chromocentres. In the primed state, HP1 $\beta$  could be recruited into

pericentromeric heterochromatin due to the removal of H3S10pho that prevents the chromodomain-H3K9me2/3 interaction (Fischle et al. 2005), or via other marks, such as H3K23me1/2/3, H1.4K26me2/3, or H4K20me3. The latter, identified so far as an HP1 $\beta$ -specific interaction partner (Bosch-Presegué et al. 2017), could explain the differences in the distribution of the HP1 $\alpha$  and HP1 $\beta$  isoforms. Alternatively, these could be caused by the differential specificity of the HP1 isoforms towards H3K9me2/3-modified chromatin (Hiragami-Hamada et al. 2016).

Furthermore, HP1 $\beta$ -interacting proteins could play a role in its targeting during differentiation. For instance, both KAP1 and CAF1 (see sections 1.4.1 and 1.4.4) are mildly upregulated in formative-state mESCs (Kalkan et al. 2017). DNMTs or other proteins involved in DNA methylation could also affect HP1 $\beta$  distribution due to the numerous links between the DNA methylation and H3K9me2/3-HP1 pathways (see section 1.2.1).

Another intriguing possibility is the involvement of RNA. Recently, the major and minor satellite repeat RNA was demonstrated to be necessary for HP1 $\beta$  localisation to heterochromatin in mESCs (Quinodoz et al. 2021). Furthermore, high levels of satellite repeat transcription have been shown to be facilitated in mESCs by NANOG (Novo et al. 2016), and thus these transcripts could be rapidly downregulated upon exit from naive pluripotency. Finally, the major satellite transcripts promote HP1 $\alpha$  LLPS and the liquid-like dynamic behaviour of chromocentres in mESCs, and their depletion leads to partitioning of heterochromatin into smaller, denser and less mobile foci (Novo et al. 2020). Thus, a potential decrease in the levels of major satellite RNA in the formative state could lead to HP1 $\beta$  relocalisation into euchromatin. Better correlation between the DNA and HP1 $\beta$  distribution in primed-state mESCs could indicate that the mechanism for HP1 $\beta$  recruitment into heterochromatin changed, and/or reflect the altered material state of chromocentres in the absence of the major satellite RNA.

Alternatively, higher amounts of HP1 $\beta$  may not be partitioned into heterochromatin simply due to saturation. Indeed, the area of the foci increases at the formative-primed transition more noticeably than their HP1 $\beta$  density, which may signal that additional binding sites are created for HP1 $\beta$  at that stage. Looking at this differently, this is

somewhat reminiscent of the concentration buffering feature of LLPS, whereby upon an increase in the total solute concentration, the volume of the condensate changes rather than the solute concentration within it. Due to the likely complex multi-phase nature of heterochromatin, rising HP1 $\beta$  concentrations may first lead to an increase in density of the foci until a certain threshold, and then the system may transition to concentration buffering behaviour.

Also in the context of LLPS, the increase in HP1 $\beta$  levels might itself promote release of HP1 $\beta$  from foci, if these are primarily formed through HP1 $\alpha$ -DNA interactions. However, in that case it is unclear why the HP1 $\alpha$  localisation appears to be unperturbed in the formative state, and how HP1 $\beta$  is redistributed in the primed state.

Finally, the targeting of HP1 $\beta$  may be influenced by its PTMs.

## **Links between global reorganisation of chromatin and HP1**

The rationale for this work was based on the global reorganisation of chromatin, observed 24h after the onset of mESC differentiation. Although multiple interesting features regarding HP1 during the exit from naive pluripotency were discovered, the link to the global increase in intermingling, DNA methylation and the transient drop in H3K27me3 is still unclear. The proposed shuttling of (developmental) genes that need to be silenced into heterochromatin by HP1 $\beta$  could cause more long-range interactions in the formative state. Alternatively, redistribution of HP1 $\beta$  into euchromatin may result in lower levels of crosslinking and stiffness, also leading to higher intermingling. This would be consistent with the altered Hi-C contact length profile of *Cbx1*,  $\beta$  double knockout mESCs. Finally, the formative state is characterised by a global increase in transcription (Laue lab, unpublished), and thus the presence of higher levels of RNA could affect HP1 $\beta$  localisation. Investigating HP1 $\beta$  distribution and dynamics in *Tet1* or *Dnmt3a/b* knockout cell lines may provide a clue for whether changes in intermingling and HP1 $\beta$  levels and localisation might be linked.

### 2.5.3 Future work

Although several exciting observations were made in this project, critical questions about their role and mechanism remain unanswered.

#### Characterisation of chromocentres at high resolution

Multiple lines of evidence point to the relationship between chromocentres/DNA-dense regions, HP1 proteins, H3K9me2/3 methylation and the H3K27me3/polycomb changing at the exit of mESCs from naive pluripotency. It is therefore necessary to characterise their relative localisation in the same cells, preferably using staining methods suitable for accurate quantification. To this end, an additional control antibody could be used to assess cell permeability in the IF experiments and account for it in the quantification. In addition, labelling with smaller-sized specific probes or nanobodies could improve penetration; alternatively, HP1 $\alpha$  could be endogenously tagged to enable labelling with small-molecule fluorophores. Furthermore, it would be beneficial to image these regions at high resolution, e.g. using SMLM or STED microscopy, especially because HP1 foci have been shown to consist of many small domains (Erdel et al. 2020). Finally, it might be useful to study the link between HP1 $\beta$  and DNA distribution and the cell cycle, e.g. using PCNA as a marker of the cell cycle stage.

Such measurements could help us compare the “areas of influence” of HP1 $\alpha$  and HP1 $\beta$ : do they overlap closely at all timepoints, forming mixed condensates, or do they form separate territories within the chromocentres? Does either of them overlap better with the DNA-dense regions, or with H3K9me2/3 foci? Does this change upon exit from naive pluripotency? In addition, these experiments could indicate whether the small HP1 speckles are decorated with H3K9me2/3, H3K27me3, or other histone marks, potentially helping to elucidate their nature.

#### Biophysical properties of HP1 foci

Since nuclei in the formative state display auxeticity (Pagliara et al. 2014), and the materials properties of the chromatin in general and heterochromatic foci in particular are

known to change in other differentiation contexts (Chalut et al. 2012; Novo et al. 2016; Novo et al. 2020), it is likely that the formative state is linked with physical changes in the nuclear environment. Thus, understanding the nature of heterochromatin and principles of its segregation is crucial to understanding the processes occurring upon exit from pluripotency. To this end, properties of HP1 $\beta$  dynamic behaviour were studied, and the findings are described in the next chapter.

Furthermore, studying the phenotype of cells devoid of HP1 $\beta$  or having higher or lower levels of it might provide further clues to the mechanism and function of its redistribution. For example, if increasing or dampening the levels of HP1 $\beta$  triggers or prevents its release from chromocentres, it would argue that altering the balance between HP1 isoforms is sufficient to cause their redistribution.

### Searching for triggers

In addition to the change in HP1 $\beta$  levels, other factors likely contribute to its redistribution. Given that H3K9me2/3 levels on major satellite repeats decrease during exit from naive pluripotency, H3S10pho, H3K23me1/2/3, H4K20me3 and H1.4K26me2/3 might be involved. ChIP-seq or similar experiments for these histone modifications during exit from naive pluripotency could help understand the observed phenomena. A different explanation for a change in HP1 $\beta$  patterns could be a new interaction partner, or a post-translational modification of the protein itself, which could be studied by proteomic approaches. Finally, the contribution of RNA could be studied by measuring the levels of the major satellite repeat transcripts during the differentiation timecourse, as well as investigating the consequences of RNA depletion (e.g. using LNA) on HP1 $\beta$  localisation and genome architecture.

Another interesting question is what causes the increase in *Cbx1* gene transcription. Examining the H3K27me3 and H3K27ac ChIP-seq data in the vicinity of this gene did not provide any clues to this, but investigating other histone marks and epigenetic features (e.g. chromatin accessibility) might be informative. A motif analysis of the accessible chromatin near the *Cbx1* promoter and enhancer could potentially reveal a binding site for a transcriptional activator, whose levels rise in the formative state. Al-

ternatively, a locus-specific chromatin proteome investigation by CasID (Schmidtman et al. 2016) could be used to identify such factors.

### **HP1 $\beta$ binding in euchromatin and the shuttle hypothesis**

To test the model for HP1 $\beta$  shuttling certain loci from euchromatin into heterochromatin, the identity of the genomic sequences associated with the HP1 $\beta$  speckles must be determined. The difficulty lies in distinguishing the speckles from the rest of HP1 $\beta$  prior to sequencing, i.e. in retaining the spatial information. A combination of genome reconstruction with snHi-C (Stevens et al. 2017) and imaging could allow one to do that, and the development of this method is described in chapter 4.

An alternative way to address this question could be to characterise the timing and dynamics of emergence of the small HP1 $\beta$  speckles, as well as their subsequent fate. This could be done by long-term time-lapse imaging of HP1 $\beta$  in mESCs and during differentiation. For example, the small puncta could emerge, but not fuse with the large HP1 $\beta$  domains in naive or formative mESCs, while an increase in the fusion might be observed in the primed state. It would also be interesting to see if the puncta emerge more often in the formative state. Furthermore, a more continuous and dynamic picture of HP1 $\beta$  redistribution could be obtained in this way rather than with fixed-cell imaging, also potentially revealing the link between these processes and the cell cycle.

# Chapter 3

## Studies of HP1 $\beta$ dynamics with single-molecule imaging

*The sptPALM approach was pioneered in the lab by Dr Srinjan Basu, using the microscope built by Dr Aleks Ponjavic. Some of the analysis scripts were written by Dr Wayne Boucher. The 4P analysis pipeline was developed by Dr Ofir Shukron and Prof David Holcman, Ecole Normale Supérieure, Paris.*

*The initial optimisation of HP1 $\beta$  3D sptPALM was performed together with Lucy Henderson, and some of the approaches to analyse the double-labelling experiments were developed by Krzysztof Herka.*

### 3.1 Introduction

The second aim of this work was to characterise the dynamic behaviour of HP1 $\beta$  in live mESCs, and link it with the features of HP1 interaction with chromatin observed *in vitro* and with the current models of heterochromatin organisation (see section 1.4). In addition, I wanted to observe whether there were any changes associated with chromatin reorganisation upon differentiation. This was achieved by employing single-particle tracking (SPT) – a type of imaging experiment, where the movement of single molecules in live cells is followed over time. It is a very information-rich technique, allowing one to investigate diffusion of molecules and understand changes in their en-

vironment, *in vivo* chromatin binding kinetics and mechanism of action. However, as is the case with many emerging techniques, it requires sophisticated data analysis based on theoretical biophysical frameworks, which are themselves currently under development.

### 3.1.1 Single-molecule features of phase separation

One of the key questions in understanding the organisation of heterochromatin is whether HP1 proteins induce LLPS to compact chromatin and/or to concentrate or deplete selected molecules (Larson and Narlikar 2018). The features unequivocally characteristic of phase separation are a matter of debate, since on one hand they might depend on the composition and material properties of the condensed phase (Larson and Narlikar 2018), while on the other hand phenomena reminiscent of LLPS can also be caused by other mechanisms (McSwiggen et al. 2019a). For HP1, the picture is even more complex, since multiple processes might take place simultaneously. Firstly, HP1 is able to induce LLPS *in vitro* through multiple mechanisms (see section 1.4.2). Secondly, a proportion of HP1 is bound to chromatin stably, and it has been proposed that a phase-separated droplet of mobile protein may form around it (Strom et al. 2017). Furthermore, stable HP1 binding might change the properties of the chromatin fiber, leading to chromatin LSPS or LLPS. Thus, in this complex system, it may not be easy to answer this simple question. Below, some potentially useful parameters that could be measured by SPT are reviewed.

#### Features of the dense phase

In a simple case of LLPS, both phases are expected to maintain a uniform concentration of each component. Super-resolution imaging of chromocentres has shown that this is not the case for HP1 (Erdel et al. 2020). However, in the presence of a long polymer like chromatin, and when a proportion of HP1 can bind to chromatin strongly, this may not necessarily indicate absence of LLPS (Larson and Narlikar 2018; Keenen et al. 2021). This also means that other liquid-like features of phase-separated droplets,

such as sphericity, might not apply in cells.

On the other hand, if LLPS takes place, the dense phase by definition must contain a higher concentration of protein. Although a cross-linked chromatin globule with an increased density of HP1 binding sites should also display this feature, the molecules in that case would be predominantly chromatin-bound. Hence, if heterochromatin is a phase-separated condensate of HP1, one would expect there to be a higher concentration of freely diffusing, non-chromatin-bound HP1 compared to euchromatin. This argument could be especially powerful, since neutral probes of a size similar to HP1 dimer are partially excluded from heterochromatin, either due to crowding or LLPS (Bancaud et al. 2009; Strom et al. 2017).

In addition, the physical environment within the condensate might be different from that outside. Naively, one might expect the protein under study to diffuse slower due to the presence of transient interactions, and the absence of a change in rotational diffusion coefficient for HP1 in heterochromatin was used to argue against the LLPS hypothesis (Erdel et al. 2020). However, this is not conclusive, since phase separation does not necessarily lead to a change in diffusion rates (Bo et al. 2021). Moreover, in a comparison between biophysical models for focus formation via LLPS or weak protein binding to a cross-linked chromatin globule, Heltberg et al. (2021) found that the latter predicted a proportional relationship between the apparent diffusion coefficients and the concentrations of molecules inside and outside the focus. In their study, a suspected LLPS system violates this relationship, with the protein displaying much faster – not slower – motion within the condensate than expected. Finally, since heterochromatin slows down diffusion of neutral probes (Bancaud et al. 2009), a lack of such an effect for HP1 might indicate that the effects of the environment are protein/particle-specific, potentially arguing in favour of LLPS.

Proteins in cells are expected to display anomalous diffusion, deviating from Brownian motion. Thus, the diffusion coefficient is not the only parameter that describes their dynamics. LLPS leads to a stronger confinement of proteins within the condensate compared to the a cross-linked globule model (Heltberg et al. 2021). This is manifested in the higher tendency of trajectories to fold back on themselves, i.e. display angles

$>90^\circ$  between their consecutive steps. However, the effect is quantitative rather than step-wise, and accurate parameter measurement is required to draw conclusions.

### **Boundary effects**

A phenomenon distinctive of LLPS occurs at the boundary between the condensate and the environment. Due to an increased number of interactions within the dense phase, leaving it is energetically unfavourable. As a result, molecules display a collective drift away from the boundary, which can be measured (Bo et al. 2021; Heltberg et al. 2021), and trajectories near the phase boundary turn back more often than those outside or deeper inside the focus (Collombet et al. 2021). HP1 was demonstrated to display lower variance in the direction of diffusion near the edges of heterochromatic foci, which might be attributed to the drift effect (Strom et al. 2017).

The presence of such a semi-permeable barrier could also decrease the diffusion rate of molecules compared to the inner regions of the liquid droplet. Indeed, this was observed for HP1 and for a YFP probe at the edge of HP1 foci (Strom et al. 2017). However, theoretical work indicates that this is not a mandatory characteristic of phase separation, and when present, it signifies a difference in the diffusion environments within the two phases (Bo et al. 2021).

As a consequence of the energetic barrier, molecules within the dense phase are expected to display preferential internal mixing – in other words, a molecule is more likely to remain within a focus than to leave. FRAP experiments indicated that this may not be the case for HP1 (Erdel et al. 2020). However, since liquid droplets are dynamic and readily exchange contents with the environment (Keenen et al. 2021), this effect may only operate on very short timescales that cannot be studied by FRAP.

Finally, many processes in cells are not at equilibrium – for example, a non-equilibrium model of LLPS during transcription was proposed recently (Henninger et al. 2021). Monitoring the behaviour of molecules at the interphase boundary may indicate whether the system in question is at equilibrium, since equal numbers of particles are expected to cross the boundary in both directions, and the distributions of their diffusion rates must also be the same (Bo et al. 2021).

### 3.1.2 Experimental design

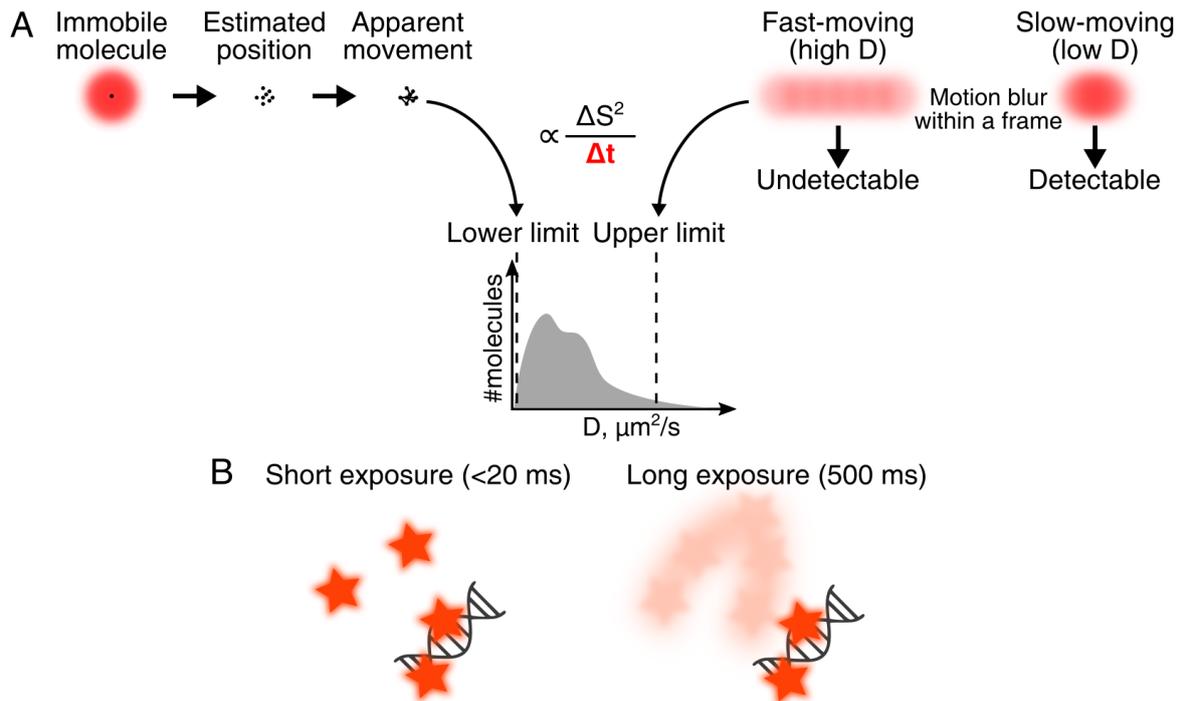
#### SptPALM

For SPT, proteins under study must be specifically labelled with a fluorophore. Furthermore, during imaging, individual emitters must be sufficiently far apart to enable one to follow their motion. To achieve this for high-copy-number proteins, photoactivatable fluorophores can be employed in an experiment known as single particle tracking photoactivatable localisation microscopy (sptPALM) (Manley et al. 2010). Illuminating the sample with UV light results in activation of a subset of photoactivatable fluorophores (see section 1.5.1), and the density of localisations can be optimised by varying the laser intensity. In this work, organic photoactivatable Janelia Fluor (PA-JF) dyes, conjugated to HaloTag ligand, were employed to achieve bright and long-lasting signal (Grimm et al. 2016).

#### The question of exposure

The range of diffusion speeds that can be detected by SPT is determined by the precision limit of the microscope and by motion blur (Figure 3.1A). In mammalian cell nuclei, an inert protein of  $\sim 30$  kDa diffuses with a coefficient of  $\sim 5\text{-}20 \mu\text{m}^2/\text{s}$  (Lippincott-Schwartz et al. 2001; Xiang et al. 2020). In 20 ms, such a molecule will move  $\sim 0.6\text{-}1.2 \mu\text{m}$  in 2D. However, these values drop significantly for larger, stickier or more positively-charged proteins (Xiang et al. 2020). Thus, for many chromatin-associated proteins, live single-molecule imaging with an exposure of  $< 20$  ms allows one to observe almost the full spectrum of motion regimes, including free diffusion, diffusion constrained by obstacles or by tethering to chromatin fiber (Etheridge et al. 2014; Carr et al. 2017; Basu et al. 2020) (Figure 3.1B). At shorter exposure times, a higher fraction of the rapidly-moving particles can be detected. However, this comes at a cost of artificially inflating the diffusion coefficients of the slow-moving molecules due to localisation imprecision (Figure 3.1A).

On the other hand, imaging with a long exposure time results in motion blurring of the freely-diffusing molecules (Etheridge et al. 2014; Basu et al. 2020). Indeed, for



**Figure 3.1:** A. Both the upper and the lower limits of detection are determined by the exposure time. The upper limit is the result of motion blurring of the fast-moving particles, and it is proportional to the squared maximum displacement in a frame that is still detectable ( $\Delta S^2$ ) (right). An immobile molecule will display apparent movement due to the imprecision of particle localisation, giving rise to the lower limit of detection, proportional to the squared precision of the microscope ( $\Delta S^2$ ) (left). The limits are inversely proportional to the exposure time ( $\Delta t$ ). In reality, they are probabilistic effects, but they are illustrated as hard boundaries for simplicity. B. The exposure time determines the populations of molecules that can be observed in an experiment. Short exposure times allow observation of both the chromatin-bound and unbound proteins in different environments, but it is not possible to characterise slow diffusion accurately. Long exposure times allow one to specifically observe the chromatin-bound population, determine its diffusion parameters and residence time.

example, if the maximum displacement within a frame that still allows the localisation to be detected is 250 nm, then with a 500 ms exposure, only particles diffusing slower than  $0.03 \mu\text{m}^2/\text{s}$  are detected (Figure 3.1B). Since the diffusion coefficient of chromatin has been estimated as  $10^{-3}$ - $10^{-4} \mu\text{m}^2/\text{s}$  (Akhtar and Gasser 2007), these will probably correspond to the chromatin-bound fraction. Overall, the observed value of the diffusion coefficient depends on the exposure time, and only “apparent” diffusion coefficients measured under identical imaging conditions can be compared (Basu et al. 2020).

In this study, two strategies were employed: 1) imaging with a short exposure of 5 ms to capture the dynamics of the fast-moving fraction of HP1 $\beta$  and measure the relative amounts of confined and unconfined molecules, and 2) imaging with a long 500 ms exposure to investigate the diffusion of chromatin-bound HP1 $\beta$  (or, in other words, HP1 $\beta$ -bound chromatin loci) and analyse the residence time of HP1 $\beta$  on chromatin (Gebhardt et al. 2013; Basu et al. 2020).

## 3.2 Short-exposure sptPALM

The goal of this set of experiments was to describe the diffusive behaviour of HP1 $\beta$  and to investigate whether it changes upon exit of mESCs from naive pluripotency in the context of chromatin reorganisation. Initially, 3D sptPALM was employed to prevent loss of molecules diffusing out of the focal plane. However, since we wanted to investigate HP1 $\beta$  behaviour separately in eu- and heterochromatin and at the boundary between them, a two-colour imaging method was developed to distinguish between the compartments. At present we can only implement this using 2D sptPALM.

A recent SPT data analysis pipeline for chromatin-binding proteins (Basu et al. 2020) was employed to classify the confined and unconfined fractions of HP1 $\beta$ , as well as to explore its dynamic features beyond the diffusion coefficient. The suitability of this novel approach for this data was assessed, and the algorithm was developed further to capture a larger proportion of experimental tracks.

Upon exit from naive pluripotency, the balance between the HP1 isoforms is altered, and the localisation of HP1 $\beta$  in the nucleus changes (see chapter 2). Using the single-

molecule approach, I set out to answer the following questions.

1. Does HP1 $\beta$  diffuse differently in eu- and heterochromatin in mESCs? How can this be interpreted?
2. Does HP1 $\beta$  dynamics display features of LLPS?
3. Does HP1 $\beta$  diffusion change upon exit from naive pluripotency, and can this be linked to the changes in genome organisation?

### 3.2.1 Method optimisation

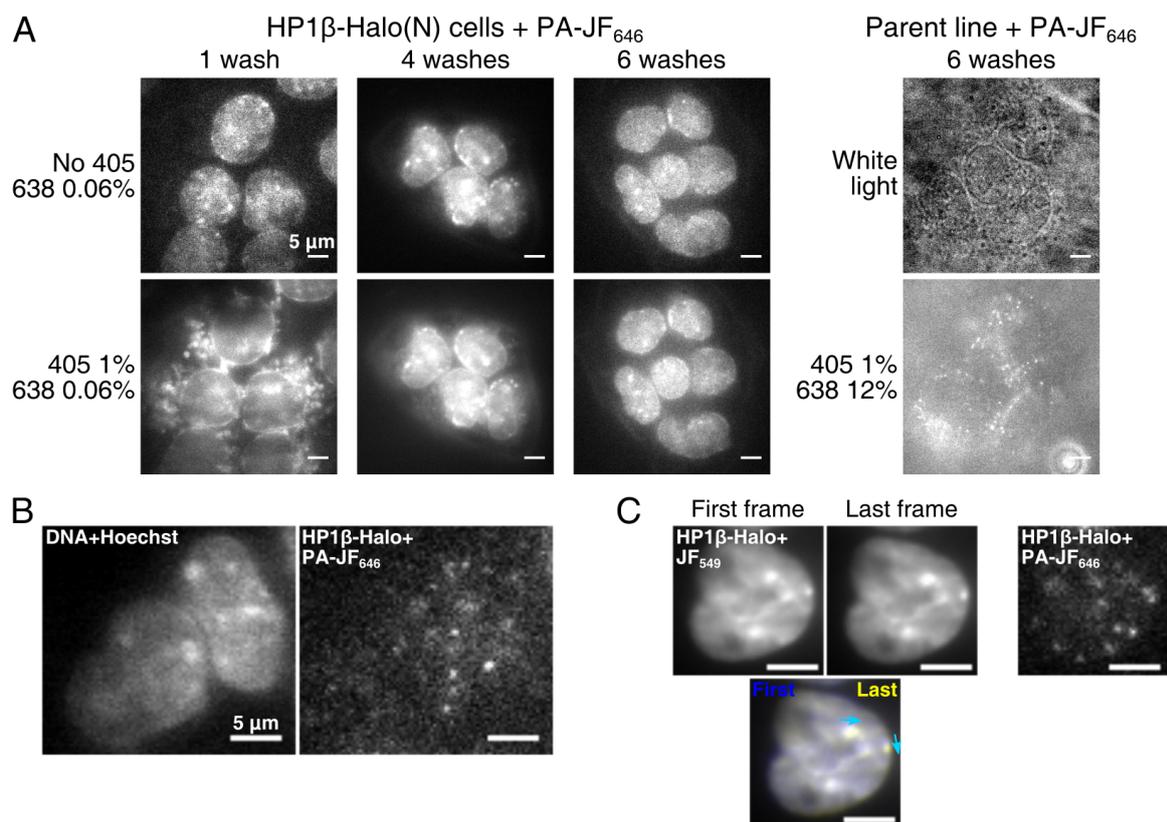
SptPALM was employed previously in the lab by Dr Srinjan Basu to explore the dynamics of the NuRD complex (Basu et al. 2020), but the method required optimisation to apply it to HP1 $\beta$  in mESCs exiting from the ground state.

#### Sample preparation

HP1 $\beta$ -Halo(N) and -Halo(C) mESCs, where the endogenous *Cbx1* gene was genetically engineered to have the HaloTag at either the N- or the C-terminus of the HP1 $\beta$  protein, respectively, were used in these experiments. For two-colour imaging, mostly presented in this chapter, HP1 $\beta$ -Halo(N) mESCs were FACS-sorted to highly enrich in cells expressing the correct levels of HP1 $\beta$ .

The basic protocol for labelling a HaloTagged protein in mESCs in serum/LIF (Grimm et al. 2016) was, unexpectedly, unsuitable for cells in 2i/LIF or N2B27 media. The PA-JF dyes tended to label the membranes non-specifically or stay trapped in the cytoplasm. This issue was resolved by increasing the number of washes (Figure 3.2A). The negative control, using cells not containing the HaloTag fusion, showed that the modified protocol resulted in very low non-specific labelling (Figure 3.2A).

To distinguish eu- and heterochromatic regions, cells were double-labelled with a second fluorophore targeting DNA (Hoechst33342) or HP1 $\beta$  (Halo-JF<sub>549</sub>) for wide-field imaging, while PA-JF<sub>646</sub> was used for simultaneous sptPALM (Figure 3.2B,C). Because



**Figure 3.2:** A. Protocol optimisation for labelling mESCs grown in 2i/LIF with PA-JF<sub>646</sub>. Cells were labelled as described in the text and washed with media once for 1h (left) or 4-6 times for 20 min each. The samples were then imaged, using the 640 nm excitation laser at the indicated intensity, with or without the 405 nm activation laser. The negative control (right) was imaged at full 640 nm laser intensity. Similar results were obtained for cells in N2B27 and for PA-JF<sub>549</sub>. The intensities of the images were adjusted. B, C. Examples of mESCs in 2i/LIF, labelled for two-colour imaging. Diffraction-limited images of DNA labelled with Hoechst (average projection of 200 frames; B, left) or HP1β labelled with JF<sub>549</sub> (average projection of 500 frames; C, left), and single localisations of HP1β labelled with PA-JF<sub>646</sub> for sptPALM (single frame; right). In C, HP1β images at the beginning and at the end of acquisition and their overlay are shown to illustrate movement of the nucleus and of the heterochromatic domains (cyan arrows). Scale bars, 5 μm.

some DNA-dense areas may not be decorated by HP1 $\beta$ , and vice versa (section 2.4.2), results obtained using both methods were compared.

### Imaging optimisation

First, laser intensities were optimised to obtain good localisation density and enable tracking individual molecules for as long as possible. Next, preliminary data for HP1 $\beta$  was acquired at 50 or 100 frames per second (fps), which indicated that the higher frame rate revealed a larger population of unconfined molecules ( $\sim 35\%$  vs  $\sim 20\%$ ), likely due to less motion blur. Thus, the 3D sptPALM imaging was performed at 100 fps, and 2D sptPALM at 200 fps, i.e. with 5 ms exposure.

Since it was not possible to collect enough data from individual cells, all observations generated from the same dish were pooled, and different dishes were treated as biological replicates. 20-25 positions with 1-6 cells in them were imaged per dish, producing several thousand long high-quality trajectories (Table 3.1).

For 2D SPT experiments, wide-field snapshots could not be collected beforehand, since even within the one-minute acquisition cycle of 5 ms sptPALM data, chromatin within nuclei and nuclei themselves moved (Figure 3.2C). Thus, simultaneous imaging in two channels was performed by using a dichroic mirror and a second camera.

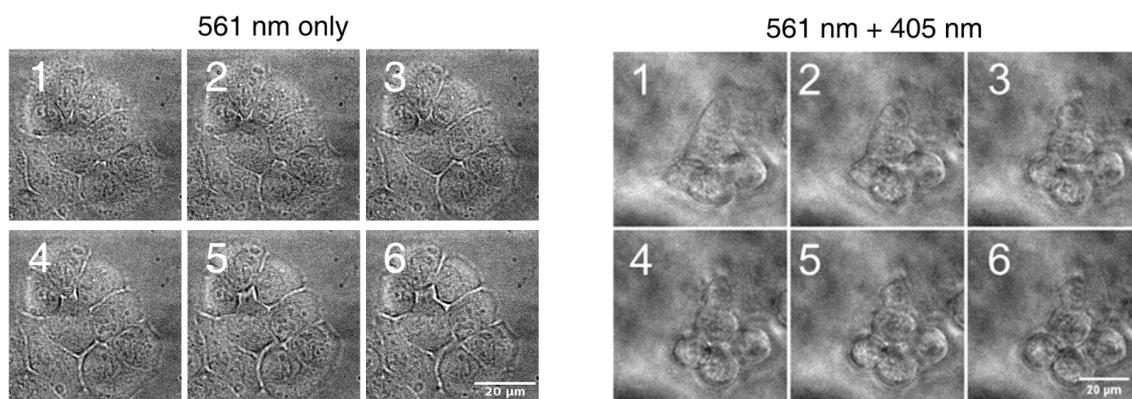
**Table 3.1:** Number of localisations and trajectories in each 2D sptPALM dataset: all localisations, all trajectories  $>2$  positions, trajectories  $>10$  positions and trajectories  $>10$  positions after quality filtering.

Diff state	Co-labelled	All loc	All traj	Traj $>10$	Traj $>10$ , filt
2i/LIF	DNA	2410667	186952	47999	18900
2i/LIF	HP $\beta$	838974	69574	17736	5986
24h	DNA	2602959	162861	53732	23859
24h	HP $\beta$	2207346	144677	49148	22309
48h	DNA	2372767	171036	47496	17822
48h	HP $\beta$	1619937	117047	37940	14073

### Phototoxicity

An important consideration for any live-cell imaging is phototoxicity. After 10 min of imaging with the laser intensities required for sptPALM, mESCs displayed morpholog-

ical characteristics of apoptosis. When imaging with the 561 nm laser only, no signs of cell death were apparent (Figure 3.3, left), while irradiation with both the 405 nm and 561 nm lasers caused cells to round up and detach from the dish after about 20,000 frames (Figure 3.3, right). This suggested that the cell damage was caused by the UV, which was therefore kept at the minimum possible power. Furthermore, the period over which one field of view was imaged was restricted to 10,000 frames (<1 min). Although the UV exposure probably exerts other more subtle effects, they are less likely to manifest themselves at such a short time scale, and the measures taken should minimise them. In addition, the medium used for imaging did not contain phenol red to suppress background and decrease phototoxicity.



**Figure 3.3:** Montages of bright-field images of cells illuminated with 561 nm laser only (left) or both the 561 nm and 405 nm lasers (right) in the sptPALM regime. The first frame of each montage shows the appearance of cells before imaging, and each subsequent frame was taken after a further 100 s (10,000 frames) of irradiation. Characteristic signs of phototoxicity are visible in images 3-6 on the right, including cellular rounding and detachment from the culture dish. Figure by L. Henderson.

### Reproducibility and biological replicates

Data was initially collected both for the HP1 $\beta$ -Halo(C) and HP1 $\beta$ -Halo(N) mESC lines to ensure that HP1 $\beta$  behaviour was not affected by the HaloTag. The preliminary results from the two lines disagreed. Since the HP1 $\beta$ -Halo(C) cells differentiated slower than either HP1 $\beta$ -Halo(N) or the parent line (section 2.2.3), I decided to concentrate on experiments in HP1 $\beta$ -Halo(N) mESCs.

Three good-quality biological replicates for each differentiation stage were then recorded with the 3D sptPALM setup. Unfortunately, the day-to-day variation in mi-

croscopy setup turned out to significantly affect the data. Ideally, all measurements for a particular comparison should have been collected on the same day without changing the layout of the microscope, but this consideration was not taken into account at first. Although the microscope was set up for every experiment in a consistent way – the lasers were aligned, and the same laser power and camera were always used – technical variation still caused systematic biases. Furthermore, for this first set of experiments, HP1 $\beta$ -Halo(N) cells were not FACS-sorted, and thus the variation in HP1 $\beta$  levels may also have influenced the data. As a result, the results were not reproducible, and this data is only reviewed briefly in this chapter.

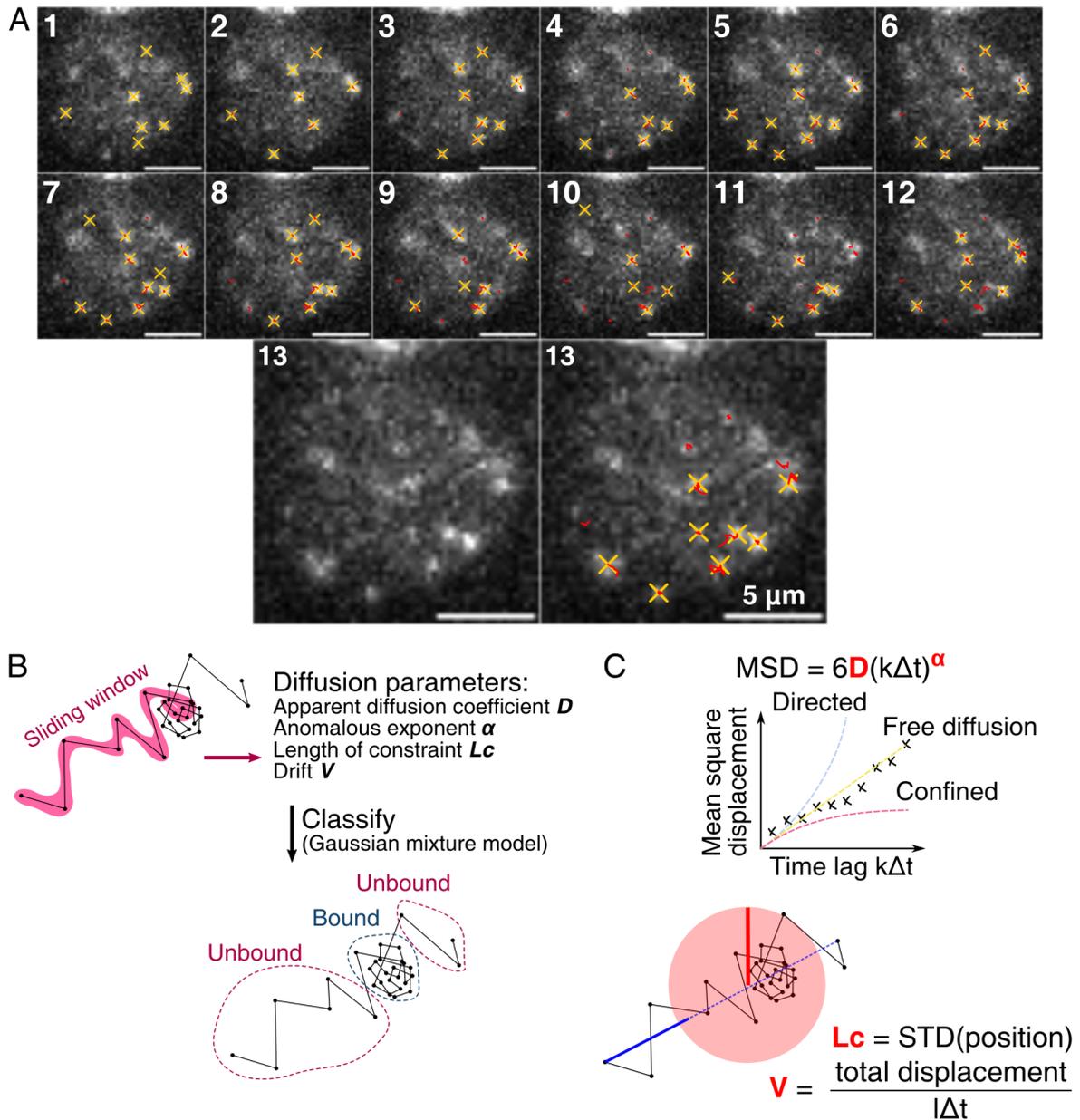
The dual-colour 2D SPT datasets were then collected for high-expressing HP1 $\beta$ -Halo(N) mESCs, with data for all differentiation stages acquired on the same day. Due to the time constraints and occasional labelling failures, only two full datasets of the three differentiation time points were obtained, one with HP1 $\beta$  double-labelling and one with DNA labelling.

### 3.2.2 Data analysis overview

In the first step, individual localisations were identified in the images, and the pathways taken by molecules were marked as trajectories (Figure 3.4A). The number of trajectories obtained is shown in Table 3.1.

Next, eu- and heterochromatin were segmented manually, using the wide-field images of DNA or HP1 $\beta$ , and the features of HP1 $\beta$  movement in eu- and heterochromatin were compared. In particular, the 4P algorithm (Shukron et al. 2019; Basu et al. 2020) was employed to separate freely diffusing molecules from confined HP1 $\beta$ , and to calculate their diffusion parameters. This step was especially important for investigation of features of freely diffusing HP1 $\beta$  in the context of the LLPS hypothesis. Furthermore, the angles between the consecutive steps were analysed to reveal the tendency for motion reversal in the trajectories.

To complement the results, the spatial variation in the features of HP1 $\beta$  diffusion was investigated and related to the local DNA or HP1 $\beta$  density without explicit seg-



**Figure 3.4:** A. Example of localisation detection (yellow crosses) and tracking (red trajectories) in 13 subsequent frames imaging HP1 $\beta$ , labelled with PA-JF<sub>646</sub>, with 5 ms exposure. The last frame is shown with and without annotation. Scale bars, 5  $\mu\text{m}$ . B. Workflow of 4P analysis. Four diffusion parameters are calculated along each trajectory, using sliding windows of a specified length. The collection of parameters is then used to classify parts of each trajectory as confined or unconfined movement, with the assumption that the confined state corresponds to molecules bound to chromatin and the unconfined state corresponds to free molecules. The characteristics of each population can then be studied separately. C. Illustrations of the diffusion parameters. In 3D, the apparent diffusion coefficient  $D$  and the anomalous exponent  $\alpha$  are calculated by fitting the curve of mean squared displacement  $MSD$  versus the time lag  $k\Delta t$ .  $\alpha = 1$  indicates free diffusion, while higher or lower values correspond to confined or directed motion, respectively. The length of confinement  $Lc$  is the standard deviation of a molecule's position, representing the radius of the sphere which the particle typically occupies (red). Drift  $V$  indicates the preference for a certain direction of movement, normalised by the observation time  $l\Delta t$  (blue).

mentation of heterochromatin.

The segmentation was also used to identify the boundaries between eu- and heterochromatin, and to seek for features consistent or inconsistent with HP1 $\beta$  LLPS.

Finally, the dynamics of HP1 $\beta$  in mESCs in different pluripotent states was compared.

### 3.2.3 Improvements and testing of the 4P algorithm

#### Overview of the 4P analysis

In this study, the four-parameter (4P) diffusion analysis algorithm (Shukron et al. 2019; Basu et al. 2020), was employed and adapted to the current data. Briefly, each trajectory was analysed using a sliding window to enable detection of a change in the molecule’s behaviour (Figure 3.4B). A set of four diffusion parameters, described below, was then calculated for each window (Figure 3.4C).

1. The apparent diffusion coefficient is defined as

$$D = \frac{MSD}{2d\Delta t}, \quad (3.1)$$

where  $MSD$  is mean square displacement,  $d = 2$  is the dimensionality, and  $\Delta t$  is the time resolution.  $D$  describes how fast a particle diffuses and, in the case of Brownian motion, depends on the size and shape of the molecule, the structure of the medium and the absolute temperature.

2. The anomalous exponent  $\alpha$  is obtained by fitting the following more general diffusion equation to the data:

$$MSD = 2dD(k\Delta t)^\alpha, \quad (3.2)$$

where  $k = 1, 2, 3, \dots$  signifies the increasing time lag, over which the  $MSD$  is computed.  $\alpha = 1$  indicates Brownian motion (linear dependence of  $MSD$  on

$k\Delta t$ ),  $\alpha < 1$  shows that the movement is confined or subdiffusive, and  $\alpha > 1$  indicates active, or superdiffusive motion.

3. The length of constraint

$$L_c = STD(A) \quad (3.3)$$

is the standard deviation of the molecule's position  $A$ . The length of constraint characterises the radius of the imaginary sphere that the particle explores.

4. The magnitude of the drift vector characterises the total displacement of a molecule in a unit of time:

$$V = \frac{\Delta L}{l\Delta t}, \quad (3.4)$$

where  $\Delta L$  is the total displacement of the trajectory, and  $l$  is the trajectory length in steps ( $l\Delta t$  is the total time of observation). The drift indicates the preference of a molecule to move in a specific direction rather than randomly diffuse.

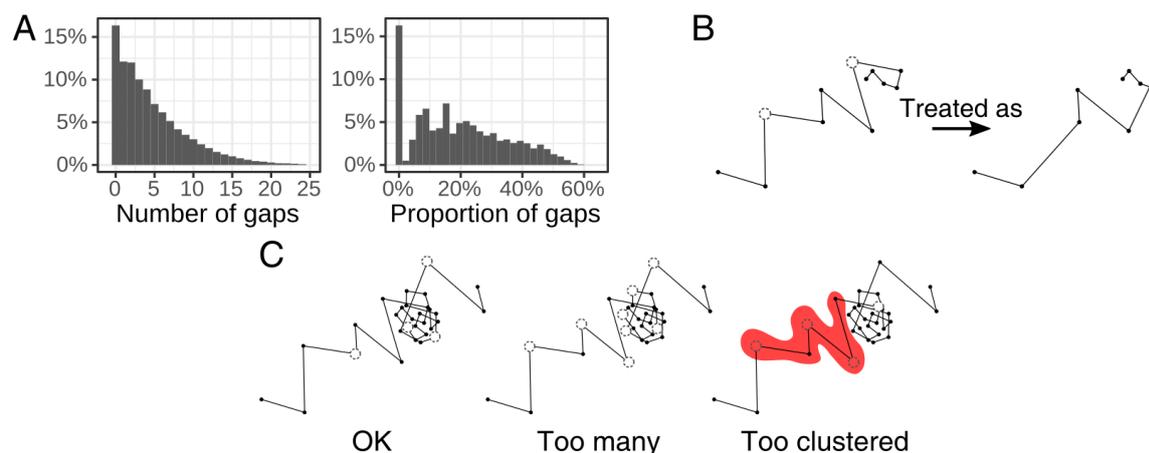
Since molecules bound to chromatin display markedly different characteristics to the ones that are not bound, these parameters can be used to classify each point as “confined” or “bound”, or as “unconfined” or “unbound”. The classification is performed using a Gaussian mixture model (GMM). Importantly, the method assumes that such two populations exist – in other words, the presence of two populations is forced upon the data. This is a reasonable expectation for a chromatin-binding protein or complex, but it is also important to verify this with an orthogonal method, such as jump distance analysis.

#### **4P algorithm adaptation for missing points**

Since localisation detection is not perfect, sometimes a molecule may not be identified in one frame, while being well-defined in the neighbouring frames. This leads to the trajectory being split in half and both halves likely not passing the length filter. However, if this point was allowed to be “missing” from the trajectory, many more trajectories could be retrieved. Importantly, this approach is only possible when the

number of localisations in a frame is low ( $<20$  in this study) to avoid erroneous joining of different trajectories.

To investigate the usefulness of such an approach, up to two missing points in a row were allowed at the tracking stage. Only 16% of trajectories did not have any gaps, and 25% had one or two missing points, suggesting that including these in subsequent analysis would substantially increase the amount of data retrieved (Figure 3.5A).



**Figure 3.5:** A. Histograms of the total number of gaps and the proportion of missing points in trajectories with  $>10$  points from all 3D sptPALM datasets. B. Without modification, the analysis of trajectories with missing points would be incorrect. C. The number and distribution of gaps needs to be considered when evaluating the quality of trajectories. To remove trajectories with clustered gaps, a sliding window of length  $s$  is used (highlighted in red).

To achieve this, the data analysis and diffusion parameter calculation had to be adjusted for trajectories with missing points (Figure 3.5B). This was accomplished in three stages:

1. Filtering out low-quality trajectories;
2. Modifying parameter calculation to take varying  $\Delta t$  into account; and
3. Classification of missing points.

**Filtering** Good-quality trajectories with few gaps were selected by first filtering out tracks with proportion of missing points  $\geq p$  ( $p = 0.25$  was used). Next, tracks were removed if in a sliding window of length  $s$  they contained  $\geq n$  gaps ( $s = \{5, 7, 9\}$  and

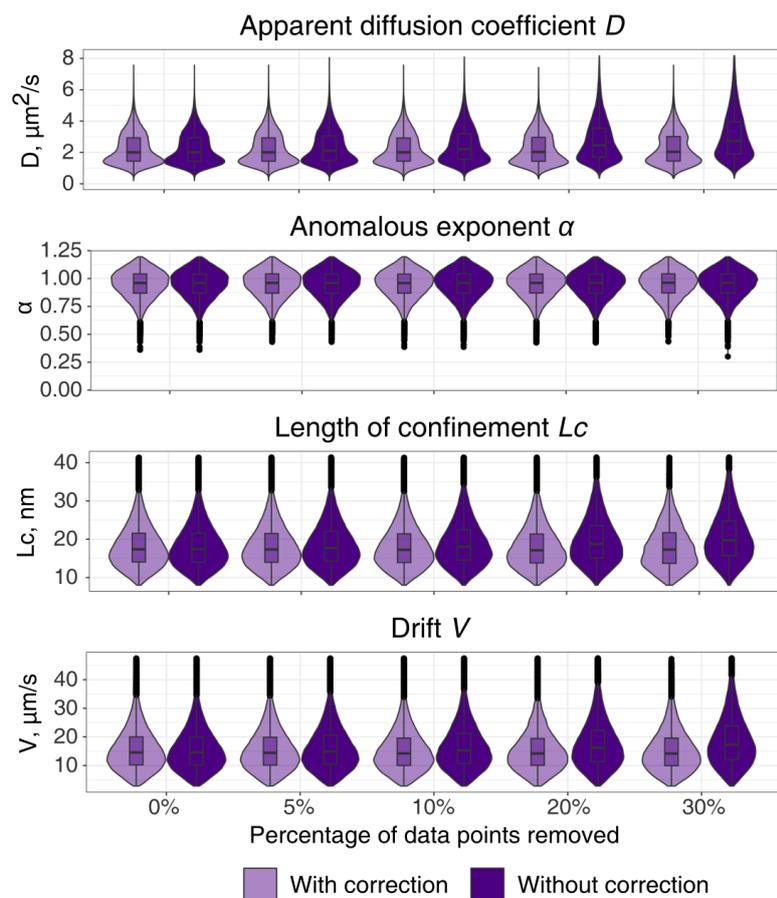
$n = \{3, 4, 5\}$ ) (Figure 3.5C). The final number of long trajectories after filtering for the datasets in this study is shown in Table 3.1.

**Parameter calculation** The diffusion parameter calculation had to be adapted to a varying  $\Delta t$  between the subsequent points in the trajectories. For the apparent diffusion coefficient, every squared displacement term in the *MSD* calculation was divided by the correct time interval. The length of confinement and the drift magnitude were normalised by the absolute trajectory length, while the anomalous exponent calculation remained unchanged (see section 6.4.5).

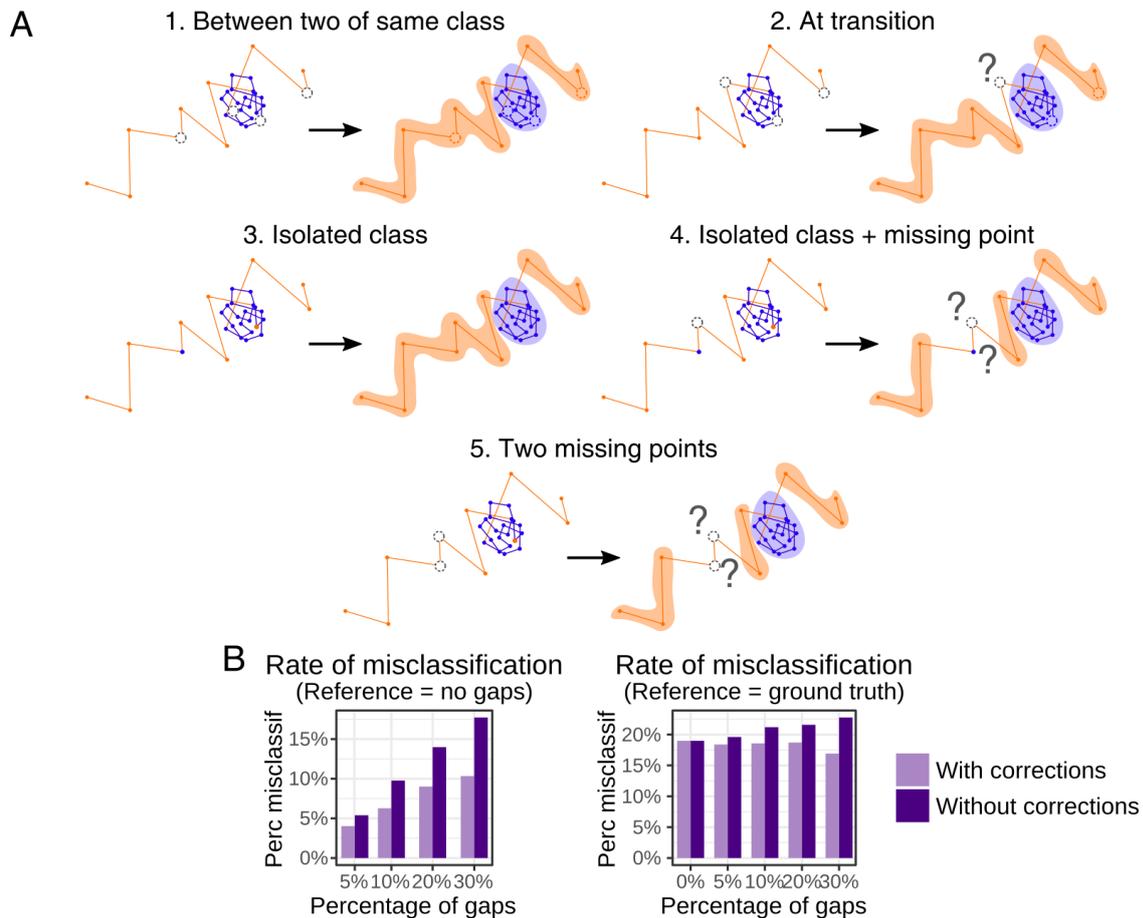
An empirical comparison of the performance of the original and the modified versions of the algorithm for a simulated data set (generated using a two-state model of diffusion from Basu et al. (2020)) with gaps indicated a very substantial increase in the robustness of parameter estimation (Figure 3.6). Tests were performed on six simulated datasets, and the results were similar across the parameter space explored (see section 6.4.5). Importantly, the results obtained from the experimental data also changed considerably when the updated algorithm was used for analysis.

**Classification** The 4P method classifies each point as confined or unconfined, based on the diffusion parameter values calculated in the window of  $\pm 5$  points around it. To infer the likely state of missing points, their neighbour points were considered. If the neighbours on both sides were of the same category, the point was also assigned to this class, while if they belonged to different classes, it remained unclassified (Figure 3.7A). The same procedure was used to convert “isolated” points, where a small transient change in diffusive behaviour lead to this point being classified differently from its neighbours, since these are likely to be artifacts.

The classification accuracy for non-missing points with an increasing proportion of gaps in the simulated data was compared between the original and the modified algorithms. Misclassification rates increased with higher proportions of missing points, but the modified algorithm appeared to cope with the task better, likely reflecting the more accurate diffusion parameter calculation (Figure 3.7B).



**Figure 3.6:** Comparison of diffusion parameter estimation by the algorithm with and without corrections for gaps. Violin plots of calculated parameters are shown for all points along the trajectories from a simulated dataset, where increasing amounts of data were removed randomly.

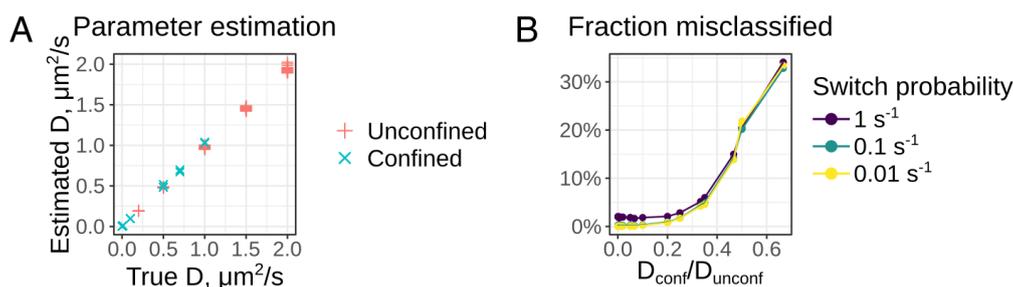


**Figure 3.7:** A. Different scenarios for missing point classification. 1. When both neighbours of a missing point are of the same class, we assume that that point also belongs to this class. 2. When the point's neighbours are of different classes, we cannot classify it. 3. To reduce the probability of an incorrect outlier classification, isolated class points are converted into the class to which its neighbours belong. 4. If the missing point is next to an isolated class point, neither of them are assigned any class. 5. Two missing points in a row remain unassigned. B. Misclassification rates by the two versions of the algorithm with different proportions of missing points introduced. The evaluation is either with respect to how the points were classified when there were no gaps (left), or the ground truth known for the simulated data (right).

## Evaluation of the 4P algorithm performance

Although the performance of the algorithm has been evaluated before (Basu et al. 2020), additional tests were conducted with simulations of data with 5 ms exposure, exploring the parameter space relevant for HP1 $\beta$  (see below). I was particularly interested in characterising the classification accuracy of the algorithm, since the initial assessment indicated that it may not be very good (Figure 3.7B). 67 datasets were simulated, varying the diffusion coefficients of the confined and unconfined fractions and the switching dynamics.

The diffusion coefficient of each population was estimated very accurately (Figure 3.8A). Since the other three parameters were not directly specified in the simulation, their calculated values could not be compared to the ground truth. The classification accuracy depended very strongly on the ratio between the diffusion coefficients of the populations (Figure 3.8B). When the unconfined population diffusion coefficient was at least three times larger than that of the confined population, the misclassification rate did not exceed 5%. Overall, the 4P algorithm was able to accurately distinguish two molecular populations displaying confined and unconfined behaviour.



**Figure 3.8:** A. Apparent diffusion coefficient estimated by the 4P algorithm vs the ground truth. Each point corresponds to the confined or unconfined population of a simulation, and the y axis indicates the median value for that population. B. The fraction of misclassified points as a function of the ratio between the diffusion coefficients of the populations, at different switch probabilities between the unconfined and confined states (equal in both directions).

### **3.2.4 HP1 $\beta$ exhibits different behaviour in eu- and heterochromatin**

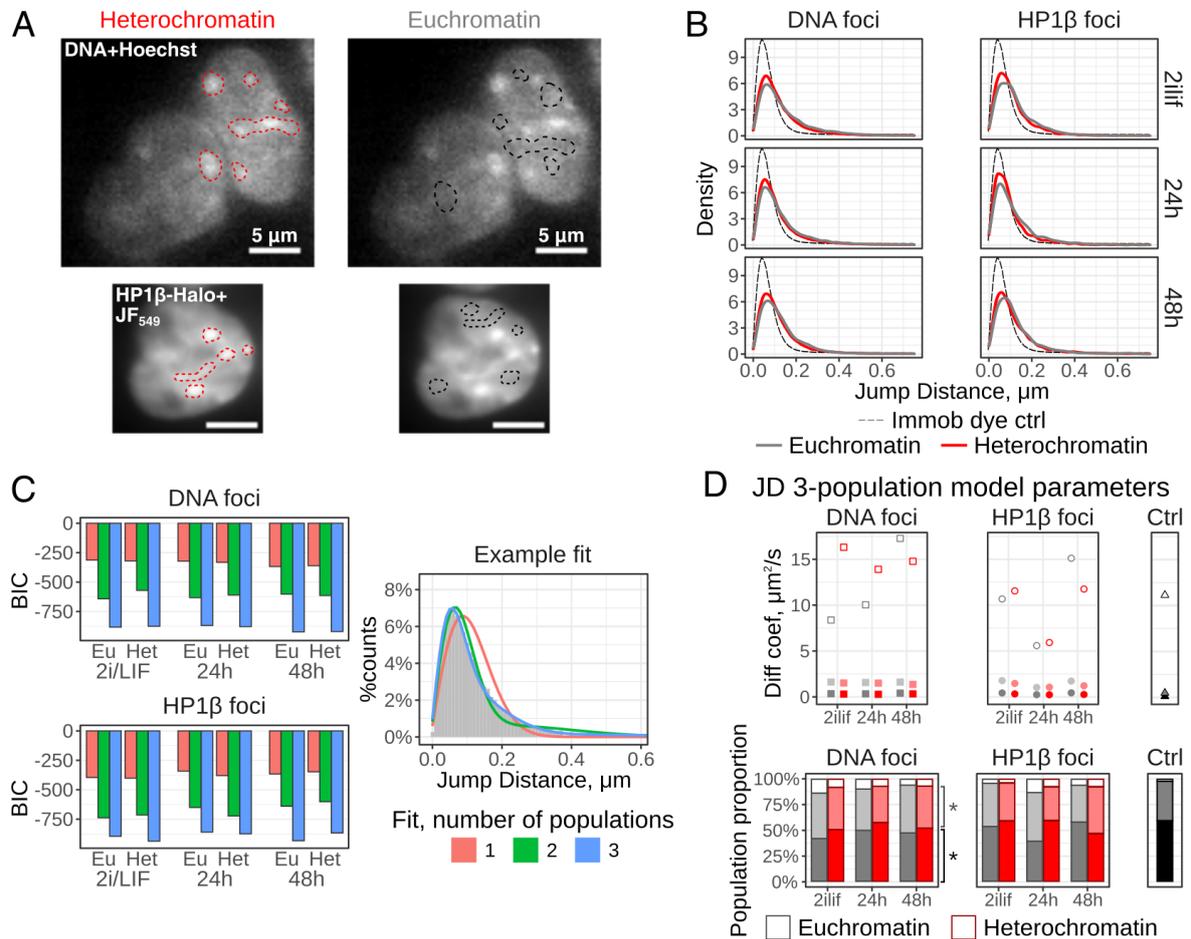
To compare the diffusion characteristics of HP1 $\beta$  in eu- and heterochromatin, the DNA or HP1 $\beta$  wide-field images were used to segment the foci (Figure 3.9A, left). Only the trajectories that were entirely within heterochromatic domains were considered. Since the masking process may have introduced a bias towards slower-moving or more confined trajectories, regions of the same size and shape were randomly selected in euchromatin for comparison (Figure 3.9A, right).

#### **Jump distance analysis reveals three populations of HP1 $\beta$**

To initially explore HP1 $\beta$  dynamics, jump distance analysis was conducted. Plotting the distributions of displacement lengths in one frame ( $\Delta t = 5\text{ms}$ ) indicated that HP1 $\beta$  in heterochromatin displayed noticeably shorter displacements than in euchromatin, although the difference was not large (Figure 3.9B).

The displacement distributions were best described by a model with three populations of HP1 $\beta$  (Figure 3.9C), with apparent diffusion coefficients of 0.2-0.4  $\mu\text{m}^2/\text{s}$  (dominated by localisation imprecision), 1-2  $\mu\text{m}^2/\text{s}$  and 6-16  $\mu\text{m}^2/\text{s}$  (Figure 3.9D). The diffusion coefficients of the populations were not significantly different between eu- and heterochromatin. However, the slowest population had a significantly higher presence in DNA-dense foci, while the result for HP1 $\beta$  clusters was less consistent.

The jump distance analysis indicated that there were differences in HP1 $\beta$  movement in eu- and heterochromatin, confirmed that there were at least two populations of HP1 $\beta$ , and suggested that potentially the relative proportions of these populations were different between the compartments. Thus, the 4P analysis was applied to separate the trajectories of chromatin-bound and freely diffusing molecules and analyse their motion in detail.



**Figure 3.9:** A. The heterochromatic domains (left) were segmented, using the DNA (top) or HP1β (bottom) wide-field images. Regions of the same size were selected in euchromatin for comparison (right). B. Density curves of jump distance distributions for HP1β in mESCs in 2i/LIF, and 24h or 48h after the onset of differentiation, in eu- or heterochromatin, or for a control experiment with immobilised PA-JF<sub>646</sub>. The differences between eu- and heterochromatin were highly significant (Kolmogorov-Smirnov test,  $p < 10^{-7}$ ). C. The Bayesian information criterion (BIC) values for one-, two- and three-population jump distance model fits to the HP1β SPT data, used to select the best model (one with the lowest BIC) (left). An example of the model predictions plotted over a jump distance histogram (2i/LIF, DNA labelling, heterochromatin; right). D. Parameters of the three-population jump distance model fits to the HP1β SPT data or immobilised dye control. The diffusion coefficient and proportion of each population are shown in the same colour (dark, light or white). The differences between the parameters for eu- and heterochromatin was tested using paired t-test (\*,  $p < 0.05$ ), and they were found to be non-significant unless indicated.

### Lower limit of sensitivity

As mentioned in section 3.1.2, microscope resolution defines the lower limit for the estimated diffusion parameters at a given frame rate. Several approaches were used to check whether it was possible to accurately measure the diffusion parameters of confined molecules.

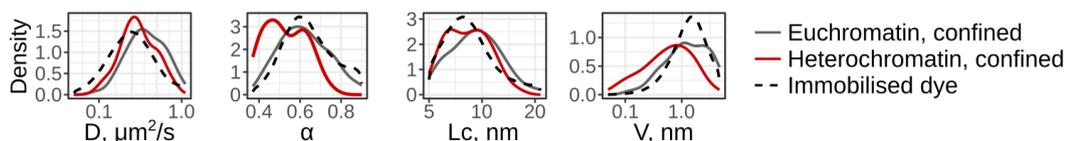
Firstly, the theoretical limit of the method was estimated. At the localisation stage of data pre-processing, a value of 50 nm was chosen as the precision threshold for PSF detection. Since precision is the standard deviation in the detected position of a molecule, for an immobile emitter, ~68% of the time its estimated position will be within  $\sigma = 50$  nm around the centre point. At 200 fps, the apparent diffusion coefficient of an immobile molecule  $\delta$  in  $d = 2$  dimensions will be in the order of

$$\delta = \frac{(2\sigma)^2}{2d\Delta t} = 0.5 \mu\text{m}^2 \text{ s}^{-1}. \quad (3.5)$$

The sensitivity of short-exposure SPT was also tested experimentally by imaging PA-JF<sub>646</sub> immobilised on a coverslip. As expected, the control data displayed noticeably shorter displacement lengths than HP1 $\beta$  (Figure 3.9B). The diffusion coefficient, estimated assuming a single population of dye molecules, was  $0.3 \mu\text{m}^2/\text{s}$ , similar to the slowest population of HP1 $\beta$  in the live data. However, surprisingly, the control data was described best (as evidenced by BIC) by a diffusion model with three populations with  $D = 0.2 \mu\text{m}^2/\text{s}$ ,  $0.5 \mu\text{m}^2/\text{s}$  and  $11 \mu\text{m}^2/\text{s}$  (Figure 3.9D). The slowest and the fastest molecules diffused similarly to the corresponding HP1 $\beta$  populations, although the abundance of the latter was very low (3%). The medium-speed particles, however, displayed a lower apparent diffusion coefficient than HP1 $\beta$ . Altogether, these observations indicate that the slowest HP1 $\beta$  population is essentially stationary, while the medium population displays noticeable motion. In addition, the fast population might be an artifact, since it also appears in the immobilised control data, although there its proportion is lower than in the HP1 $\beta$  SPT.

The control data was then analysed using the 4P algorithm and compared to the measurements for the confined population of HP1 $\beta$  (Figure 3.10). All parameters of

HP1 $\beta$  diffusion displayed values similar to or even lower than the dye control, suggesting that characterising this population was not possible with this frame rate and spatial resolution. Similar results were obtained by comparing 3D sptPALM data for confined HP1 $\beta$  with immobilised PA-JF<sub>549</sub> or labelled fixed cells.

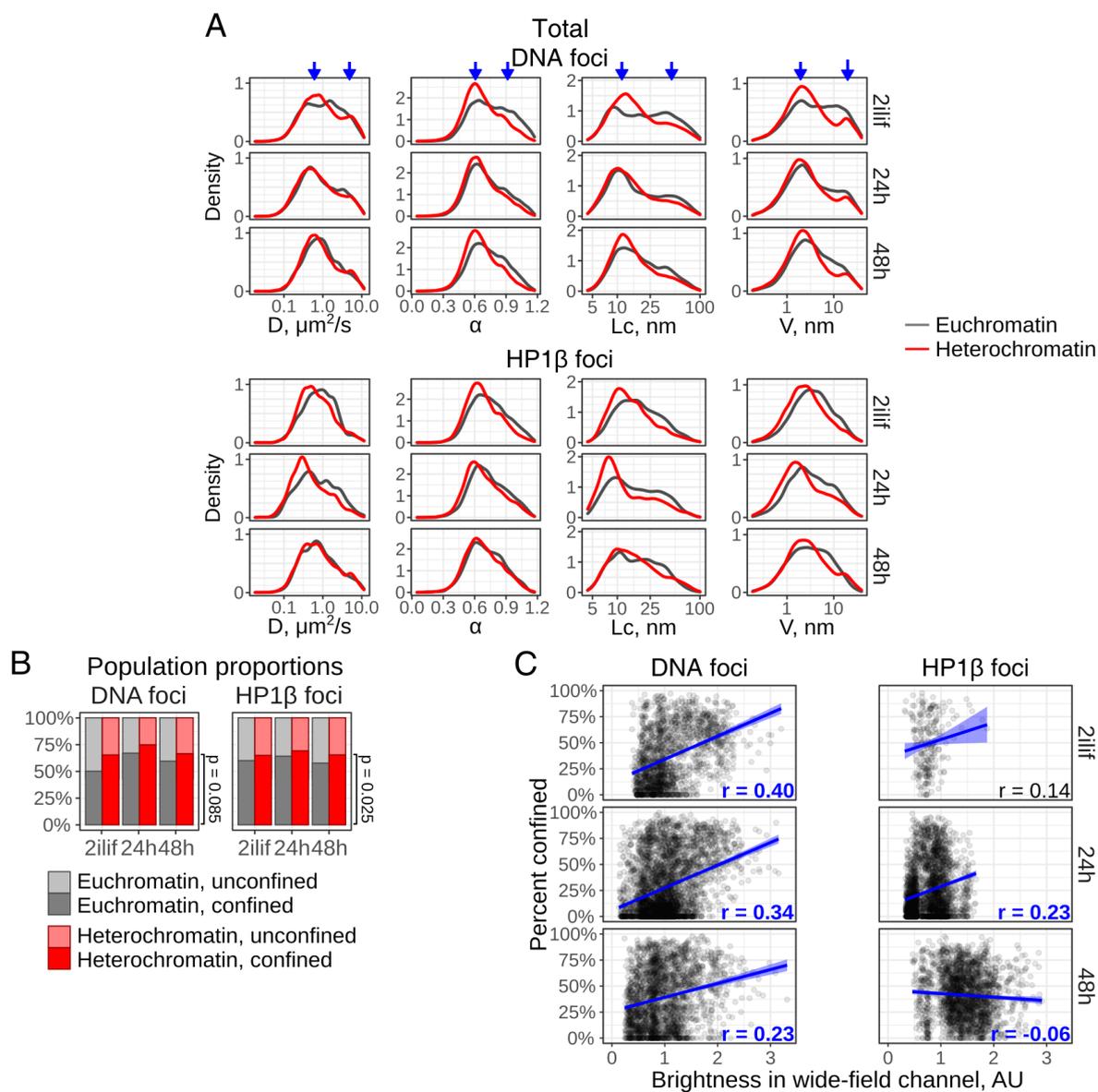


**Figure 3.10:** Comparison of the distributions of apparent diffusion coefficient  $D$ , anomalous exponent  $\alpha$ , length of confinement  $Lc$  and drift magnitude  $V$  values for the confined population of HP1 $\beta$  inside HP1 $\beta$  foci or in euchromatin of formative-state mESCs with immobilised PA-JF<sub>646</sub> dye. In this and subsequent graphs, the distribution of  $\alpha$  is shown on a linear scale, and the other parameters are shown on a log scale.

#### 4P analysis indicates a higher percentage of bound HP1 $\beta$ in heterochromatin

Next, the 4P analysis was applied to the HP1 $\beta$  trajectories in eu- and heterochromatin. Plotting of the diffusion parameters revealed that a) the data had very wide distributions and often double peaks, further justifying the classification procedure; b) the peak corresponding to the faster population generally appeared more pronounced in euchromatin; c) the apparent diffusion coefficient was not the most discriminative feature of HP1 $\beta$  diffusion in eu- or heterochromatin (Figure 3.11A).

Classification of molecules confirmed that the proportion of chromatin-bound HP1 $\beta$  was higher in heterochromatin (Figure 3.11B). The difference was more pronounced for DNA foci ( $\sim 9\%$ ) than for HP1 $\beta$  foci ( $\sim 5\%$ ), but due to higher variability it was not statistically significant in the former case. To support this finding with an approach independent from segmentation of foci, the relationship between the percentage of confined molecules and the concentration of DNA or HP1 $\beta$  was investigated. A clear correlation was found between DNA density and percentage of confined HP1 $\beta$  molecules, while the results were less consistent for HP1 $\beta$  density (Figure 3.11C).



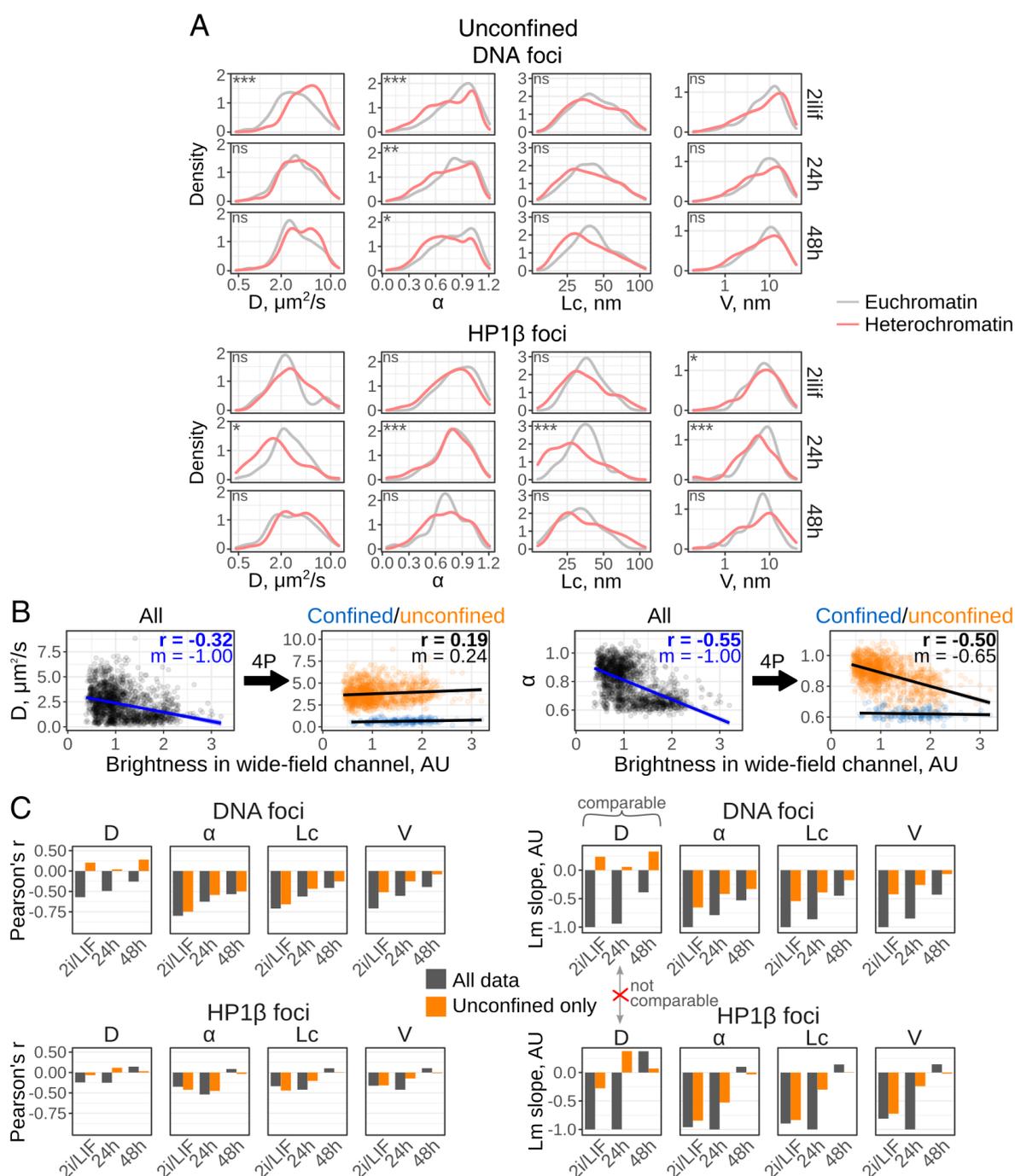
**Figure 3.11:** A. Distributions of the diffusion parameters for HP1 $\beta$  in eu- and heterochromatin in naive, formative and primed mESCs. Arrows point to the two peaks apparent in some of the plots. The Kolmogorov-Smirnov test indicated that the differences for all datasets were highly significant due to large sample size. B. Proportions of the confined and unconfined populations in eu- and heterochromatic trajectories of HP1 $\beta$  in mESCs during exit from naive pluripotency. P-values from paired t-test; for two datasets together,  $p = 0.005$ . C. Scatter plots of the pixel brightness values in the wide-field channel (either DNA or HP1 $\beta$ ) and the proportion of molecules within that pixel classified as confined. Only pixels with at least 10 trajectories within them were considered. A linear fit to the data is shown, as well as Pearson correlation coefficient  $r$  (in blue and bold if significantly different from 0).

## Unbound HP1 $\beta$ moves in heterochromatin with a similar diffusion coefficient, but is more confined

To investigate whether on top of containing a higher proportion of bound HP1 $\beta$  protein, heterochromatin also slowed down the movement of freely diffusing molecules, the diffusion parameters of the unconfined fraction of HP1 $\beta$  were compared inside and outside foci. The differences were small – the distributions of the anomalous exponent and length of confinement for heterochromatin tended to be slightly shifted to the left, and the drift magnitude had a slightly larger dispersion – and their statistical significance depended largely on the sample size (Figure 3.12A). The trends for diffusion coefficient were inconsistent, but in some cases HP1 $\beta$  inside foci appeared to move even faster than outside.

The correlation between the diffusion parameters and the concentration of DNA or HP1 $\beta$  was studied next. The Pearson correlation coefficient and the slope of the linear fit between each parameter and DNA/HP1 $\beta$  signal intensity were calculated, either for all trajectories before classification, or for confined and unconfined molecules separately (Figure 3.12B). As expected due to the higher percentage of bound HP1 $\beta$  in heterochromatin, all parameters showed an overall anticorrelation with DNA and HP1 $\beta$  density, although in the latter case it was weaker and less consistent (Figure 3.12C, gray bars).

Investigating solely the unconfined HP1 $\beta$  trajectories, however, revealed an important distinction between the parameters. The apparent diffusion coefficient  $D$  showed no or positive correlation with DNA or HP1 $\beta$  density (Figure 3.12C, orange bars). Thus, the negative correlation for all data was predominantly caused by the higher percentage of confined molecules in heterochromatin. In contrast, the anomalous exponent  $\alpha$ , the length of constraint  $Lc$  and the drift magnitude  $V$  were still strongly anticorrelated with DNA/HP1 $\beta$  signal, although the consistency of the correlation (correlation coefficient  $r$ ) and the magnitude of the relationship (slope of the linear fit  $m$ ) were somewhat lower. Therefore, chromatin-unbound HP1 $\beta$  in heterochromatin diffuses with a similar instantaneous velocity compared to euchromatin, consistent with a



**Figure 3.12:** A. Distributions of the diffusion parameters for unconfined HP1β in eu- and heterochromatin in naive, formative and primed mESCs. Due to additional trajectory length filters upon classification (see Methods), the HP1β-double-labelling dataset (bottom) contained little data (<150 trajectories per curve). The statistical significance of the differences between the distributions, determined by Kolmogorov-Smirnov test, is indicated: ns, non-significant; \*,  $p < 0.05$ , \*\*,  $p < 0.01$ , \*\*\*,  $p < 0.001$ . B. Examples of scatter plots of the pixel brightness values in the wide-field channel (either DNA or HP1β) and the median value for the apparent diffusion coefficient  $D$  and the anomalous exponent  $\alpha$  for trajectories within the pixel (2i/LIF, DNA labelling), for all trajectories (left) or after classification with 4P algorithm (right). Pearson's correlation coefficient  $r$  (in bold if significantly different from 0) is indicated (on the right, for the unconfined population).  
*Legend continued on the next page.*

previous characterisation of HP1 $\alpha$  rotational diffusion (Erdel et al. 2020), but displays increased confinement over longer distances.

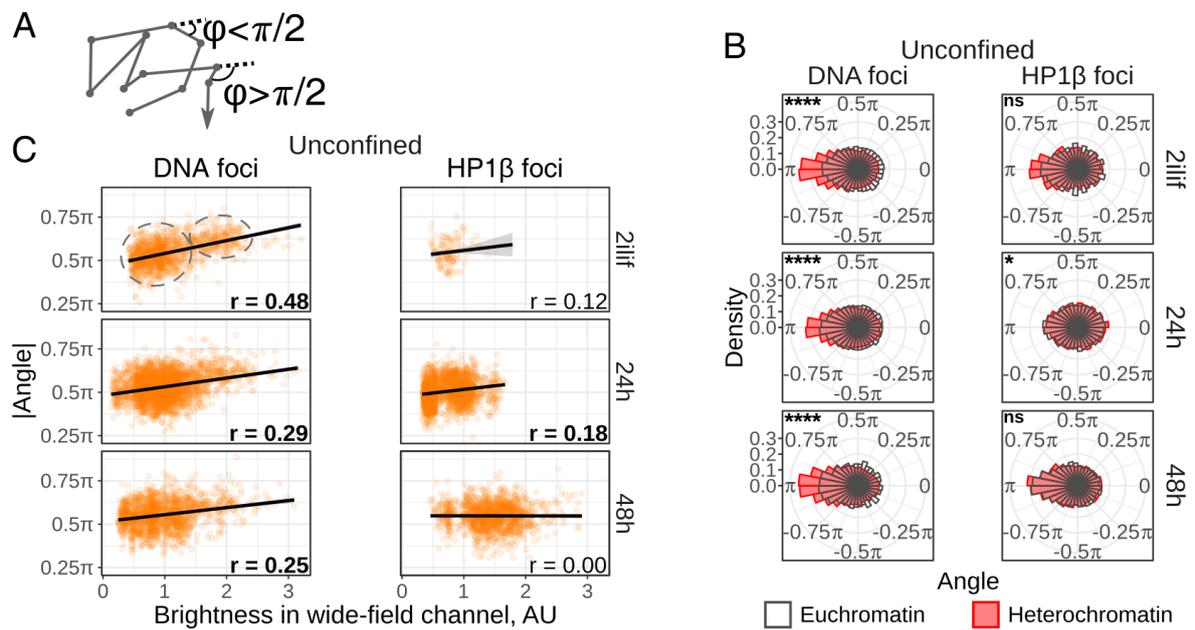
### Angle analysis confirms increased confinement in heterochromatin

Measurement of the distribution of the angle between the consecutive steps in a trajectory can also be used to assess confinement of molecules (Figure 3.13A), and the LLPS model was shown to display a stronger enrichment in angles  $>\pi/2$  compared to the polymer binding model (Heltberg et al. 2021). Since the study considered the motion of unbound molecules, the angle distributions were plotted for the unconfined HP1 $\beta$  trajectories, in eu- or heterochromatic regions. There was a clear bias for motion reversal of HP1 $\beta$  inside the DNA foci (Figure 3.13B, left). In contrast, there was no difference between the HP1 $\beta$  foci and euchromatin (Figure 3.13B, right), consistent with the weaker trends in this dataset observed previously.

A positive correlation between the mean absolute angle for unconfined HP1 $\beta$  trajectories and the DNA density was also observed, and a significant correlation with HP1 $\beta$  density was also found in one of the datasets (Figure 3.13C).

Interestingly, the points on the scatter plots of angle vs DNA staining brightness appeared to form two clouds (Figure 3.13C, left), and a similar pattern was observed for  $\alpha$  (Figure 3.12B, right) and other confinement parameters (not shown) vs DNA. The dependence of the diffusion features on the DNA density in each cloud seemed weaker compared to the overall trend. This might indicate a step-wise transition between the diffusive behaviour of HP1 $\beta$  in eu- and heterochromatin, rather than a gradual change.

**Figure 3.12:** A linear fit to the data is also shown alongside the slope of the line  $m$ , where  $m$  has been normalised by the maximum value within each dataset for simplicity. C. Pearson correlation coefficients (left) or slopes of the linear model fit (right) between the diffusion parameters of HP1 $\beta$  and the density of DNA (top) or HP1 $\beta$  (bottom), as shown in B. The slope has been normalised by the maximum value within each dataset for simplicity, and it is not comparable between datasets.



**Figure 3.13:** A. An illustration of the angle  $\varphi$  between the consecutive steps in a trajectory. B. Angle distributions for unconfined trajectories inside and outside DNA or HP1 $\beta$  foci. The statistical significance of the differences was determined by the Kolmogorov-Smirnov test: ns, non-significant; \*,  $p < 0.05$ , \*\*\*\*;  $p < 0.0001$ . C. Scatter plots of the pixel brightness values in the wide-field channel (either DNA or HP1 $\beta$ ) and the mean absolute angle  $\varphi$  for unconfined trajectories within that pixel. Only pixels with at least 10 trajectories within them were considered. A linear fit to the data is shown, as well as Pearson's correlation coefficient  $r$  (in bold if significantly different from 0). On the top-left graph, the two putative point clouds are indicated (grey dashed ellipses).

### 3.2.5 HP1 $\beta$ at the inter-compartment boundary

After characterising the diffusion of HP1 $\beta$  in eu- and heterochromatin, I set out to test the predictions of the LLPS model regarding HP1 $\beta$  diffusion near the phase boundary.

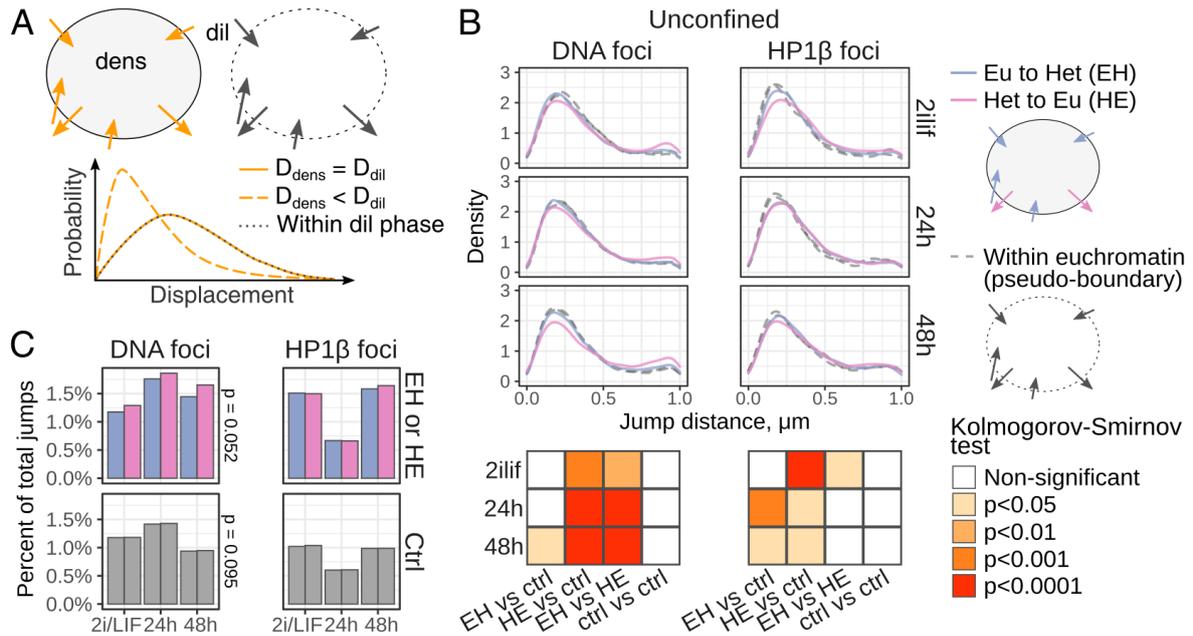
#### Displacements across the boundary

The jump distance distribution of particles crossing the inter-compartment boundary differs from that within one phase, when the diffusion coefficients of molecules in different phases are not equal (Bo et al. 2021) (Figure 3.14A). This phenomenon has been used before to gain insights into the nature of certain condensates (McSwiggen et al. 2019b; Collombet et al. 2021). In addition, comparison of the displacement distributions for molecules entering and exiting the condensate can reveal whether the system is in equilibrium (Bo et al. 2021).

Thus, the jump distance distributions for molecules crossing the boundary between eu- and heterochromatin were compared to the movement across analogous pseudo-boundaries within euchromatin (as shown in Figure 3.9A). Since the model by Bo et al. (2021) considers only a binary mixture without any stable binding of particles to chromatin, this analysis was performed on the unconfined trajectories.

Unconfined HP1 $\beta$  molecules entering heterochromatin did not show consistent deviations from the control, but the molecules leaving the foci displayed significantly longer displacements (Figure 3.14B). These distributions tended to be enriched for very fast molecules, potentially corresponding to the third HP1 $\beta$  population identified using jump distance analysis. This result indicates that the diffusion coefficient of HP1 $\beta$  within heterochromatin is similar to that in euchromatin, consistent with the findings described earlier. However, this is at odds with the previously measured slow-down of HP1 $\alpha$  at the inter-compartment boundary (Strom et al. 2017).

Furthermore, the displacement distributions for HP1 $\beta$  crossing the boundary from eu- to heterochromatin and the other way round were significantly different, especially when heterochromatin was defined using DNA density (Figure 3.14B). Importantly, the control curves for molecules entering or leaving the pseudo-foci were identical. This



**Figure 3.14:** A. Schematic of the jump distance curves, expected for molecules crossing the interphase boundary when the diffusion coefficients in the dense ( $D_{\text{dens}}$ ) and dilute ( $D_{\text{dil}}$ ) phases are identical or different, compared to the distribution for molecules crossing a random boundary within the dilute phase. B. Top, displacement distributions for unconfined HP1 $\beta$  molecules crossing the boundary between euchromatin (eu) and heterochromatin (het), defined either as DNA-dense (left) or HP1 $\beta$ -dense (right) regions. The control distributions for molecules crossing a pseudo-boundary within euchromatin are shown in gray. Bottom, results of statistical testing for differences between displacement distributions for molecules entering (EH) or exiting (HE) heterochromatin and the respective euchromatic controls (ctrl), as well as between the molecules entering and exiting heterochromatin or between the molecules entering and exiting the control pseudo-foci. C. The percentage of jumps crossing the real or the control compartment boundary in either direction, colour scheme same as in B. The difference between the entering or exiting trajectories was tested with paired t-test ( $p \gg 0.05$  for HP1 $\beta$  foci).

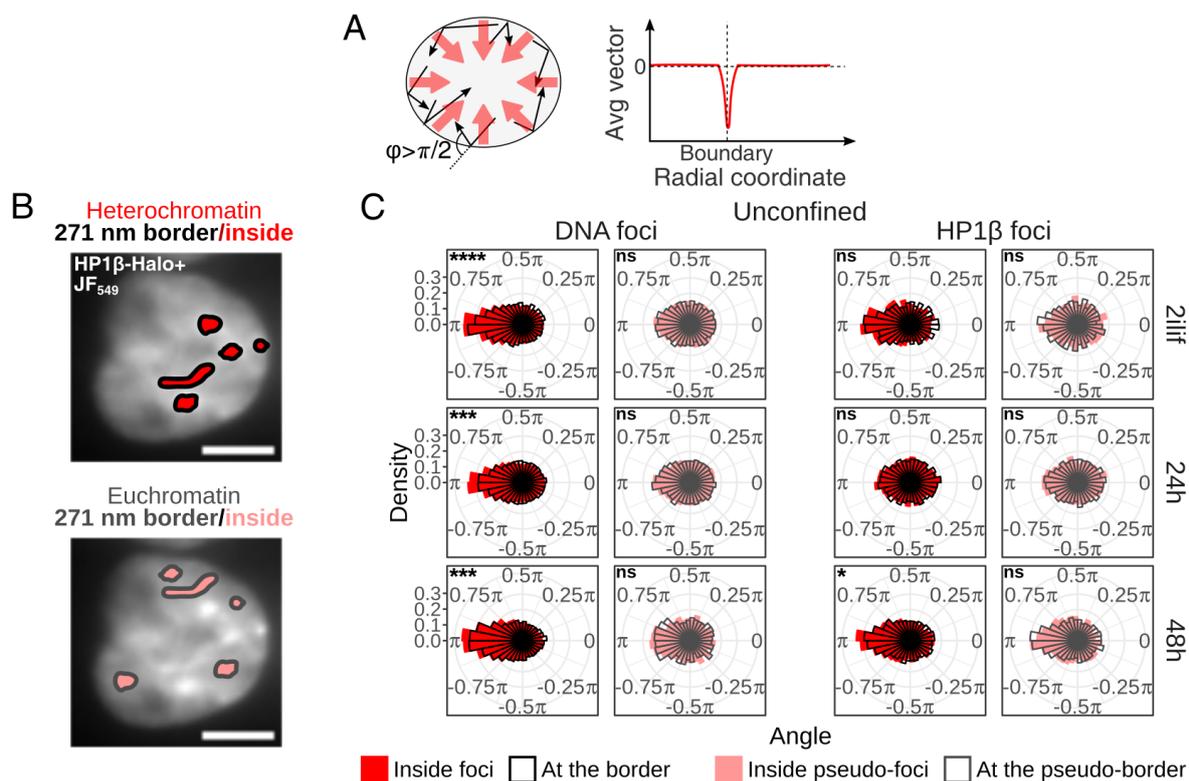
suggested that heterochromatin in mESCs might be in a non-equilibrium state with respect to HP1 $\beta$  flux. In addition, the proportion of molecules leaving the DNA foci was slightly higher than the proportion entering the foci, again potentially pointing to a non-equilibrium state (Figure 3.14C).

### **“Bouncing off” the interphase boundary**

The LLPS model predicts that molecules within the condensate should “bounce off” the interphase boundary, since leaving the dense phase is energetically unfavourable (Figure 3.15A). This results in particles displaying drift – i.e. coordinated movement – away from the boundary of the focus (Bo et al. 2021; Heltberg et al. 2021). On the other hand, the “bouncing” effect can also be detected as a shift in the distribution of the angles between trajectory steps (Collombet et al. 2021).

Constructing spatial maps of drift at 250 nm resolution did not reveal any effects on the level of individual foci, and the current data did not allow constructing such maps at higher resolution. Investigating the boundary-proximal drift effects on average, rather than for individual cells and foci, also did not indicate any significant differences from the control. Overall it appeared that the current design of the SPT experiments was not optimal for such an analysis due to a) poor resolution of the foci boundaries, b) localisation precision, and c) the amount of data from independent molecules that was collected in a given area.

Next, the angle distributions at the borders of the foci were compared with the heterochromatic interior. In general, there was a slightly stronger tendency for motion reversal inside the foci compared to at the border, which was not observed in the control regions (Figure 3.15B). Interestingly, this finding is opposite to the predictions of the LLPS model, although it is possible that the distributions could be skewed due to imprecise definition of the boundary.



**Figure 3.15:** A. Left, a schematic explaining two effects observed due to “bouncing off” of molecules (black arrows) near the phase boundary: an average drift vector pointing away from the boundary (large red arrows) and the tendency of the angle  $\phi$  to exceed  $\pi/2$ . Right, an illustrative plot of the radial drift vector vs the radial coordinate with respect to the centre of the focus. Note that the particular values and shape of the curve will depend on the parameters of LLPS. B. The 271 nm (one pixel), proximal to the foci borders from the inner side (black) were analysed separately from the inside (red). A similar procedure was repeated for regions of the same size and shape in euchromatin (dark gray/pink). C. Angle distributions of unconfined HP1 $\beta$  trajectories at the boundary or further inside the DNA or HP1 $\beta$  foci, or the corresponding pseudo-foci in euchromatin. The statistical significance of the differences was determined by the Kolmogorov-Smirnov test: ns, non-significant; \*,  $p < 0.05$ ; \*\*\*,  $p < 0.001$ ; \*\*\*\*,  $p < 0.0001$ .

### 3.2.6 HP1 $\beta$ displays slower dynamics in formative-state mESCs

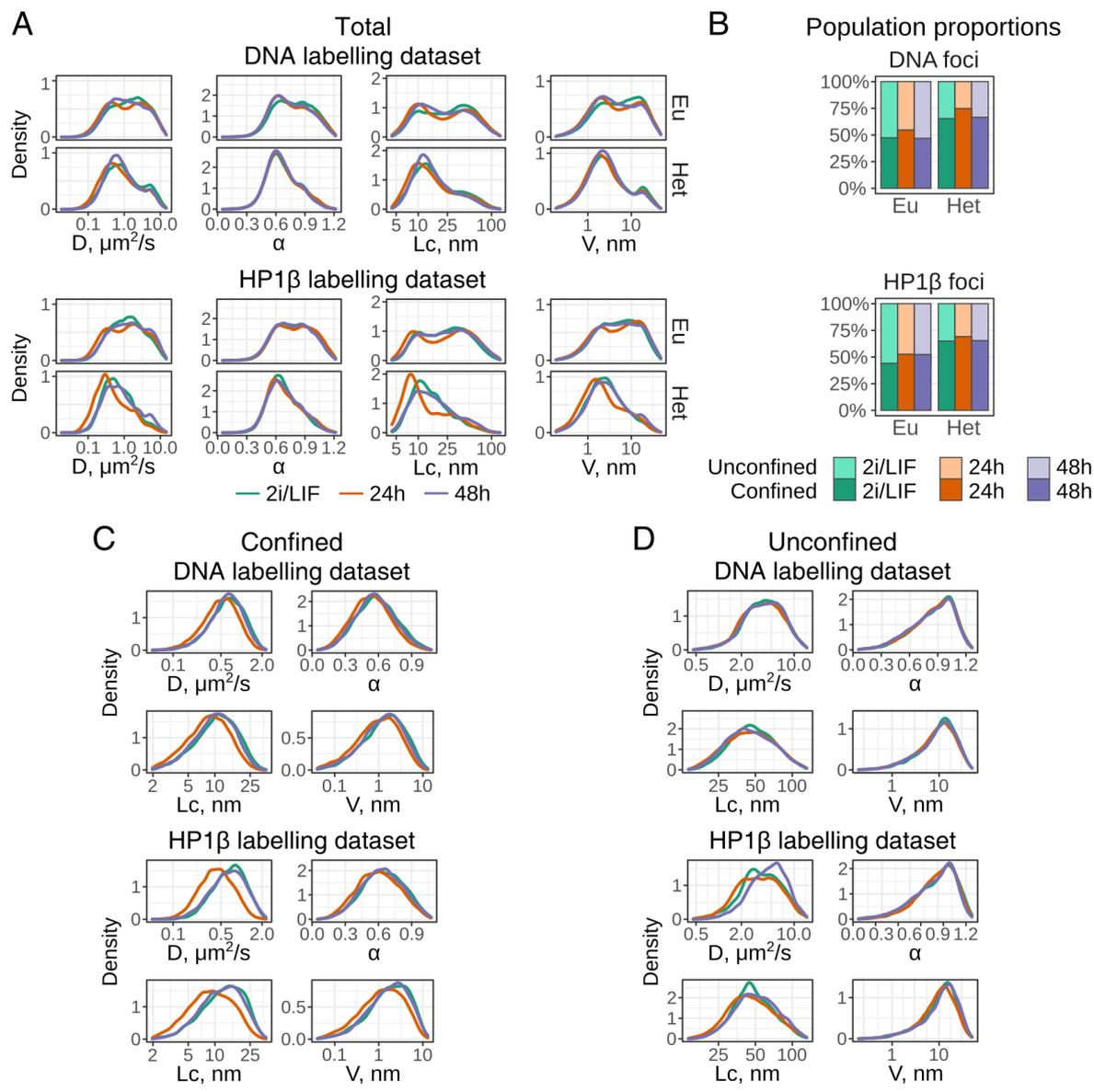
Finally, the diffusion of HP1 $\beta$  was compared between the three pluripotent states. Plotting the four diffusion parameters for all molecules suggested that there were subtle differences between the time points (Figure 3.16A). Curves for the formative-state mESCs for all parameters but the anomalous exponent were slightly shifted to the left, indicating that HP1 $\beta$  might diffuse slower at this stage. The 4P classification showed that there were consistently more confined HP1 $\beta$  molecules 24h after the onset of differentiation in both eu- and heterochromatin (Figure 3.16B). Notably, the same phenomenon was observed in the 3D SPT dataset.

Analysis of the confined and unconfined HP1 $\beta$  trajectories suggested that the change in dynamics in the formative state was present in both populations, although it was perhaps more consistent in the chromatin-bound fraction (Figure 3.16C,D). Thus, it appeared that overall HP1 $\beta$  diffusion became subtly slower at this time point, leading to more points classified as confined and a shift in all diffusion parameters.

### 3.2.7 Summary

Overall, in this section the single-molecule dynamics of HP1 $\beta$  inside and outside heterochromatic foci was examined. As expected, heterochromatin was found to contain a higher proportion of chromatin-bound HP1 $\beta$ . The non-chromatin-bound HP1 $\beta$  molecules did not display lower apparent diffusion coefficient in heterochromatin, but their motion was more confined, manifested in lower anomalous exponent, length of constraint and drift. Furthermore, HP1 $\beta$  trajectories in DNA-dense regions showed a strong tendency for reversal, although the data was less consistent for HP1 $\beta$  foci. Plotting the median/mean diffusion parameters in a pixel vs the DNA labelling intensity in that pixel revealed a non-linear relationship between the two, potentially indicating a qualitative difference between diffusion in eu- and heterochromatin.

Analysis of HP1 $\beta$  behaviour on the boundary between the chromatin indicated a possible non-equilibrium state of the system. No evidence for molecules “bouncing



**Figure 3.16:** A. Distributions of the diffusion parameters for all HP1 $\beta$  in eu- and heterochromatin in naive, formative and primed mESCs. All euchromatic trajectories are shown here, rather than the ones within pseudo-foci. Due to large sample sizes, differences between all conditions are statistically significant. B. Percentages of confined and unconfined molecules in each compartment during differentiation. C, D. Distributions of the diffusion parameters for the confined (C) or unconfined (D) trajectories. The data was not partitioned between eu- and heterochromatin to retain a large sample size. Note the difference in scales between C and D. The formative-state distributions for nearly all parameters are highly significantly different from the naive or primed state (Kolmogorov-Smirnov test,  $p < 0.001$ ).

off” the heterochromatin border was found. In contrast, trajectories appeared to show a weaker tendency for reversal at the boundary compared to heterochromatic interior.

Finally, a subtle change in HP1 $\beta$  diffusion was observed in mESCs 24h after the onset of differentiation, coinciding with an increase in HP1 $\beta$  concentration and change in localisation, as well as a global reorganisation of chromatin.

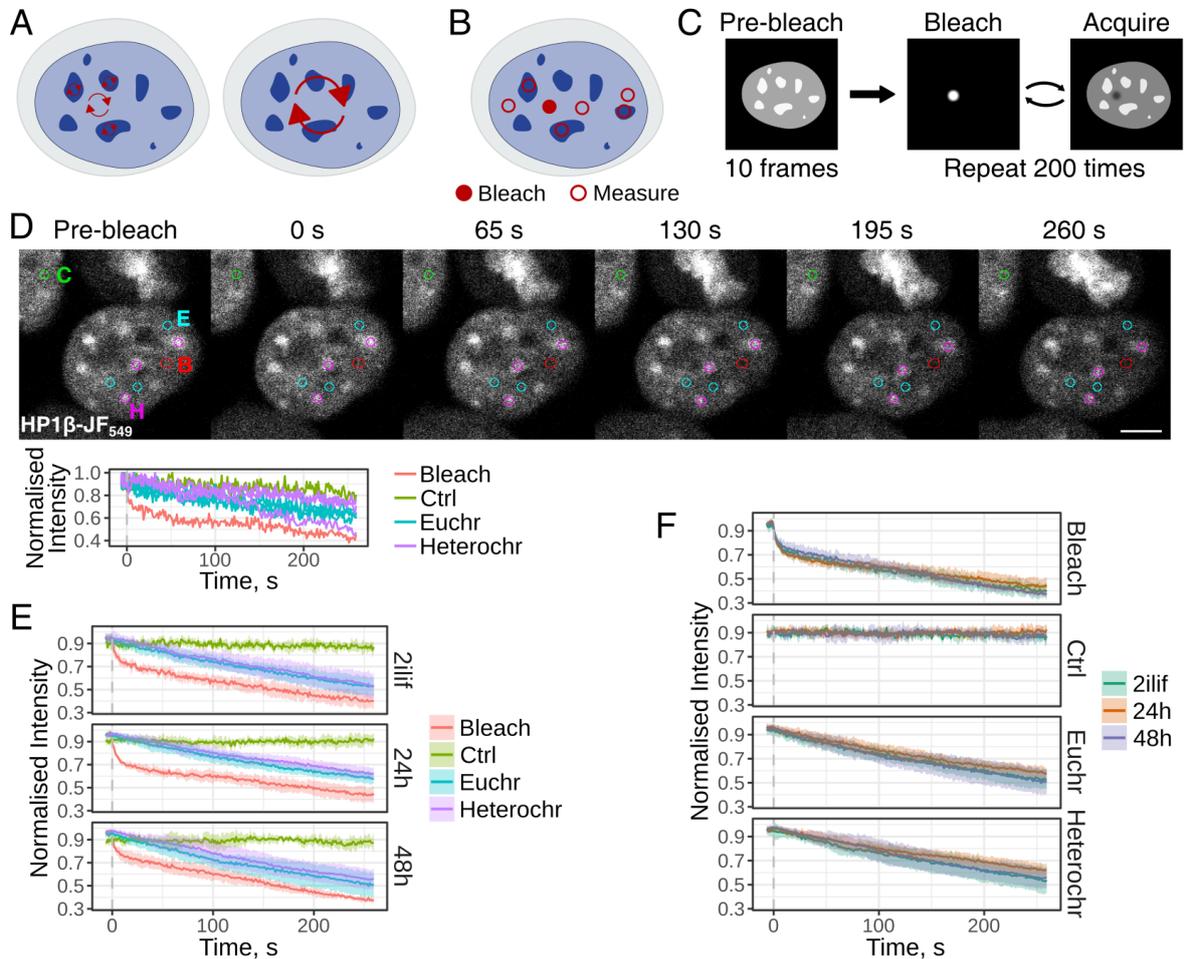
### 3.3 Analysis of HP1 $\beta$ mobility with FLIP

Although single-molecule imaging is a powerful technique for investigating the dynamic behaviour of molecules, ensemble features can also be studied by other methods. LLPS condensates might display preferential internal mixing (Figure 3.17A) (Erdel et al. 2020). To test whether HP1 $\beta$  in chromocentres displayed these properties, the fluorescence loss in photobleaching (FLIP) technique was employed. Similarly to FRAP, it involves bleaching the fluorophores, associated with the protein of interest, in a small area within the cell. Bleaching is typically performed continuously during acquisition (Figure 3.17C). A different area to that being bleached is monitored in the same cell, the loss in fluorescence intensity indicating the diffusion of bleached fluorophores into it. Thus, FLIP is a perfect tool for studying mixing of material between compartments.

#### 3.3.1 Overview

Due to the small size of chromocentres and their fast movement in mESCs it was not possible to compare diffusion within a heterochromatic compartment and between it and euchromatin. Thus, the experiment was designed to investigate mixing by bleaching a spot in euchromatin while monitoring several areas both in eu- and heterochromatin (Figure 3.17B).

HP1 $\beta$ -Halo(N) cells with normal HP1 $\beta$  expression were labelled with JF<sub>549</sub> and imaged using a laser-scanning confocal microscope (Figure 3.17D, top). A spot in euchromatin was bleached repeatedly, followed by taking a snapshot of the cell (Figure 3.17C). Due to the mobility of chromatin foci and entire cells, it was not possible to gain



**Figure 3.17:** A. An illustration of the preferential internal mixing within eu- (light blue) and heterochromatin (dark blue) (left), as opposed to unrestricted mixing throughout the nucleus (right). B. The experimental design to study HP1 $\beta$  mixing in euchromatin and between eu- and heterochromatin. C. FLIP experiment layout: after the initial imaging of 10 frames, targeted bleaching and data acquisition were alternated, with the cycle repeated 200 times. D. Top, images from an example FLIP sequence for HP1 $\beta$ -Halo labelled with JF<sub>549</sub> in naive mESCs. The location of the bleaching laser is shown in red (B), and the regions where fluorescence intensity was measured are shown in cyan for euchromatin (E), magenta for heterochromatin (H) and green for a control in a neighbouring cell (C). The B, E and C regions were static, while the H regions traced the movement of chromocentres during the acquisition. Scale bar, 5  $\mu$ m. Bottom, fluorescence intensity traces for the regions shown above. The intensity is normalised by dividing by the maximum. 0 s signifies the onset of bleaching (gray dashed line). E. Relative fluorescence intensity traces for eu- and heterochromatin, as well as the bleaching spot and the neighbouring cell control, averaged across multiple cells in the same conditions. The shaded area represents the standard deviation. F. The same traces as in D, but comparing the trends between the differentiation time points.

spatially-resolved diffusion measurements. Instead, for every cell, three euchromatic and three heterochromatic regions were selected (Figure 3.17D). To accurately measure the loss in fluorescence in heterochromatin, the movement of foci was tracked semi-manually. The field of view also always contained a portion of a neighbouring cell, used as a control to ensure that the decrease in intensity was the result of intracellular diffusion.

### **3.3.2 Exchange between eu- and heterochromatin is marginally slowed down**

While the data for individual regions and individual cells was quite heterogeneous (Figure 3.17D, bottom), averaging across many regions yielded much smoother curves (Figure 3.17E). As expected, the control regions were not affected by the bleaching, while the bleached spot showed a rapid decrease in intensity. The loss in fluorescence in other regions was slower, reflecting the fact that this was dependent on protein diffusion.

Mixing of HP1 $\beta$  within euchromatin and between the compartments was not drastically different, especially considering variability in the data. However, in all three states (less so in 2i/LIF), diffusion between the compartments was marginally slower. Thus, there might be a slight difference in mixing within euchromatin and between compartments, potentially becoming more pronounced as mESCs exit from naive pluripotency.

#### **Slower dynamics in formative state**

Comparison between the decay curves for naive-, formative- and primed-state mESCs also revealed that both in eu- and heterochromatin, HP1 $\beta$  diffusion was slower 24h after the onset of differentiation, in agreement with the earlier SPT results (Figure 3.17F).

## 3.4 Diffusion of chromatin-bound molecules and residence time analysis

Short-exposure SPT allows observation of all (or most) of the different populations of protein molecules, whether they are freely diffusing, confined by barriers, or stably bound to chromatin. However, it is not possible to measure the dynamics of the slow-diffusing chromatin-bound populations using this method. Furthermore, it is not easy to distinguish the different populations, and sophisticated analysis approaches such as the 4P algorithm are required for this.

To circumvent these limitations, long-exposure imaging is typically used to investigate the dynamics of chromatin-bound proteins (Etheridge et al. 2014; Basu et al. 2020) (Figure 3.1B). In the formative state, the mESC genome structure decondenses and displays higher intermingling (see section 1.3.2), and the nuclei at this stage display altered mechanical properties (Pagliara et al. 2014). Thus, it is plausible that chromatin might have increased mobility in this state. Consistently, faster histone protein dynamics has been measured by FRAP in NANOG-low cells compared to NANOG-high cells in serum/LIF conditions (Chalut et al. 2012). Since HP1 $\beta$  is distributed throughout the genome, the diffusion properties of its chromatin-bound fraction can be used as a proxy for chromatin movement. In addition, wide-field imaging of HP1 $\beta$  or DNA in parallel with the SPT data acquisition could allow one to further distinguish the processes occurring in eu- and heterochromatin.

Furthermore, since molecules become “invisible” owing to motion blur upon dissociation from chromatin, it is possible to measure the characteristic residence time of proteins on chromatin, using long-exposure SPT (Basu et al. 2020). HP1 $\beta$  localisation changes at the exit from naive pluripotency (see section 2.4.2), and so do the relative fractions of the confined and unconfined molecules (see section 3.2.4). This could be the result of a change in the mode of interaction between HP1 $\beta$  and chromatin, potentially detectable through residence time analysis. In addition, characterisation of the dissociation kinetics of HP1 $\beta$  from chromatin in live mESCs and comparison to the *in vitro* studies may offer new insights into heterochromatin organisation.

### 3.4.1 Method optimisation

#### Experimental design and replicates

To acquire this data, an exposure time of 500 ms was used (Etheridge et al. 2014; Basu et al. 2020). Since HP1 $\beta$  is a very abundant protein, the sptPALM approach was also employed to observe a subset of proteins at any given time. Thus, the labelling strategies developed in previous experiments were also applicable in this case, and HP1 $\beta$  in HP1 $\beta$ -Halo(N) mESCs was labelled with PA-JF<sub>549</sub> or PA-JF<sub>646</sub>.

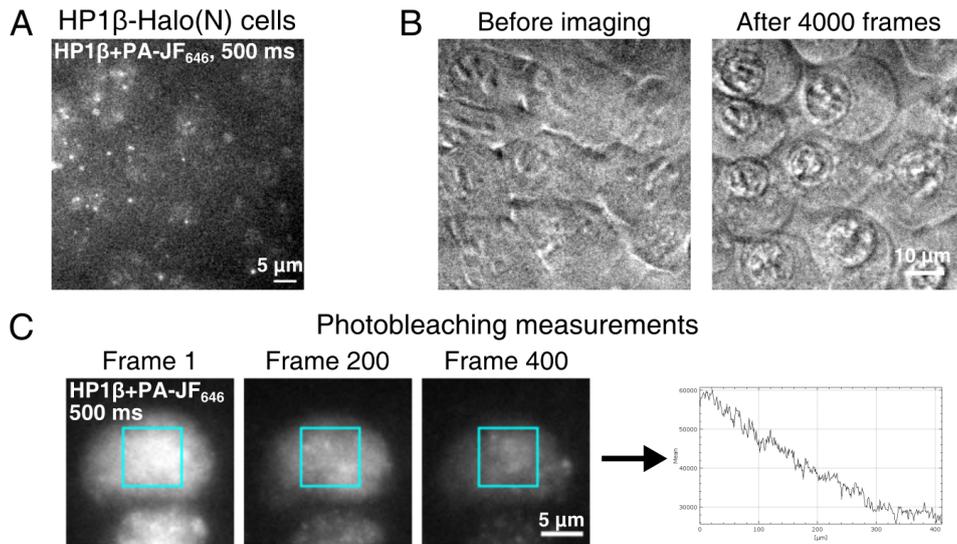
Accurate estimation of the residence time relies on the assumption that molecules disappearing from view dissociate from chromatin. However, two additional processes may lead to the same observation: photobleaching and movement of the molecule out of the focal plane. 3D imaging helps to circumvent the latter issue by having a large depth of focus. One dataset with three biological replicates for each time point was collected using DH-PSF sptPALM. However, the data in different conditions was not recorded on the same day, and thus it may have suffered from day-to-day variation (see section 3.2.1).

Thus, another dataset was acquired with three biological replicates per condition, controlling for the day-to-day variation. In addition, HP1 $\beta$  was imaged in the wide-field mode in parallel to investigate processes occurring in eu- and heterochromatin separately. Due to this, the SPT data was acquired in 2D rather than 3D. Unfortunately, due to technical problems with the 561 nm laser line, the diffraction-limited images were not of good enough quality to segment the heterochromatic foci.

#### Imaging conditions and phototoxicity

Longer exposure time meant that, compared to fast imaging, lower excitation and activation laser intensities were required. At the 2 fps rate, both photoactivatable dyes displayed sufficient autoactivation to enable imaging without the 405 nm activation laser (Figure 3.18A). To minimise photobleaching, the excitation laser intensities were adjusted to the minimum required to localise the molecules.

The absence of the UV allowed imaging live cells for longer periods of time due to



**Figure 3.18:** A. An example frame from 2D sptPALM 500 ms data for HP1 $\beta$ . B. Cells round up and detach from the dish after 4000 frames (33 min) of long-exposure imaging with 561 or 640 nm laser. C. Bulk loss in intensity of PA-JF<sub>549</sub> or PA-JF<sub>646</sub> under the long-exposure imaging conditions prior to the single-molecule regime was used to measure the bleaching rate. Shown is one example of a region selected for measurement (cyan square) and the corresponding intensity decay curve.

lower phototoxicity. Cells started displaying clear signs of apoptosis after 4000 frames (33 min) of imaging (Figure 3.18B). Thus, 1000 frames (8.5 min) were acquired per position, since cells' appearance remained normal during this time, when illuminated with much higher intensities in a previous experiment (Figure 3.3). However, it cannot be excluded that laser exposure might have perturbed cellular processes even at this time scale. 10-12 positions were collected per replicate.

### Bleaching correction

Two approaches exist to correct the dissociation constant estimates for the effects of photobleaching. The first one involves explicitly measuring the bleaching rate constant and including it into the model fitted to the SPT data. This method is simple, but it assumes that the bleaching constant is invariable between conditions. The second possibility is time-lapse imaging with constant exposure time, but variable interval (Gebhardt et al. 2013). This approach is in principle more reliable, however, collecting the timelapse data takes a long time, and thus may suffer from technical variability. Furthermore, fitting a complex model to the data may not always be possible.

In this work, I attempted using both methods. However, fitting the timelapse data using the procedure described in Gebhardt et al. (2013) was unsuccessful. Thus, the bleaching rate constant was estimated independently by illuminating labelled cells with laser intensities used for SPT (Figure 3.18C). The loss in bulk fluorescence signal (before achieving the single-molecule regime) was then measured across several areas for 400 frames, and the bleaching rate was determined by fitting to a single exponential.

## Data analysis

The pre-processing and diffusion analysis of the long-exposure SPT data was performed similarly to the short-exposure SPT. The number of localisations and trajectories collected is summarised in Table 3.2. All trajectories (>2 points) were used in residence time calculations, while long trajectories were used for diffusion analysis.

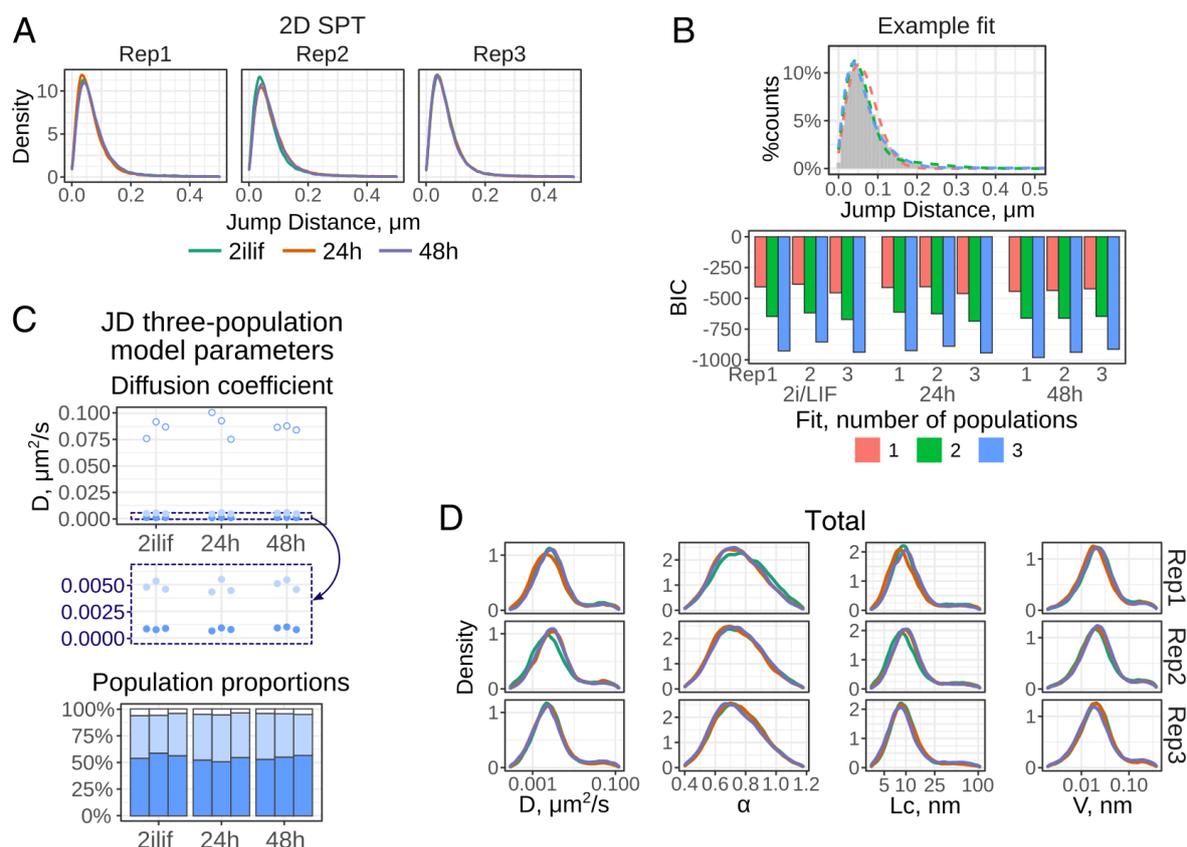
**Table 3.2:** Number of localisations and trajectories in each dataset: all localisations, all trajectories >2 points, trajectories >10 points and trajectories >10 points after quality filtering.

Dataset	All loc	All traj	Traj>10	Traj>10, filt
2ilif_1	79712	904	855	146
2ilif_2	124063	15021	1386	253
2ilif_3	140711	1203	1080	204
24h_1	49128	583	533	103
24h_2	87927	1056	822	148
24h_3	129073	1152	1010	216
48h_1	96632	1561	1193	216
48h_2	121892	1464	1303	206
48h_3	102313	1107	989	188

### 3.4.2 Dynamics of HP1 $\beta$ -bound chromatin does not change upon differentiation

To compare the dynamics of HP1 $\beta$ -bound chromatin upon exit of mESCs from naive pluripotency, the jump distance analysis was first conducted. In the first 3D SPT dataset, the replicates showed inconsistent results. The second 2D SPT dataset indicated that there were no reproducible differences in HP1 $\beta$ -bound chromatin movement

between the three differentiation time points (Figure 3.19A). The three-population diffusion equation fitted the data the best (Figure 3.19B). However, the fastest population had a very low abundance and displayed much faster diffusion than the other two, fitting the long tail of the distribution and potentially corresponding to the rare erroneous joining of trajectories of different molecules (Figure 3.19C).



**Figure 3.19:** A. Displacement curves for HP1 $\beta$  long-exposure imaging in mESCs and upon removal of 2i. B. An example a jump distance histogram (48h rep1) with the one-, two- and three-component diffusion model predictions overlaid (top). The BIC values for the different models fitted to the HP1 $\beta$  SPT data (bottom). C. The apparent diffusion coefficient  $D$  and the proportions of slow (blue), fast (light blue) and very fast (white) populations of HP1 $\beta$ , predicted by the three-component diffusion model. Each column corresponds to a replicate. The zoom-in panel shows the  $D$  of the two slow populations on a different scale. Differences in the parameters between the conditions are not significant (repeated-measures ANOVA). D. Distributions of diffusion parameters, calculated by the 4P algorithm for the 2D SPT dataset.  $\alpha$  is shown on a linear scale, while the other parameters are shown on a log scale.

Next, the four diffusion parameters were calculated, using the 4P approach (Figure 3.19D). Unlike for the short-exposure data, the feature distributions were smooth and had a clear single peak.  $D$ ,  $Lc$  and drift had long tails, which potentially corresponded

to the erroneous pasting of trajectories and were identified as “unconfined” by the 4P classification. Thus, HP1 $\beta$ -bound chromatin appeared to display a single mode of diffusion, or multiple very similar ones, as indicated by the jump distance analysis.

The anomalous exponent displayed by chromatin was surprisingly large, with a median of 0.7-0.8, while it has previously been estimated to be in the range of 0.4-0.6 (Shukron et al. 2019). This could be explained by the nuclear movement dominating over local dynamics at 2 fps repetition rate (Shukron et al. 2019). Indeed, coordinated movement of molecules in the same direction as the entire nucleus could be observed in the raw data.

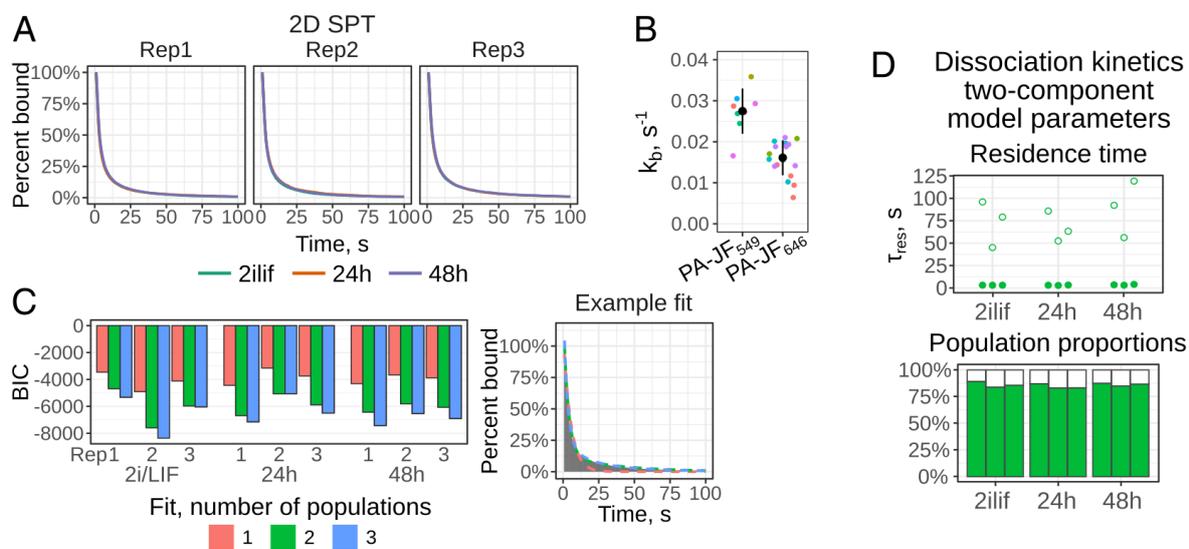
No consistent differences were found between the pluripotency states (Figure 3.19A,C,D). This could either indicate a lack of change in chromatin dynamics upon exit from naive pluripotency, or that the values of diffusion parameters were predominantly determined by the large-scale nuclear movement.

### 3.4.3 HP1 $\beta$ has at least two modes of chromatin binding

Finally, the residence time of HP1 $\beta$  on chromatin was estimated, using the collected long-exposure data. Plotting the decay curves for each differentiation time point suggested that there was no change in HP1 $\beta$  dissociation kinetics upon exit of mESCs from naive pluripotency (Figure 3.20A).

The bleaching constant for PA-JF<sub>646</sub> under the 2D sptPALM conditions was estimated to be  $k_b = 0.016\text{s}^{-1}$ , corresponding to a characteristic bleaching time of 62.5 s (Figure 3.20B). Interestingly, a stronger bleaching constant for PA-JF<sub>549</sub> was measured under the 3D sptPALM conditions, consistent with the shorter uncorrected residence times measured in those experiments (not shown). This likely reflected the fact that DH-PSF microscopy required higher excitation laser intensities. Furthermore, this underscored the importance of accounting for bleaching in residence time measurements.

Fitting of the decay curves to the appropriate models revealed at least two components in the dissociation kinetics of HP1 $\beta$  from chromatin (Figure 3.20C). Although the BIC values for the three-component fit were lower for some datasets, the difference



**Figure 3.20:** A. Decay curves, showing the percentage of HP1 $\beta$  molecules visible (bound to chromatin) in the long-exposure SPT experiment for at least the given amount of time. B. Bleaching constants  $k_b$  for PA-JF<sub>549</sub> and PA-JF<sub>646</sub>, measured by bulk bleaching in HP1 $\beta$ -Halo(N) live cells under conditions similar to the SPT experiment conditions. Each point corresponds to a different region of the image sequence, and different fields of view are shown in different colours. The mean  $\pm$  standard deviation is shown in black. C. The Bayesian information criterion values for one-, two- and three-component model fits to the HP1 $\beta$  2D SPT residence time decay curves (left), and an example fitting plotted over the data (48h, replicate 1; right). D. The characteristic bleaching-corrected residence times and the relative percentages of the binding modes between HP1 $\beta$  and chromatin from the two-component fitting.

was small, and thus a simpler model was preferred. Taking bleaching into account, the major process seemed to have a characteristic residence time of 2.8-4.0 s (Figure 3.20D). A much longer residence time ( $\tau_{\text{res}} = 45\text{-}120$  s) was characteristic for  $\sim 11\text{-}17\%$  of molecules, but it could not be measured accurately due to photobleaching.

### 3.5 Discussion

In this chapter, the dynamics of HP1 $\beta$  in mESCs and upon exit from naive pluripotency were characterised. Single-molecule microscopy approaches were used extensively, enabling one to:

- study different types of molecules – such as chromatin-bound and unbound molecules – separately in the same cells;
- explore distributions of various diffusion parameters, and take into account sub- or superdiffusive behaviour;
- study the spatial variation in diffusion, and link it to the local density of DNA or proteins;
- measure residence times of molecules on chromatin in live cells.

However, this freedom comes at a cost of complicated setup and therefore sensitivity to technical variation. Careful experimental design is thus required to produce comparable results and draw robust conclusions. Furthermore, a variety of non-biological factors, such as localisation precision, time resolution and bleaching, must be taken into account to correctly interpret the data.

A variety of analysis techniques can be applied to harness the information hidden in the data. In this project, conventional jump distance analysis was compared to the novel 4P algorithm, which makes good use of the single-molecule nature of the data (Shukron et al. 2019; Basu et al. 2020). Although any differences between samples or conditions should manifest themselves in the displacement curves, the 4P analysis allows to investigate these differences in much more detail. For example, changes in

confinement can be distinguished from changes in diffusion coefficient, as was demonstrated here for HP1 $\beta$  in heterochromatin. In addition, the classification functionality of the 4P method can be used to separate the chromatin from free protein movement, which is crucial to study the LLPS hypothesis for HP1 $\beta$  and distinguish it from chromatin phase separation.

Furthermore, additional analysis was performed in this work, inspired by the predictions of the LLPS model for heterochromatin segregation. This was particularly informative due to the ability to distinguish eu- and heterochromatin, using dual-colour imaging. Calculating the angle between the consecutive steps in trajectories uncovered a link between the tendency for motion reversal and the concentration of DNA in the area. Potential non-equilibrium processes were revealed by studying molecules crossing the compartment boundary. On the other hand, some analysis, such as searching for increase in drift at phase boundaries, was inconclusive, potentially requiring a different type of experiment.

HP1 $\beta$  diffusion was also investigated in bulk using FLIP, and the exchange of HP1 $\beta$  between eu- and heterochromatin was found to be marginally slower than mixing within euchromatin. Interestingly, the conventional processing of this experiment was not possible due to movement of cells and especially movement of chromatin within cells.

SPT at a different time scale was employed to study the chromatin-bound fraction of HP1 $\beta$ . Given the substantial increase in chromosome intermingling and breakdown of compartments observed in the formative state (see section 1.3.2), and the reported links between chromatin structure and dynamics (Basu et al. 2020; Shaban et al. 2020; Barth et al. 2020), it was surprising that no change in chromatin movement was found upon exit of mESCs from naive pluripotency. Furthermore, the dissociation kinetics of HP1 $\beta$  from chromatin could also be deduced from this data, and was also found to stay constant upon differentiation.

In my view, the critical next steps required to interpret the experimental results in the context of the models of heterochromatin organisation is development of appropriate biophysical models. These models should be realistic, so that assumptions made do not contradict what is expected to happen in cells. Moreover, they should generate un-

ambiguous predictions for experimentally measurable parameters, taking into account the caveats and biases present in the data. For example, a fraction of HP1 clearly binds chromatin stably, and LLPS might occur on top of that through multivalent transient interactions. However, models describing such a system quantitatively and in the context of observables from SPT data are not available. Furthermore, basic features of chromatin phase separation and its effects on movement of chromatin-binding proteins are also poorly understood. Furthermore, better ways of accounting for localisation imprecision in single-molecule experiments are needed.

### **3.5.1 Models of heterochromatin organisation**

Two competing models explaining the enrichment of HP1 and other proteins in heterochromatin exist: the LLPS model (Larson and Narlikar 2018; Strom et al. 2017) and the crosslinked chromatin globule model (Erdel et al. 2020). The latter suggests that HP1 foci are formed by a locally elevated density of binding sites on a collapsed chromatin globule, which is either crosslinked by an external factor (potentially HP1 itself) or represents a fundamental conformation of chromatin fiber itself. The LLPS model, on the other hand, proposes that enrichment of HP1 and other heterochromatic proteins, exclusion of other factors and potentially chromatin condensation are mediated by a network of transient multivalent interactions between HP1, histone tails and DNA, forming a semi-permeable barrier. It is possible that both of these models are applicable to a certain extent, and whether that is the case is the subject of active research. In this section, I will attempt to bring my findings into the context of these hypotheses.

#### **Unconfined HP1 $\beta$ is retained in heterochromatin**

The LLPS model predicts a higher concentration of freely diffusing HP1 $\beta$  in heterochromatic foci, while in the absence of phase separation and in the limit of fast diffusion, the concentration should be uniform throughout the nucleoplasm. The unconfined fraction measurements, combined with the estimation of HP1 $\beta$  concentration inside

and outside of foci, suggested that there was about 16% more unbound protein within heterochromatic domains (Table 3.3), seemingly supporting the LLPS model.

**Table 3.3:** The ratio of total HP1 $\beta$  concentration in hetero- and euchromatin, calculated using the wide-field HP1 $\beta$  images from dual-colour SPT data. The indicated values are mean $\pm$ SD of values for individual cells. The ratio of unconfined HP1 $\beta$  concentrations in the two compartments was estimated by multiplying the total ratios by the ratio of the unconfined fractions (%u) in the two compartments.

	$[\text{HP1}\beta]_{\text{Het}}/[\text{HP1}\beta]_{\text{Eu}}$	$\%u_{\text{Het}}/\%u_{\text{Eu}}$	$[\text{HP1}\beta]_{\text{u, Het}}/[\text{HP1}\beta]_{\text{u, Eu}}$
<b>2i/LIF</b>	1.34 $\pm$ 0.17	0.88	1.17
<b>24h</b>	1.35 $\pm$ 0.11	0.86	1.16
<b>48h</b>	1.39 $\pm$ 0.15	0.83	1.16

However, if HP1 $\beta$  diffusion is not fast enough to immediately equilibrate its concentration across the nucleus, a gradient could be expected in the vicinity of a cross-linked chromatin globule with a high concentration of HP1 $\beta$  binding sites. A theoretical description of such a scenario is needed to find the expected concentration field at given binding and diffusion rates, and to compare these with the LLPS hypothesis predictions and with the experimental results.

Although the concentration estimates based on brightness in wide-field images are imprecise, they are supported by more accurate measurements from the confocal images, which suggest an even greater enrichment of HP1 $\beta$  in foci. In addition, the proportion of the unconfined molecules could be underestimated due to diffusion out of the focal plane or motion blur (Hansen et al. 2018). Although methods to account for the former in jump distance analysis exist, further developments are needed to apply them to the 4P algorithm. Alternatively, a method combining the 3D DH-PSF live tracking with volumetric imaging of foci could be developed.

### **HP1 $\beta$ movement in heterochromatin is confined, but fast**

The unconfined HP1 $\beta$  molecules in heterochromatin display subdiffusive motion with lower anomalous exponent and radius of confinement, indicative of a different environment. However, unexpectedly, the diffusion rates of the unbound molecules in heterochromatin were similar or even slightly higher than in euchromatin.

Speculatively, this could be an argument against the LLPS hypothesis, as suggested

by Erdel et al. (2020). Crosslinked heterochromatic nucleosome arrays of MDa size can be viewed as large obstacles for the small 50 kDa HP1 $\beta$  dimer, and thus a higher confinement (lower  $\alpha$ ) could be expected in either model. Formation of transient multivalent interactions, as proposed in the LLPS hypothesis, might be expected to increase the effective viscosity and reduce the diffusion coefficient. In contrast, if the heterochromatic globules are percolated by the normal nucleoplasm, in the absence of LLPS the movement of HP1 $\beta$  in between the obstacles may be similar to that in euchromatin. Analytical models and/or molecular dynamics simulations are needed to explore the different features of diffusion that would be characteristic for either scenario.

On the other hand, crowding in heterochromatin slows down the movement of inert probes (Bancaud et al. 2009), and the absence of this with HP1 $\beta$  is hard to explain by the crosslinked globule model. However, since those experiments were conducted in other cell types, it would be beneficial to repeat them in the mESC system.

### **Relationship between HP1 $\beta$ confinement and chromatin density needs to be studied further**

The spatial correlation analysis provided a segmentation-independent verification of the differences between eu- and heterochromatin. As expected, the proportion of confined HP1 $\beta$  molecules correlated with the DNA or HP1 $\beta$  staining intensity. The motion reversal bias for unbound HP1 $\beta$  also correlated with chromatin density, suggesting stronger confinement within heterochromatin.

However, more can be learned from examining this relationship further. The simple LLPS model predicts that motion reversal bias should be uniform within the condensate, and stronger at the border due to molecules “bouncing off”. This effect was not observed in the comparison of the angle between trajectory steps on the border of the foci and the interior; however, this result may be influenced by imprecise border definition. Furthermore, heterochromatic foci might consist of multiple phases, not distinguishable by wide-field imaging. On the other hand, for the crosslinked globule model, one might expect the level of confinement to scale with heterochromatin density. The scatter plots of diffusion parameters vs DNA density seem to suggest that

that is not the case, and that there might be two types of regions with little scaling between chromatin density and confinement. Better-resolution DNA/HP1 $\beta$  images are however needed to define the chromocentre borders more precisely and to investigate these relationships further. In addition, this could allow one to test whether there is a linear relationship between the apparent diffusion coefficient and the concentration of HP1 $\beta$  molecules, which is characteristic of the crosslinked globule model (Heltberg et al. 2021).

### **Heterochromatin in mESCs might be in non-equilibrium state**

Comparison of displacement lengths and frequencies of molecules crossing the boundary from eu- to heterochromatin and the other way round suggested that the foci may not be in equilibrium. Perhaps the most surprising aspect of this finding was that the flux of molecules appeared to go out of the heterochromatic regions, meaning that the foci were slowly shrinking rather than growing. Speculatively, this could be a manifestation of the Ostwald ripening effect, where the small phase-separated droplets shrink over time as the solute transitions through the dilute phase into the large droplets.

As with other measurements, likely due to the subtleness of the phenomena, but also the imprecision in boundary definition, the confidence in these results is not high. Since differences between entering and exiting molecules scale with the frame rate (Bo et al. 2021), one potential way observe these effects more clearly would be to introduce time intervals between frames. Such a modification should not affect the populations of molecules detected, because the extent of motion blur will remain the same.

### **Stable interactions between HP1 $\beta$ and chromatin**

The long-exposure SPT experiments indicated that HP1 $\beta$  has at least two characteristic residence times on chromatin: 83-89% of molecules were bound for 2.8-4.0 s, and the remaining for >45 s. These values are substantially longer than those reported in the *in vitro* binding experiments between HP1 $\beta$  and H3K9me3-modified nucleosome arrays, where 95% of molecules had a characteristic residence time of 0.19 s, and the rest about

3.6 s (Kilic et al. 2015; Bryan et al. 2017). However, measurements by other techniques have also previously revealed that stably chromatin-bound HP1 populations exist in cells. In NIH 3T3 cells, 74% of total HP1 $\beta$  was found to be bound to chromatin with a characteristic residence time of 2.5 s, and 8% was bound for >60 s (Müller et al. 2009). Other studies in various cell types also discovered a fraction (5-30%) of immobile HP1 $\beta$  with residence time of >120 s (Schmiedeberg et al. 2004; Festenstein et al. 2003).

These differences could be explained by the interaction being stabilised in cells by additional factors, such as proteins binding the HP1 $\beta$  CSD, or RNA. In this context, the two (or more) populations with different dissociation kinetics could represent various HP1-containing complexes. Measuring the residence times of HP1 point mutants, or wild-type HP1 in the context of point mutations of its interactors, could provide some insight, although side effects of the mutations must be carefully considered. Alternatively, HP1 binding stability could depend on the environment, e.g. the density of H3K9me2/3 or potentially H3K23me1/2/3, H1.4K26me2 or H4K20me3, as well as the accessibility of histone tails.

The difference could also arise from the different techniques used. The *in vitro* studies acquired data at 20 fps, which was possible due to the spaced-out immobilised chromatin fibers and fast unbound HP1 diffusion (Kilic et al. 2015; Bryan et al. 2017). The current experiments were conducted at 2 fps to ensure blurring of the freely diffusing molecules in cells, and thus transient binding events may have been missed. Thus, obtaining *in vitro* data with longer exposure time, as well as *in vivo* data with shorter exposure time and potentially applying the 4P classifier to extract the chromatin-bound HP1 fraction is needed to directly compare these measurements.

The presence of very long-lived interactions between HP1 $\beta$  and chromatin once again underscores that the pure LLPS model for heterochromatin cannot be accurate and must be adjusted to include stable binding.

### **Preferential internal mixing**

The FLIP experiments conducted for HP1 $\beta$  suggested that a subtle difference might exist between mixing within euchromatin and that between eu- and heterochromatin.

The presence of a weak boundary for HP1 $\beta$  diffusion might be one potential explanation for this result. However, it could also be caused by the increased percentage of chromatin-bound HP1 $\beta$  proteins in heterochromatin. Unfortunately, the bulk nature of FLIP measurements does not allow one to distinguish between these hypotheses.

In addition, the energetic barrier should affect proteins within the dense phase rather than those outside. This can be understood better by considering that at equilibrium, an equal number of molecules transition between the phases in both directions. Due to a higher concentration in the dense phase, more molecules arrive to the inner side of the boundary, which “must” turn back to maintain the equilibrium. Thus, experiments where a spot within heterochromatin is bleached and the loss in fluorescence in another region of the focus and in euchromatin is compared need to be carried out. This is technically challenging due to the fast movement of chromocentres in mESCs. To achieve this, either the bleaching laser must somehow follow the chromocentres, or the data must be collected much faster.

### **Limitations and future work**

Overall, the single-molecules measurements described in this chapter provided insights into HP1 $\beta$  dynamics in heterochromatin, but were hard to interpret due to the limitations of the technique, the complexity of the system, the subtlety of the effects, and the lack of clear predictions informed by appropriate biophysical models.

The first confounding factor was the low resolution of the wide-field imaging of DNA or HP1 $\beta$  density and, as a result, poor definition of heterochromatin borders. This made the measurements of any parameters across or in the vicinity of the compartment boundary – displacement length, average drift vector, angle distribution – imprecise, and prevented drawing robust conclusions. This issue could be addressed by combining SPT with a higher-resolution imaging technique, such as light-sheet or super-resolution radial fluctuations (SRRF) microscopy.

Furthermore, the type of SPT experiment employed here, aiming to produce long trajectories and detect changes in molecular behaviour, may not be the optimum strategy in all cases. In particular, high-resolution spatial mapping of the average drift vec-

tor requires a high number of measurements from independent molecules in the area, rather than a few long trajectories. The SPT experiment must therefore be modified. Higher excitation laser intensities and/or less photostable fluorophores, combined with stronger activation, could be employed to achieve faster bleaching and higher data density in sptPALM. Alternatively, spontaneously blinking dyes, such as HM-JF<sub>526</sub> (Zheng et al. 2019), could be used to avoid the damage from the UV exposure. In addition, SPT variations, such as stroboscopic photoactivation SPT (spaSPT) or single-molecule displacement mapping (SMdM), could be employed to collect more data faster and with better resolution for map construction (Hansen et al. 2018; Xiang et al. 2020). Furthermore, this could even alleviate the need for precise border definition, since the density of super-resolution data could be sufficient to observe local effects at individual foci. However, these techniques do not allow one to distinguish between the chromatin-bound and freely diffusing molecules, which is the strong side of the current approach. A crude classification could be achieved by only considering displacements longer than a threshold value. Alternatively, simulations could be used to generate predictions for a combined LLPS and stable binding model, which could be tested without classification.

Another important factor that must be taken into account is localisation precision. For example, for a particle whose movement at a given repetition rate is on a finer scale than localisation precision, the angle analysis will display an enrichment of apparent motion reversal. Thus, the expected distribution even for Brownian motion may not be totally symmetric. Although corrections for diffusion coefficient estimation are known, and some limitations can be deduced from immobile controls, the effect of localisation precision on other measurements – such as the angle between trajectory steps – must be studied further.

Matching quantitative predictions from theory and the experimental results may be challenging due to the relative nature of the latter. In particular, as described in section 3.1.2, the exposure time used in the experiment affects the outcome, and measurements performed with different repetition rates are not easily comparable. A suite of experiments with a range of exposure times and intervals between frames could potentially be used to probe the entire range of molecular movement and attempt obtaining ab-

solute values for diffusion and kinetic parameters. Alternatively, theoretical models or simulations could incorporate the effects of the exposure time, and/or generate relative predictions.

### 3.5.2 Differentiation

In the previous chapter, HP1 $\beta$  levels were found to increase in formative mESCs. Moreover, HP1 $\beta$  was “released” from foci at this stage, and was concentrated back later. Finally, DNA and HP1 $\beta$  densities were better correlated in the primed state, indicating a relocalisation of HP1 $\beta$ .

The SPT and FLIP data revealed that globally, HP1 $\beta$  diffuses slightly slower in the formative state. In addition, despite a marked change in nuclear architecture, the dynamics of HP1 $\beta$ -bound loci did not change, at least at the scale sampled in this work. The residence time of stably bound HP1 $\beta$  on chromatin was also unchanged. In this section, I will bring these observations together to better understand heterochromatin reorganisation during exit from naive pluripotency.

#### Viscosity of formative mESC nuclei might be increased

There are two potential reasons why a protein’s diffusion could slow down: 1) a larger radius of gyration, either due to a conformational change or binding to another protein; 2) an effective increase in the viscosity of the environment, e.g. due to longer transient interactions between the protein and components of the environment. The latter effect could be general or specific to the protein of interest.

The change in HP1 $\beta$  diffusion observed in the formative state is very subtle (e.g. the median  $D$  of the unconfined population decreases by only 6%). Since HP1 $\beta$  is a small protein (around 50 kDa as a dimer), binding of HP1 to another protein would substantially increase the mass of the former. The diffusion coefficient of a particle is inversely proportional to its radius of gyration, which in its turn roughly scales with the cubic root of the mass. Therefore, binding of an HP1 $\beta$  dimer to e.g. Dnmt3a (130 kDa), which increases in expression in the formative state, would be expected to

decrease the diffusion coefficient of HP1 $\beta$   $\sim$ 1.53 times. Formation of a complex with a single Suv39h1 protein (48 kDa) should bring the diffusion rate down 1.25 times, and a similar difference would be expected between HP1 monomer and dimer. Thus, the magnitude of the effect does not support the hypothesis that HP1 $\beta$  might move slower in formative-state mESCs due to forming a complex with another protein, or due to an enhanced tendency to dimerise. A conformational change, e.g. due to a PTM, is a more plausible explanation in this context. This could also potentially be responsible for HP1 $\beta$  redistribution in the formative state.

The transient slow-down in HP1 $\beta$  diffusion coincides with large-scale changes in chromatin structure and nuclear mechanical properties (Pagliara et al. 2014). Thus, it is tempting to speculate that the relaxed chromatin conformation in the formative state might cause the change in HP1 $\beta$  dynamics. Indeed, the decondensed chromatin fiber might occupy a larger volume, and the probability of HP1 $\beta$  colliding with it may increase. Furthermore, “sticky” regions of chromatin might become exposed upon relaxation and contribute to stronger transient interactions with HP1 $\beta$ , especially with its unstructured regions.

If this hypothesis is correct, a similar phenomenon should be observed for other proteins, although it might be more or less pronounced depending on their propensity to interact with chromatin non-specifically. To check this, diffusion of other proteins must be investigated using SPT or other methods. It would be interesting to compare other chromatin-binding proteins, e.g. components of PRC1/2 or transcription factors, with an inert probe such as HaloTag or GFP. On the other hand, to prove a causal link, dynamics of proteins could be measured in mESCs with artificially decondensed chromatin. Interestingly, FRAP experiments in CHO cells indicated an increase in HP1 mobility upon artificial chromatin decondensation, contrary to the mechanism discussed here (Cheutin et al. 2003). However, this could be the result of an increase in chromatin-unbound fraction of HP1.

## Confined HP1 $\beta$ might interact with chromatin non-specifically

For simple bimolecular reactions, upon increasing the concentration of a component, the proportion of that component not bound to the other component must also increase. The SPT results in this work seem to contradict this rule: although levels of HP1 $\beta$  rise upon differentiation, and the concentration of its classic binding site H3K9me3 drops, the proportion of unconfined HP1 $\beta$  does not increase.

This can be rationalised in the case when the concentration of one component is much larger than that of the other:

$$K_d = \frac{[A]_{\text{free}}[B]_{\text{free}}}{[AB]} = \frac{[A]_{\text{free}}([B]_{\text{total}} - [AB])}{[AB]} \stackrel{[B]_{\text{total}} \gg [AB]}{\approx} \frac{[A]_{\text{free}}[B]_{\text{total}}}{[AB]} = \frac{[A]_{\text{free}}[B]_{\text{total}}}{[A]_{\text{total}} - [A]_{\text{free}}}, \quad (3.6)$$

$$\text{then } \frac{[A]_{\text{free}}}{[A]_{\text{total}}} \approx \frac{K_d}{K_d + [B]_{\text{total}}} = \text{const}. \quad (3.7)$$

Here,  $A$  is HP1 $\beta$  and  $B$  is its binding site on chromatin. To have a substantial proportion of free molecules, the affinity between the two components must be rather weak, with  $K_d$  in a range similar to  $[B]_{\text{total}}$  – e.g. to reach 55%, as seen in euchromatin of mESCs in 2i/LIF,  $K_d = 1.22[B]_{\text{total}}$ . Given that 1) H3K9me3-HP1 $\beta$  interaction is not weak ( $K_d \approx 2\text{-}6 \mu\text{M}$ , further strengthened by the chromatin context – see section 1.4.1), and 2) the concentration of H3K9me3 drops in formative state, while the proportion of free HP1 $\beta$  does not increase, it appears more likely that the confinement captured by short-exposure SPT is primarily caused by non-specific binding of HP1 $\beta$  to the chromatin fiber – potentially to the H3 tails, DNA, histone H1, etc. This would tie the slight increase in the bound fraction of HP1 $\beta$  in formative mESCs to chromatin decondensation. Furthermore, the same transient interactions happening on the timescale  $< 5$  ms might result in the apparent slow-down in HP1 $\beta$  diffusion. In support of these calculations, the observed difference in the proportion of confined HP1 $\beta$  between eu- and heterochromatin corresponds to slightly more than two-fold difference in chromatin density, in agreement with the confocal microscopy measurements.

Other explanations are also possible, such as that the number of HP1 $\beta$  binding sites

other than H3K9me2/3 mark increases proportionally with HP1 $\beta$ . However, such an exact match appears less likely than the simple interpretation of confined molecules as unstably bound to chromatin. To distinguish between these possibilities, one could potentially investigate the effect of removal of histone marks that HP1 $\beta$  binds to on the percentage of confined molecules by knocking out the respective enzymes.

Overall, it appears that sptPALM with 5 ms exposure allows distinguishing between truly unbound HP1 $\beta$  and molecules transiently interacting with chromatin, which was not possible before (Müller et al. 2009). On the other hand, one might be more interested in detecting the specific binding events, which would require measurements at a different time scale. Since it is undesirable to prolong the exposure time due to motion blurring, one could adopt a time-lapse approach with varying interval lengths between 5 ms (or shorter) frames. Combined with the 4P algorithm, this would allow probing the dynamics of HP1 $\beta$  chromatin binding with varying stability. This could be especially informative to understand the mechanism of HP1 $\beta$  redistribution in formative-state mESCs.

### **Factors other than H3K9me3 might be important for stable HP1 $\beta$ binding**

*In vitro* experiments have demonstrated that in a simple system of purified nucleosome arrays, HP1 residence time on chromatin depends very strongly on the density of H3K9me2/3 modifications due to divalency of the HP1 dimer (Kilic et al. 2015). In addition, due to facilitated dissociation caused by competition, HP1 residence time is expected to decrease as its concentration increases. However, the stable HP1 $\beta$  residence time on chromatin was unaffected by a two-fold drop in H3K9me3 levels in pericentromeric heterochromatin and a rise in concentration at the 24h time point. Potentially, H3K9me2/3 density and/or extent of chromatin compaction in mESCs is not sufficient to observe binding via two chromodomains. However, this should lead to much shorter residence times of HP1 $\beta$  on chromatin in mESCs than in other cell types, which was not the case (Müller et al. 2009). Alternatively, other direct or indirect interactions between HP1 $\beta$  and chromatin might define its dwell time, as suggested previously (Thiru et al. 2004).

A lack of change in stable binding is also puzzling in the context of HP1 $\beta$  relocalisation. The hypotheses for its mechanism, outlined in the previous chapter, broadly fall into two categories: altered HP1 $\beta$ -chromatin interaction (due to a change in histone marks, RNA, protein partners or PTMs of HP1 $\beta$ ), or biophysical consequences of the increased HP1 $\beta$  concentration. Although the former appeared more plausible due to the lack of HP1 $\alpha$  redistribution, unchanged dwell time on chromatin potentially argues in favour of the LLPS model. However, it is possible that interactions of intermediate duration – i.e. more stable than those probed by short-exposure imaging, but less stable than those probed at 2 fps – might be the key. Furthermore, measuring the residence time separately in eu- and heterochromatin might reveal compartment-specific processes.

### **Chromatin dynamics**

No change was observed in HP1 $\beta$ -bound chromatin diffusion properties upon exit of mESCs from ground pluripotency. This could mean that a) the chromatin reorganisation observed by Hi-C is a slow process, not involving local chromatin dynamics, or b) 500 ms is the wrong time scale to observe the potential effects due to it being dominated by large-scale movements of chromatin and cells. The latter hypothesis could be addressed by using a shorter exposure time, e.g. 100 ms, to image HP1 $\beta$ . It is unclear whether the motion blurring would be sufficient to only observe chromatin-bound molecules in these conditions, but they could potentially be identified, using the 4P algorithm. Alternatively, endogenous proteins predominantly bound to chromatin, e.g. histones, a stably binding probe, such as dCas9 (Gu et al. 2018), or fluorescent nucleotides incorporated into DNA during a short pulse in S phase (Strickfaden et al. 2020) could be employed to investigate chromatin movement directly.

### **3.5.3 Conclusion**

In summary, the work described in these two chapters revealed several interesting features of heterochromatin and, more specifically, HP1 $\beta$  in mESCs and upon their

exit from naive pluripotency. Due to the complexity of the problem, imperfection of methods, lack of quantitative biophysical models and sensitivity of the system to perturbations, it is hard to interpret the observations unambiguously and propose a coherent model. Below, several non-exclusive hypotheses are outlined, together with the evidence supporting or contradicting them.

1. HP1-mediated LLPS contributes to the formation and segregation of heterochromatin. HP1 $\alpha$  is the primary phase-separating component, while HP1 $\beta$  cosegregates with it. Increase in HP1 $\beta$  concentration in the formative state destabilises the droplets, causing redistribution of HP1 $\beta$ .

- The observations potentially supporting the LLPS model include the retention of unbound HP1 $\beta$  in heterochromatin, the stronger confinement of HP1 $\beta$  there, and the slightly impeded mixing between eu- and heterochromatin.
- However, none of the effects expected for LLPS due to “bouncing off” of molecules were observed at the foci boundaries.
- The interpretation of the lack of a difference in diffusion coefficients between the compartments is controversial.
- Development of suitable biophysical models, as well as improvement of boundary definition and designing experiments better suited for measurement of certain parameters, e.g. the boundary effects, is necessary to make more robust conclusions.
- The absence of a change in stable HP1 $\beta$ -chromatin interactions suggests that they are not responsible for HP1 $\beta$  redistribution in the formative state, indirectly supporting the LLPS dissolution hypothesis. However, it is then unclear why HP1 $\alpha$  localisation is unaffected, and why HP1 $\beta$  concentrates in the DNA-dense foci at the 48h timepoint.
- Puncta of HP1 $\beta$  can arise without an increase in DNA density, indicating that formation of HP1 condensates might be possible in the absence of chromatin condensation. Possibly, such puncta might nucleate heterochromatin

formation, or fuse with the large heterochromatic domains.

2. Heterochromatin is a cross-linked globule, and HP1 $\beta$  is enriched there due to a higher density of its binding sites. HP1 $\beta$  redistribution in formative-state mESCs is caused by a change in its interaction with chromatin.

- Fast movement of unconfined HP1 $\beta$  in heterochromatin might suggest that HP1 $\beta$  is not slowed-down by self-interactions, and that heterochromatin is percolated by the same nucleoplasm as euchromatin (Erdel et al. 2020). However, then it remains to be explained how potential exclusion of other proteins and transcriptional silencing are achieved in such an accessible environment.
- Although some of the arguments outlined above favour the LLPS hypothesis, they might also be compatible with the crosslinked globule model. For example, retention of free HP1 $\beta$  in heterochromatin might also be explained by fast transient binding and relatively slow diffusion (Heltberg et al. 2021). Theoretical studies are needed to investigate the different scenarios.
- The absence of boundary effects supports the cross-linked globule model. However, further studies, both experimental and theoretical, must explore the relationship between chromatin or HP1 concentration and HP1 diffusion and confinement.
- Changes in the abundance and distribution of H3K9me2/3, H3S10pho, H3K23me1/2/3, H4K20me3, H1.4K26me2 and other histone marks, linker histone H1, DNA methylation, RNA, HP1 $\beta$ -interacting proteins or PTM of HP1 $\beta$  itself might all lead to HP1 $\beta$  relocalisation to euchromatin in formative mESCs, and to chromocentres in primed cells. Differences in affinities for the different molecules might explain the different patterns observed for HP1 $\alpha$  and HP1 $\beta$ .
- However, the relocalisation of HP1 $\beta$  to chromocentres in the primed state, despite low levels of H3K9me3, is puzzling in this context.

3. Heterochromatic foci are not in equilibrium.

- For several datasets, the differences between the displacements of molecules crossing the inter-compartment boundary, as well as their frequency, suggested that HP1 $\beta$  might be biased towards leaving the heterochromatic domains.
- Repeating experiments with enhanced border definition and/or introducing intervals is needed to confirm this. If it proves to be true, non-equilibrium models must be developed to describe heterochromatin.

4. Viscosity of the nucleoplasm transiently increases in the formative state due to chromatin decondensation.

- This effect might be specific to HP1, e.g. due to its weak binding to histones, RNA and DNA, or could affect all nuclear proteins. This could be verified by studying the dynamics of other nuclear proteins and exogenous neutral probes.
- The consequences (if any) of this phenomenon, and whether it is related to HP1 $\beta$  redistribution, are unclear.

# Chapter 4

## Development of whole-nucleus 3D SMLM in combination with single-nucleus Hi-C

*This project was performed in close collaboration with Ziwei Zhang. The work described in this chapter is either my own or was carried out together with Ziwei, unless otherwise specified. The experimental support on the Hi-C side came from Dr David Lando and Dr Xiaoyan Ma, while the computational work in genome structure modeling and overlay with the imaging data was done by Dominic Hall, Dr Xiaoyan Ma and Dr Wayne Boucher. A lot of advice on the imaging side was gained from Dr Aleks Ponjavic, University of Leeds, and Prof David Klenerman.*

### 4.1 Introduction

Reconstructing the genome fold of individual cells from the snHi-C data has for the first time enabled us to study the global organisation of the genome with sequence-specific information, allowing investigation of the links between sequence and structure and exploration of the heterogeneity present on the single-cell level (Stevens et al. 2017).

However, further developments are needed to improve the quality of these struc-

tures. One way forward is to define additional restraints on top of the snHi-C-derived data. The current method assumes, as a first approximation, that the DNA density across the nucleus is homogeneous. Thus, knowing the real distribution of DNA density would add a new layer of information and lead to more accurate genome models. In addition, it would be possible to study the link between the DNA compaction and sequence on a genome-wide, but single-cell level.

Moreover, knowing the location of chromatin-binding proteins can provide great insight into their function. Although this can be achieved by mapping the ChIP-seq profiles onto the snHi-C structures, it is unclear to what extent the ChIP-seq traces are applicable to single cells. An alternative approach is to use microscopy to determine the spatial distribution of a protein of interest, and to overlay it on the genome structure. This can then be combined with ChIP-seq to provide additional restraints for structure calculation, or used directly to study the protein function.

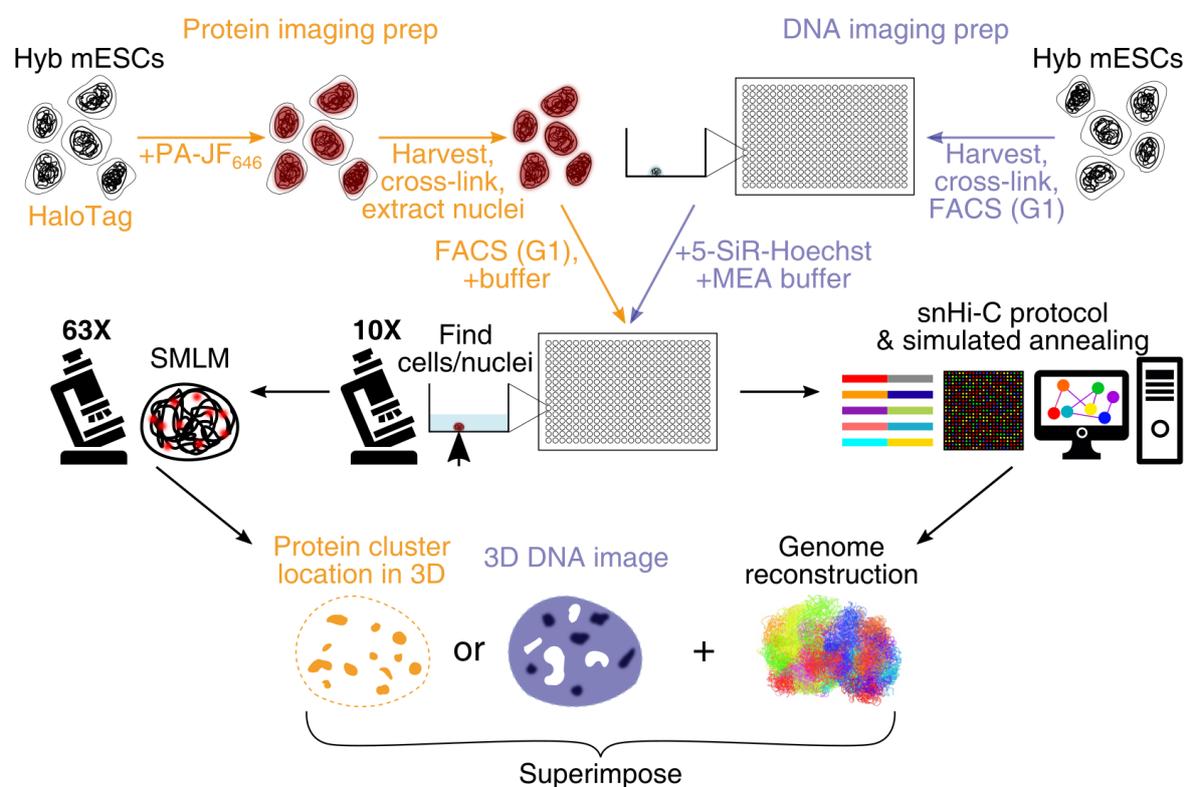
In this chapter I will discuss the work that has been done towards combining high-resolution imaging of the DNA and proteins with snHi-C for single mouse ES cells.

#### 4.1.1 General workflow

The cells used for snHi-C were F1 hybrid mESCs between *Mus musculus domesticus* Black6 and *Mus musculus castaneus* strains, called CB9 or BC8 (Strogantsev et al. 2015), which had a sufficient number of single-nucleotide polymorphisms (SNPs) to enable unambiguous mapping of the number of sequencing reads required for genome structure calculation.

The snHi-C protocol involves sorting individual fixed cells or extracted nuclei into wells by FACS, which are then processed in parallel. To combine the Hi-C protocol with imaging, cells are labelled prior to or immediately after fixation and sorting into multi-well plates (Figure 4.1). The plates are then taken to the microscope and imaged for the next 18-24 hours, since delaying Hi-C processing for longer results in decreased efficiency (Lando et al. 2018). Since recording many frames is needed to reconstruct a high-resolution image, typically 10-15 cell images are recorded in this time window.

After imaging, the Hi-C protocol is continued.



**Figure 4.1:** General workflow for combining SMLM and Hi-C. When imaging proteins (orange), F1 hybrid mESCs with the protein of interest fused to the HaloTag are labeled with PA-JF<sub>646</sub> live. The cells are then harvested and fixed. The nuclei are extracted and sorted into individual wells of a glass-bottom 384-well plate, enriching for cells in G1 phase, followed by addition of extra Tris-HCl buffer. For DNA PAINT (blue), cells are FACS-sorted straight after harvesting and cross-linking. The MEA- and 5-SiR-Hoechst-containing Tris-HCl buffer is then added. At the microscope, a low-magnification objective lens is used to scan the plate in the bright-field mode and locate the cells or nuclei in the wells. A high-magnification objective lens is then used for SMLM imaging, producing a high-resolution map of protein or DNA density. After the imaging is completed, snHi-C is performed on the same cells, and genome models are reconstructed by simulated annealing.

### 4.1.2 Imaging: 3D DH-PSF SMLM

SnHi-C-based genome structures are constructed using a polymer model with beads representing 25-100 kb of sequence. Thus, their physical diameter is in the order of 50-200 nm (based on data in Mateo et al. (2019), and the ratio of the physical and reconstructed nuclear diameters), meaning that ideally, the corresponding protein or DNA images should have a comparable or better resolution. Furthermore, to enable one to overlay them with the entire structures, the images must be three-dimensional

and contain the entire nuclear volume.

SMLM and STED super-resolution microscopy techniques provide the highest resolution attainable by light microscopy. Although STED is much faster, the bleaching induced by 3D STED usually makes it impossible to image the entire nucleus. Multiple 3D SMLM variations exist, including astigmatic, double-helix, tetrapod and biplane SMLM. Of these, DH-PSF microscopy has the largest z-range with relatively uniform resolution and compact PSF (Sage et al. 2019), and thus 3D DH-PSF SMLM (see section 1.5) was chosen for combination with snHi-C. However, the potential of employing other microscopy methods, such as spinning disk confocal, is explored in the lab.

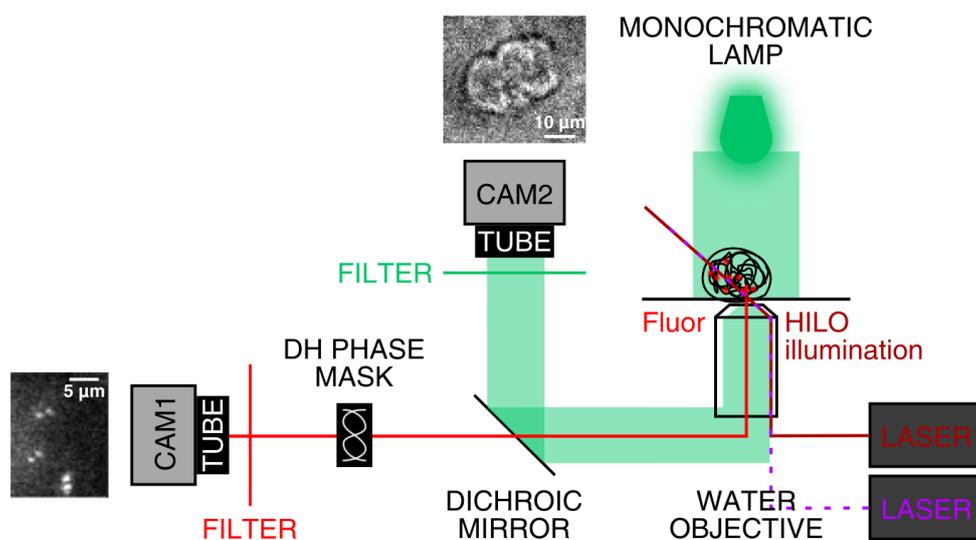
To perform this imaging, the same custom-built microscope is used as for 3D SPT experiments (see section 6.4.2). The 384-well plate is initially scanned in bright-field mode, using a 10 $\times$  objective lens, to find the single cells or nuclei in wells and record their coordinates (scripts and workflow developed by Ziwei Zhang)(Figure 4.1).

After the cells are located, the microscope is switched into the super-resolution imaging mode (Figure 4.2). The typical illumination setup involves a 638 nm excitation laser and a 405 nm activation laser, used in the highly inclined and laminated optical sheet (HILO) mode to minimise the background, as well as a monochromatic 530 nm lamp. Splitting the emission with a dichroic mirror allows imaging in two colours simultaneously by using two cameras: 3D DH-PSF SMLM in red, and bright-field imaging for drift correction in green (see section 4.6).

## 4.2 Data processing pipeline

### 4.2.1 3D localisation detection and determination of coordinates

To reconstruct a super-resolved 3D DH-PSF SMLM image, individual double-peak-shaped localisations in the data must first be detected, and then be used to accurately determine the molecules' x, y and z coordinates. The x and y positions are simply found by taking the midpoint of the PSF and correcting it appropriately. To enable



**Figure 4.2:** Schematic representation of the super-resolution microscope. The sample is illuminated with a high-power 638 nm laser, and in addition a 405 nm activation laser is used for PALM. HILO illumination is employed to reduce out-of-focus fluorescence. The emission from blinking fluorophores (Fluor) is collected through the objective lens. The PSF is reshaped when the light passes through a phase mask, and the light is focussed on an EMCCD camera detector by a tube lens (TUBE). Additionally, the sample is illuminated with collimated monochromatic 530 nm light from above, and a bright-field image of cells is recorded with the second camera for drift correction. Examples of the resulting images are shown next to the cameras (scale bars, 5 and 10  $\mu\text{m}$ ). Note that this schematic is illustrative and thus not entirely accurate.

finding the z-coordinate, a calibration series is recorded prior to every experiment, using fluorescent beads (Figure 4.3A, left). The stage is moved vertically in defined steps, so that the pitch of the double-helix can be matched with the fluorophore position. Special analysis software is employed to determine the coordinates of the molecules from the raw data, using the calibration series.

### Comparison of software

The well-established and widely used easy-DHPSF software package (Lew et al. 2013) was found to be very slow and also prone to detecting false-positives. Therefore, alternative options were explored: DHPSFU, a new software written by Dr Aleks Ponjavic (unpublished; available at <https://github.com/TheLaueLab/DHPSFU>); and SMAP (Li et al. 2018; Ries 2020), which was ranked the best in a comprehensive assessment of 3D SMLM software packages using realistic simulated data (Sage et al. 2019).

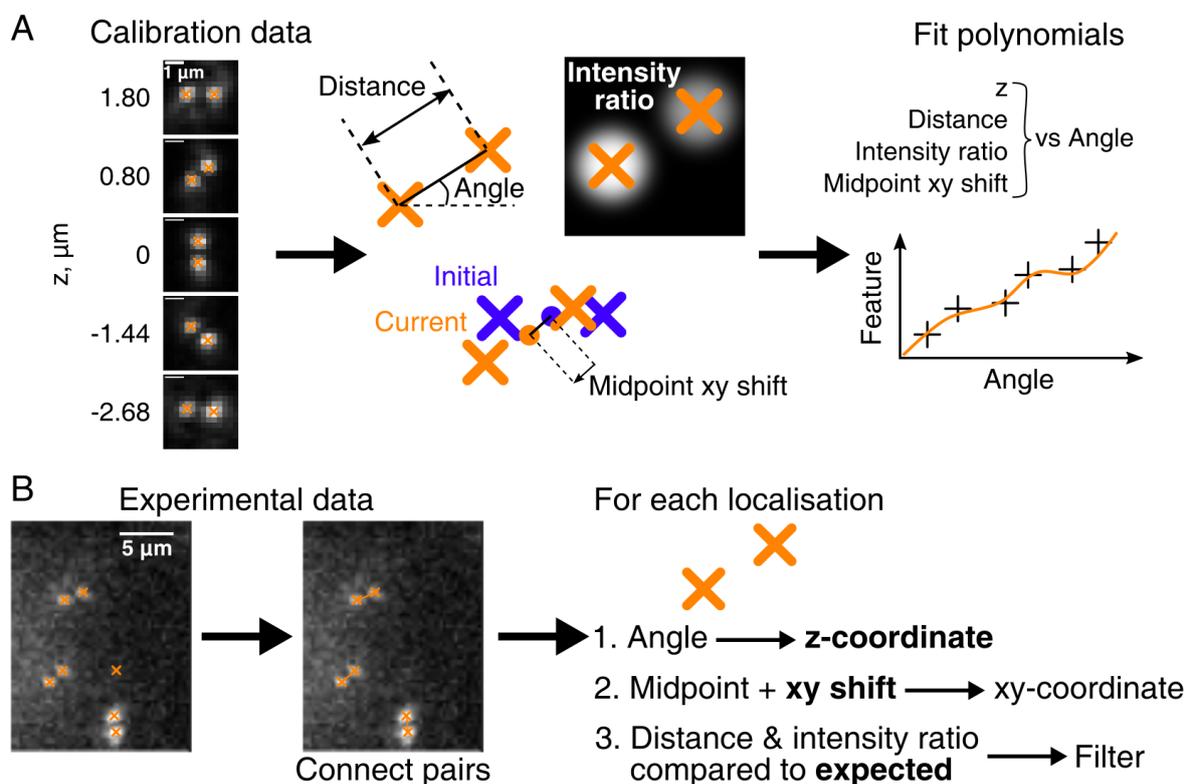
**Easy-DHPSF** Easy-DHPSF is a two-step algorithm. First, template matching, using a selection of templates from the calibration series, is performed to identify potential localisations. A user-defined threshold of goodness of the match is used to select robust hits. Next, the z-coordinate of each molecule is determined by fitting the candidate to the double-Gaussian model of the DH-PSF, which is constructed using the calibration series. The fits are then filtered based on how well they match the data.

**DHPSFU** Unlike the other two packages, DHPSFU does not take the raw data as an input. Instead, the 2D Gaussian peaks in the data must first be identified using a different software, such as GDSC PeakFit (Herbert 2014), and the list is then passed on to DHPSFU. Using a calibration sequence, DHPSFU determines the standard pitch, midpoint shift, intensity ratio and the distance between the peaks at each z-coordinate (Figure 4.3A). Polynomial functions are then fitted to this data, enabling one to predict all features, using just the angle.

For each frame in the experimental dataset, peaks likely originating from the same molecule are paired up (Figure 4.3B). The pitch is calculated for these pairs and used to determine the z-coordinate. The x-y-coordinate of the molecule is determined by calculating the midpoint between the lobes, and correcting it by the expected shift at this z-coordinate. The expected parameters of the PSF at this z-position are also computed, and localisations that deviate too far from the expectation are discarded, using tunable filters.

**SMAP** SMAP (Ries 2020) includes a universal fitting tool for any PSF modulation, including DH-PSF (Li et al. 2018). To achieve this, a filter of a PSF-matching diameter is initially applied to both the calibration and the data, blurring the PSFs. Next, Gaussian peaks above a certain (tunable) intensity threshold are located in the images. The original calibration series is then used to construct a 3D model of the PSF by 3D spline fitting. The z-coordinates of the molecules found in the experimental data are determined using this model.

An experimental PAINT dataset was used to test the three programmes. Several



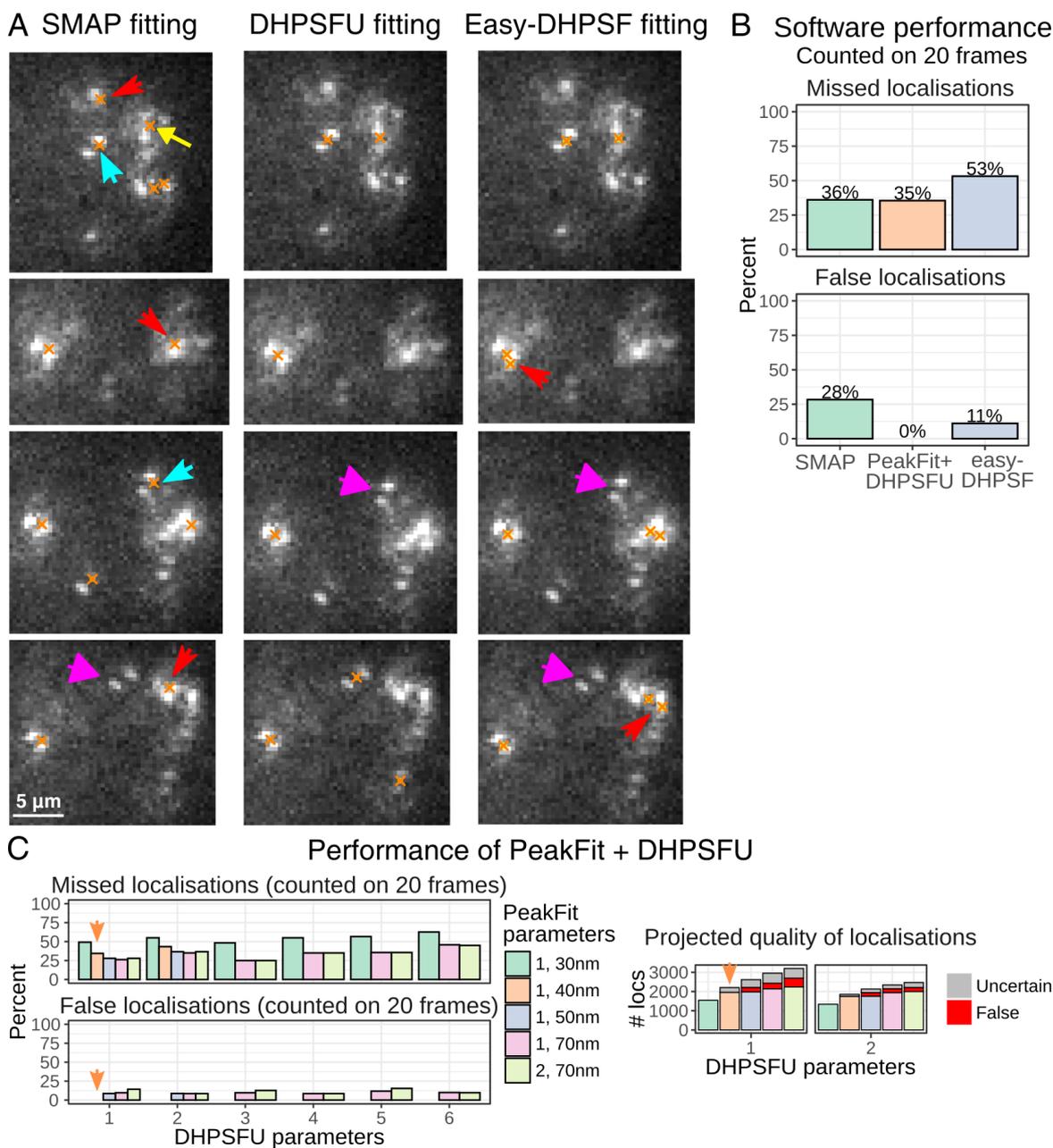
**Figure 4.3:** Principle of the DHPSFU algorithm. A. Peaks in the calibration series (scale bars, 1  $\mu\text{m}$ ) are found using an external 2D localisation software (orange crosses). For each frame in the calibration dataset, the distance between the points, the angle between the line connecting them and the horizon (the pitch), their intensity ratio, as well as the shift of the midpoint relative to the starting position are calculated. These parameters are linked to the z-coordinate known from the calibration sequence, and polynomial functions (orange line) are fitted to the data (black crosses) to describe the dependence of z, the intensity ratio and the midpoint shift from the angle. B. In the experimental dataset, after locating the individual peaks, the lobes of the DH-PSF are connected based on pairwise distances. This procedure also results in exclusion of unpaired noise localisations. For each pair of peaks: 1. the z-coordinate is predicted based on the angle; 2. the x- and y- coordinates are determined from the midpoint and corrected for PSF-induced shift; 3. localisations that have large deviations from expected distance between the lobes and intensity ratio at this angle are filtered out. The features in bold are predicted from the pitch of the localisation, using the polynomial functions fitted to the calibration data. Scale bar, 5  $\mu\text{m}$ .

inherent problems of the SMAP algorithm became apparent upon applying it to the DH-PSF data (Figure 4.4A). Firstly, since the only threshold applied is based on the total intensity of the peaks in the filtered image, many false-positives were found in the experimental data. These included both strong background noise and malformed and/or poorly fitted localisations. Furthermore, since the x-y-coordinate of the localisation is defined as the peak location in the filtered data, it was often shifted due to uneven intensity of the lobes or noise. Finally, when two PSFs were located close, they merged into one single peak upon filtering. All of these issues occurred rather frequently even with carefully tuned parameters and with post-filtering of data falling outside the well-defined range of the calibration (Figure 4.4A, B). This suggests SMAP is less useful for fitting DH-PSF data, at least of the density and quality obtained in our experiments.

Easy-DHPSF had a tendency to fit additional localisations when the PSF was bright and/or the shape of the lobes deviated from a Gaussian (Figure 4.4A). Similarly to SMAP, this led to a significant number of artifacts. Although when more relaxed thresholds were used, easy-DHPSF could retrieve the maximum number of true positive localisations among the three packages (data not shown), the number of artifacts was also very large. Therefore, stringent thresholds had to be applied, leading to a high number of missed molecules (Figure 4.4B). In addition, easy-DHPSF ran much slower than SMAP or PeakFit+DHPSFU. Thus, a combination of GDSC PeakFit and DHPSFU packages showed the best performance among the three algorithms.

### **Optimisation of parameters for PeakFit+DHPSFU**

Next, an optimal set of parameters was identified following a comprehensive evaluation of software predictions in comparison to the localisations identified by eye (Figure 4.4C). The best performance was achieved by using a rather stringent threshold for localisation precision of individual Gaussian peaks in PeakFit, combined with the most lenient set of DHPSFU filtering parameters (for intensity ratio and inter-lobe distance) among those tested. At least in the test dataset, this combination resulted in no clearly false localisations and only a small number of potentially erroneous ones, while 66%



**Figure 4.4:** A. Examples of the same images analysed using SMAP, PeakFit+DHPSFU and easy-DHPSF. Orange crosses show the x-y-coordinates of the detected localisations. Malformed PSFs and artifacts picked up by the software (narrowest red arrowheads), localisations with a shifted x-y-position (medium cyan arrowheads), fused localisations (yellow arrow with triangular head), and obvious DHPSFs missed (large magenta arrowheads) are indicated. Scale bar, 5  $\mu\text{m}$ . B. Performance of the three packages, manually evaluated on the same 20 frames of a test dataset. Shown are the percentage of true localisations missed (top) by each programme and the percentage of erroneous localisations among those detected (bottom). Shown is the best result from a panel of parameters tested. C. Optimisation of PeakFit and DHPSFU parameters. The left panel shows the same metrics as B for each set of PeakFit parameters (1 or 2, with the indicated precision threshold) and DHPSFU parameters (1-6) used. The right panel shows the number of localisations found in the entire test dataset, using some of the parameter combinations. The expected number of true (coloured by PeakFit parameters), false and uncertain localisations among them was determined by multiplication by the percentages show on the left. The orange arrow indicates the chosen parameter set.

of the true localisations were picked up. Interestingly, by combining these estimates with the total number of localisations detected in the dataset, it became apparent that further relaxation of the precision cutoff lead to an increase in false and uncertain localisations, but almost no improvement in the number of true positives. Thus, this set of parameters was used for all analysis.

Furthermore, although ideally, the DH-PSF should become very dim beyond its angular range of  $\pm 90$  degrees from the middle, in reality it is often not the case. This can lead to localisations above or below the DH-PSF z-range being detected and erroneously mapped. To avoid this, the localisations at the extremes of the DH-PSF z-range should not be used. The option to crop the calibration into the desired range was thus introduced into DHPSFU, which also gave the user control over the thickness of the resultant layer. In this study, the z-range of 3.2  $\mu\text{m}$  was typically used.

## 4.2.2 Duplicate removal

The same molecule can be detected in multiple consecutive frames, yielding a pseudo-cluster of localisations. To correct for this, a procedure similar to that used for tracking molecules in SPT was also applied to the SMLM data. Localisations were joined into “trajectories”, if they occurred on consecutive frames (allowing for gaps of up to 50 frames) and were within 500 nm from each other. To generate the final output, the x-, y- and z-coordinates of each trajectory were averaged, increasing the precision of SMLM (Sage et al. 2019).

## 4.3 Optimisation of 3D protein SMLM

### 4.3.1 Overview

To enable detection of specific proteins, genes encoding these proteins were modified to yield a fusion with the HaloTag, which in its turn could be specifically labelled with a variety of small-molecules fluorophores. Since the DH-PSF transformation decreases the intensity of individual emitters, it was critical to harness the increased brightness

of organic dyes compared to fluorescent proteins by adopting this strategy. HaloTag was chosen following earlier difficulties in the lab when labelling SNAP-tagged proteins in naive mESCs. Moreover, it has been reported to give a much brighter signal when labelled with a variety of fluorophores including JF<sub>646</sub> (Erdmann et al. 2019). Endogenous knock-ins were generated by CRISPR-Cas9 to ensure that correct levels of the chromatin factors under study were maintained in cells and not to perturb the chromatin conformation or cell state.

The most popular SMLM approaches rely on stochastic blinking of fluorophores (e.g. STORM) or on photoactivation (e.g. PALM). We settled on the photoactivation approach, where due to rare occurrence of prolonged dark states and multiple blinking events it was easier to count molecules accurately and quantify clustering (Shivanandan et al. 2014). For labeling, we used bright photoactivatable Janelia Fluor (PA-JF) fluorophores, mainly Janelia Fluor 646 (PA-JF<sub>646</sub>) (Grimm et al. 2016).

### **Proteins of interest**

The combination of PALM and snHi-C was developed by imaging three proteins. A panel of proteins, which display a non-homogeneous pattern in mESCs and whose distribution is of biological interest – such as naive transcription factors and chromatin modifiers – were considered, and the ones for which generation of HaloTag fusion cell lines was successful were used in the experiments.

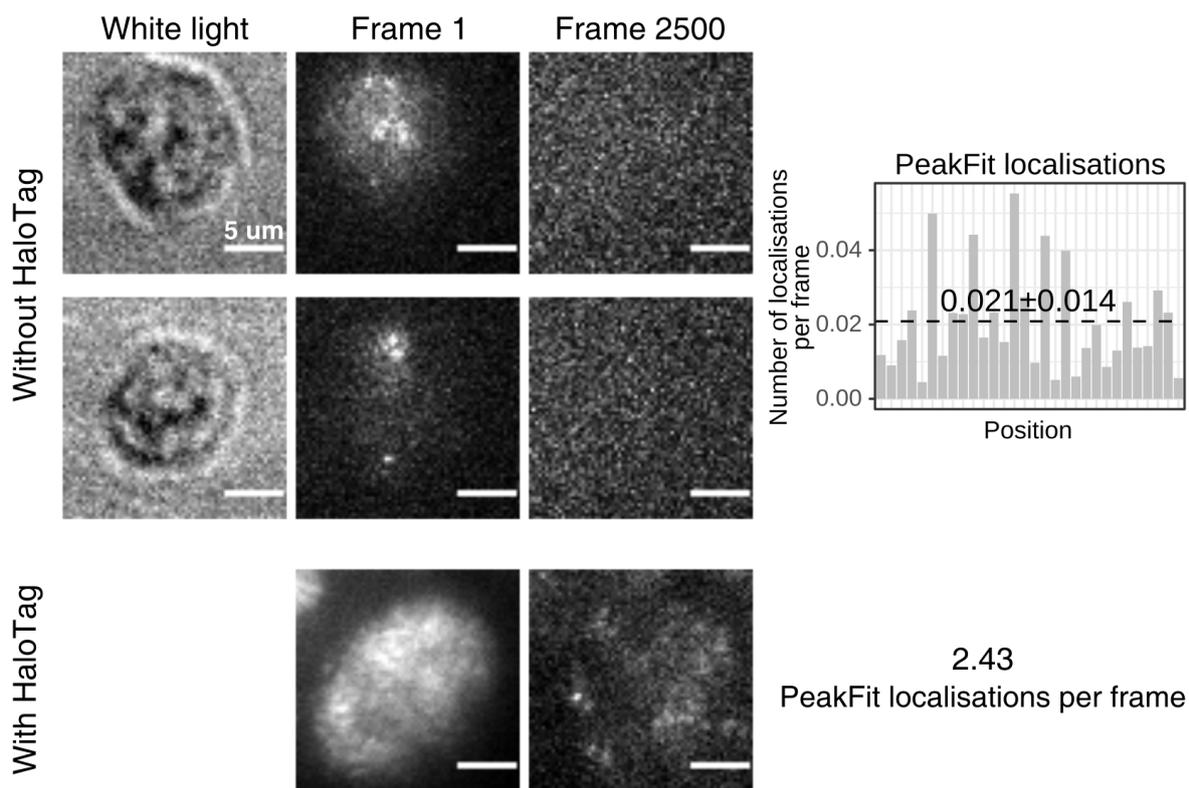
**NANOG** is a well-known master transcription factor, supporting the naive pluripotent identity of mESCs (Hackett and Surani 2014). Overlaying NANOG ChIP-seq peaks onto the snHi-C structures revealed clustering, albeit not very strong, of NANOG-bound loci (Stevens et al. 2017). Thus, the NANOG pattern in the nucleus might potentially be heterogeneous enough to distinguish NANOG-rich regions of the genome. *Nanog* gene in CB9 cells was heterozygously targeted by CRISPR-Cas9 by Dr Robin Floyd to create a fusion with HaloTag.

**SOX2** , one of the Yamanaka factors, is also a crucial master regulator of pluripotency. SOX2-bound regulatory elements are known to group spatially in mESCs, and SOX2 itself forms clusters upon exit of mESCs from naive pluripotency (Liu et al. 2014; Veneri et al. 2020). SOX2 is also an abundant protein, and thus, due to higher density of localisations, its super-resolved images could be easier to obtain and interpret than e.g. those of NANOG. The HaloTag sequence was inserted into the endogenous *Sox2* gene on one of the chromosomes in BC8 cells by Dr Robin Floyd.

**HP1 $\beta$**  has been introduced in the previous chapters, and the HP1 $\beta$ -Halo(N) line constructed by Dr Maike Wiese was also used for PALM imaging in this chapter. Since HP1 $\beta$  is very abundant and forms large foci, it is very well-suited for super-resolution imaging in combination with Hi-C. The drawback here is that a large proportion of it is located in the pericentromeric heterochromatin, and with the present genome reference it is hard to map reads (from either ChIP-seq or Hi-C) to these regions. However, attempts are underway in the lab to resolve this issue by sequencing the genome of F1 hybrid mESCs, using PacBio sequencing.

### 4.3.2 Sample preparation

The protocol for labeling mESCs with PA-JF dyes from Grimm et al. (2016) was modified to yield better specificity and shorten the preparation time (see section 6.6.1). It was also combined with nuclear extraction for downstream Hi-C processing. A control experiment labelling and imaging CB9 mESCs lacking a HaloTagged protein was performed to check the specificity of the altered method (Figure 4.5). 30 nuclei were imaged for 10,000 frames, using the same setup and settings as for the imaging-Hi-C experiments. Using the optimal (stringent) PeakFit parameters, on average 0.021 localisations per frame were detected per nucleus in the cells lacking the tag, while 2.43 localisations per frame were found in the positive control, imaged for 3,000 frames. Thus, the modified protocol was used to label, fix and permeabilise HaloTagged mESCs with PA-JF<sub>646</sub> before nuclear extraction and cell sorting (Figure 4.1).



**Figure 4.5:** The top two rows show two representative examples of cells lacking HaloTagged protein, labelled with PA-JF<sub>646</sub>, prepared and imaged under the same conditions as during an experiment, while the bottom row shows a cell expressing the NANOG-Halo fusion as a positive control. On the right, the number of simple Gaussian localisations identified by PeakFit is shown for all 30 negative control nuclei (10,000 frames each), as well as the mean  $\pm$  standard deviation, and for the positive control (3,000 frames). The intensity of each image was adjusted; scale bar, 5  $\mu$ m.

### 4.3.3 Imaging optimisation

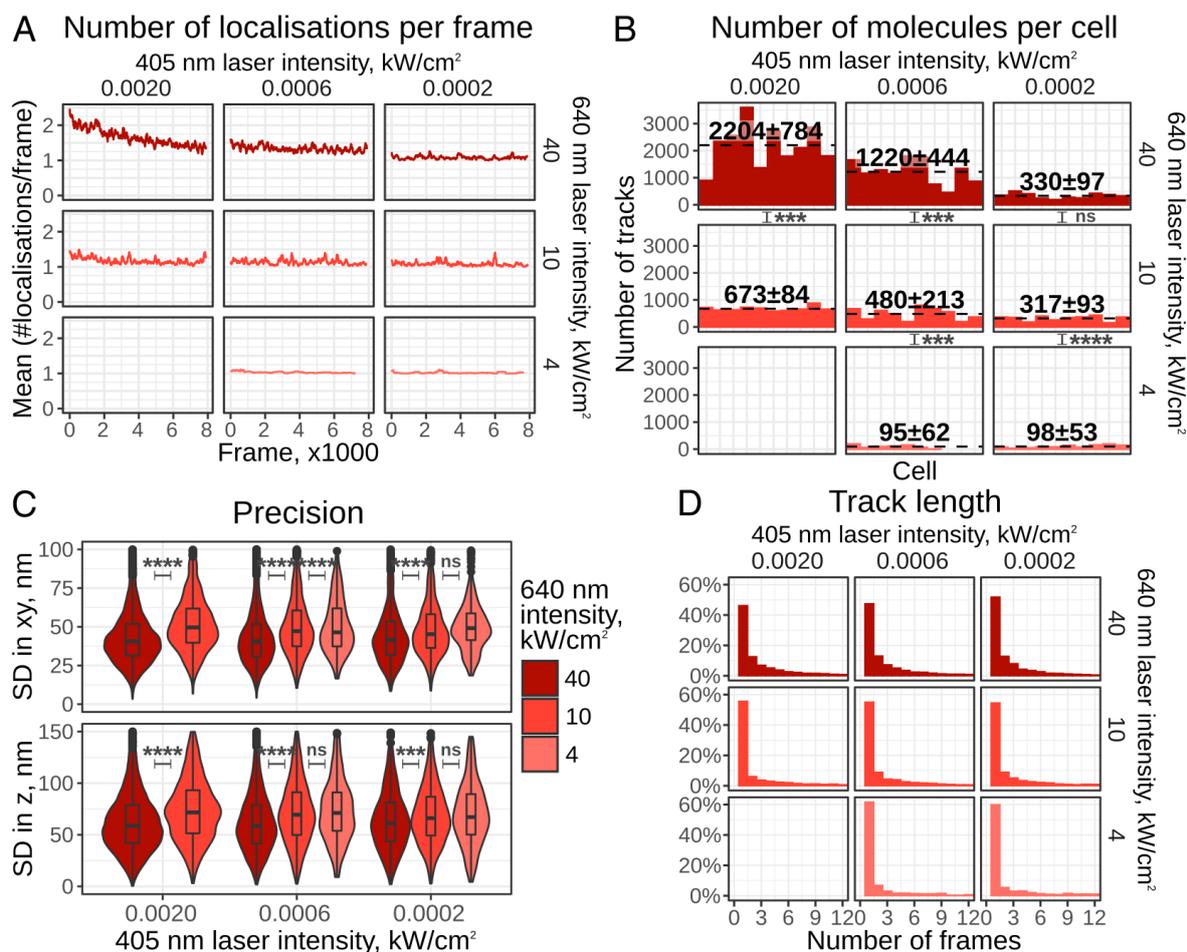
#### Laser intensity and exposure time

The laser intensities used in SMLM must be optimised to gain the maximum amount of data in the shortest amount of time. Naively, the excitation power should be as high as possible to bleach molecules quickly and make space for newly activated fluorophores. However, very strong excitation can lead to loss of data due to molecules bleached before they are imaged (Thevathasan et al. 2019), and to high background. Activation laser power, on the other hand, needs to be low enough for single molecules to be easily distinguished, which is especially important in DH-PSF SMLM (Sage et al. 2019). Within this limit, however, the higher the density of molecules, the faster the data is acquired.

These parameters were varied in a trial imaging experiment of NANOG, labeled with PA-JF<sub>646</sub>. The excitation laser intensity was varied between the maximum (40 kW/cm<sup>2</sup>), 25% and 10% of the maximum, and the activation laser power was probed around the empirically estimated optimum (0.0006 kW/cm<sup>2</sup>), and about 3 times higher or lower. Prior to imaging, each cell was illuminated with maximum-power excitation laser for 500 frames to pre-bleach the spontaneously activated fluorophores.

The number of molecules per frame should decay exponentially as imaging progresses. Indeed, this trend was clear at the highest laser powers, but at lower intensities the decay rate was very slow (Figure 4.6A). Higher numbers of localisations per frame were associated with both higher 638 nm and 405 nm laser intensities (Figures 4.6A,B). However, at the highest activation rate, clashes between individual PSFs were frequent.

The localisation precision, as expected, was best for the strongest excitation, while the activation intensity did not affect it (Figure 4.6C). Overall, the strongest excitation and intermediate activation laser powers were chosen as optimal to maximise the total number of detected molecules and the speed of data acquisition without compromising its quality. Importantly, the perfect settings for SMLM can vary, depending on the setup of the microscope, the protein imaged, etc. Thus, these laser intensity values



**Figure 4.6:** A. The number of localisations per frame during acquisition with different laser intensities. The values are averaged between >10 cells, and smoothing using a rolling mean across 100 frames is used for clarity. B. The final number of molecules obtained for each cell (after duplicate removal). The dashed line shows the mean value, and the mean  $\pm$  standard deviation is indicated. C. Localisation precision, defined as the standard deviation of the localisation within the same track (track length  $\geq 3$ ). In B and C, different excitation intensities (at the same activation intensity) were compared using Welch's ANOVA test with Games-Howell post-hoc: ns, non-significant; \*\*\*,  $p < 0.001$ ; \*\*\*\*,  $p < 0.0001$ . D. Histograms of track lengths at different laser intensities.

were used as a starting point, and were further fine-tuned during every imaging session.

The exposure time for this experiment was chosen to be 20 ms from empirical observations. The majority of localisations only appeared in one frame and were quickly bleached, suggesting that this regime was suitable (Figure 4.6D). Interestingly, stronger excitation lead to longer “trajectories”, despite stronger bleaching. This was likely the result of a higher percentage of emitters detected in these conditions, but could also be due to an increase in blinking. Potentially, the current settings could benefit from increasing the excitation intensity even further, but this was not possible with the current setup.

### **Glass thickness and DH-PSF setup**

The thickness of the coverslip glass, used for imaging, influences the optical aberration of the system. High-numerical aperture (NA) objectives are usually optimised for 0.170-0.190 mm-thick glass, and have built-in adjustable collars for correcting this effect. The 384-well plates used for Hi-C had a non-standard 0.188 mm-thick bottom, made from optical cyclic olefine. Thus, it turned out to be crucial to set the microscope up and record the DH-PSF calibration series, using an identical 384-well plate with fluorescent beads added to one of the wells.

## **4.4 Optimisation of 3D DNA PAINT**

*The early experiments in this section were performed by Dr Kirti Prakash and Alzbeta Roeselova.*

### **4.4.1 Overview**

The PAINT approach was chosen to attempt obtaining 3D super-resolution images of DNA (see Section 1.5.1). Note that in this case DNA PAINT does not mean the widespread antibody and oligonucleotide-based imaging method, but rather the general approach of super-resolution imaging with reversible binders, applied here to DNA. As

discussed in the Introduction, PAINT is particularly well-suited for SMLM of large polymers and structures like the genome. To our knowledge, similar approaches have only been used to image DNA in two dimensions (Bucevičius et al. 2020; Zhang et al. 2019), and never in 3D, let alone for the entire nuclear volume. However, 3D STED images of DNA in entire nuclei have been obtained previously (Spahn et al. 2018; Noa et al. 2021).

For this application, the fluorophore of choice was 5'-carboxy silicon rhodamine, conjugated via the carboxy moiety to a reversible DNA binder Hoechst 33258 (5-SiR-Hoechst) (Bucevičius et al. 2019). Compared to the conventional 6'-regioisomer Hoechst conjugates, this fluorogenic dye displays a single-mode binding to the DNA minor groove and increased brightness, which makes it well-suited for SMLM. Furthermore, SiR-Hoechst conjugates have a low ( $\mu\text{M}$  range) affinity for DNA, which makes them good candidates for PAINT probes.

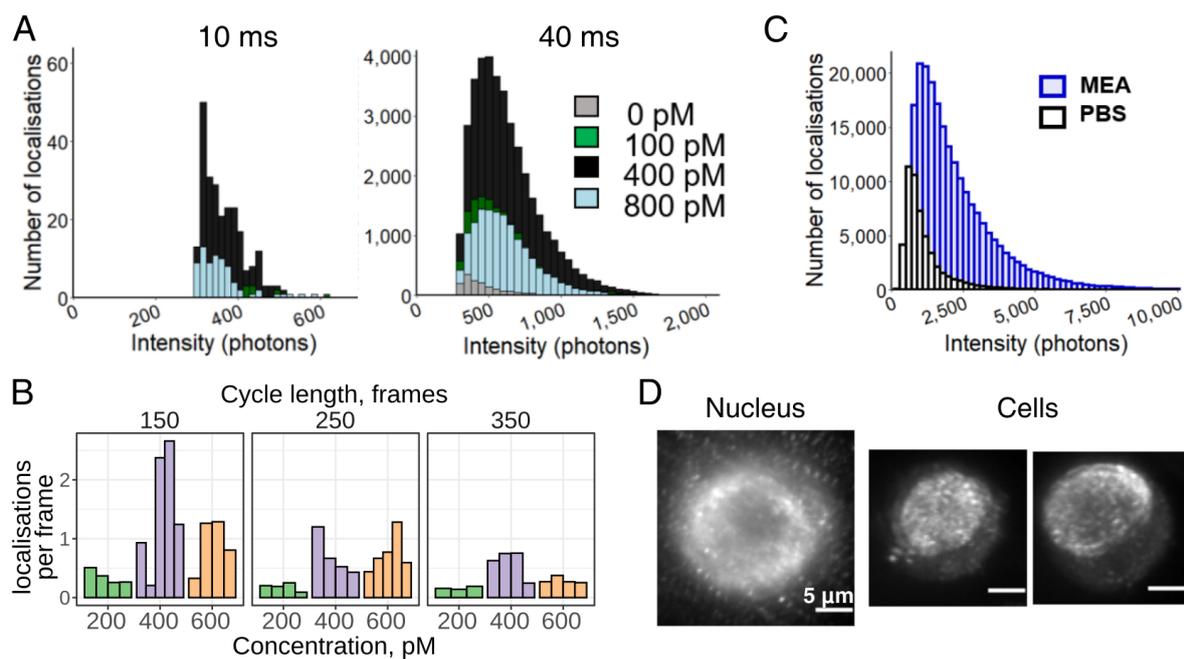
#### 4.4.2 Sample preparation

Initially, the sample preparation for DNA PAINT imaging combined with snHi-C involved cell fixation with formaldehyde and nuclear extraction, followed by FACS and addition of buffer containing 5-SiR-Hoechst to each well. This basic protocol was then optimised in multiple ways.

##### Optimising 5-SiR-Hoechst concentration

To enable imaging in the single-molecule regime, the optimal concentration of 5-SiR-Hoechst had to be determined. This was initially performed by Alzbeta Roeselova, who found the optimum  $[5\text{-SiR-Hoechst}] = 400 \text{ pM}$  for 2D imaging, corresponding on average to  $\sim 170$  molecules per nuclear volume (Figure 4.7A). The negative control using buffer without 5-SiR-Hoechst showed that the observed localisations were specific to dye binding events rather than random noise.

The localisations in DH-PSF imaging take up more space within a frame than a single Airy disk. Furthermore, when acquiring images of whole nuclei, the sample



**Figure 4.7:** A. Histograms of intensities of localisations from a 2D DNA PAINT experiment, using different concentrations of 5-SiR-Hoechst and two exposure times. B. Bar plots of the number of localisations per frame, captured in a whole-nucleus 3D DNA PAINT experiment, using different concentrations of 5-SiR-Hoechst and cycle lengths. Each bar represents an individual cell. C. Histogram of intensities of localisations from 2D DNA PAINT imaging either in PBS or an MEA-containing buffer. D. Standard deviation projections of 5000 frames of a 3D DNA PAINT single-slice imaging trial. An extracted nucleus is shown on the left, and two cells, treated with 0.1% Triton-X100, are shown on the right. The left cell displays uniform labeling throughout its nucleus. Scale bars, 5  $\mu\text{m}$ . Experiments and figures in A and C were made by A. Roeselova, images in D were acquired by D. Shah and Z. Zhang.

is scanned in  $z$  in multiple cycles, which leads to a different dynamics compared to continuous imaging (see Section 4.5). Finally, reasons that are unclear, FACS-sorted cells in a 384-well plate had fewer localisations compared to cells fixed on a dish. Therefore, a dye titration was performed again for sorted cells in a 384-well plate in the 3D regime (Figure 4.7B). Importantly, since cycle length affects the density of localisations (see Section 4.5.1 and Figure 4.8C), this testing was performed at three different cycle lengths. After testing concentrations 200-600 pM, 400 pM of 5-SiR-Hoechst was confirmed to yield a higher number of localisations than 200 pM and a similar number to 600 pM.

### **MEA buffer**

Reducing agents and triplet state quenchers such as 2-mercaptoethanol or mercaptoethylamine (MEA) alter the intensity and photostability of fluorophores (Gong et al. 2019). However, since 5-SiR-Hoechst was primarily developed for live-cell imaging, their effect on this dye has not been tested before. Thus, PAINT imaging of DNA in mESCs in PBS or an MEA-containing Tris-HCl buffer was compared by Alzbeta Roeselova (Figure 4.7C). Indeed, the intensity of localisations increased substantially in MEA-containing buffer, leading to a better signal-to-noise ratio, a four-fold increase in the number of detected localisations, and better precision.

### **Cell fixation and treatment**

Initial experiments using extracted mESC nuclei suggested that either the exposed nuclear envelope bound 5-SiR-Hoechst, or the dye could not penetrate into the centre of the nucleus before being bleached. While this effect was very prominent in all extracted nuclei, whole cells appeared to be less susceptible to this (Figure 4.7D). Guided by the optimised *Janelia Fluor* dye labelling method (Grimm et al. 2016), permeabilisation and blocking steps were introduced into the protocol. This did not lead to an improvement for extracted nuclei, but permeabilisation with TritonX-100 increased the proportion of homogeneously labelled whole cells. Therefore, for DNA

PAINT experiments, the cells were fixed and permeabilised before sorting, the intact cells were imaged, and nuclear extraction was performed only after that at the start of the Hi-C protocol (Figure 4.1).

### 4.4.3 Imaging optimisation

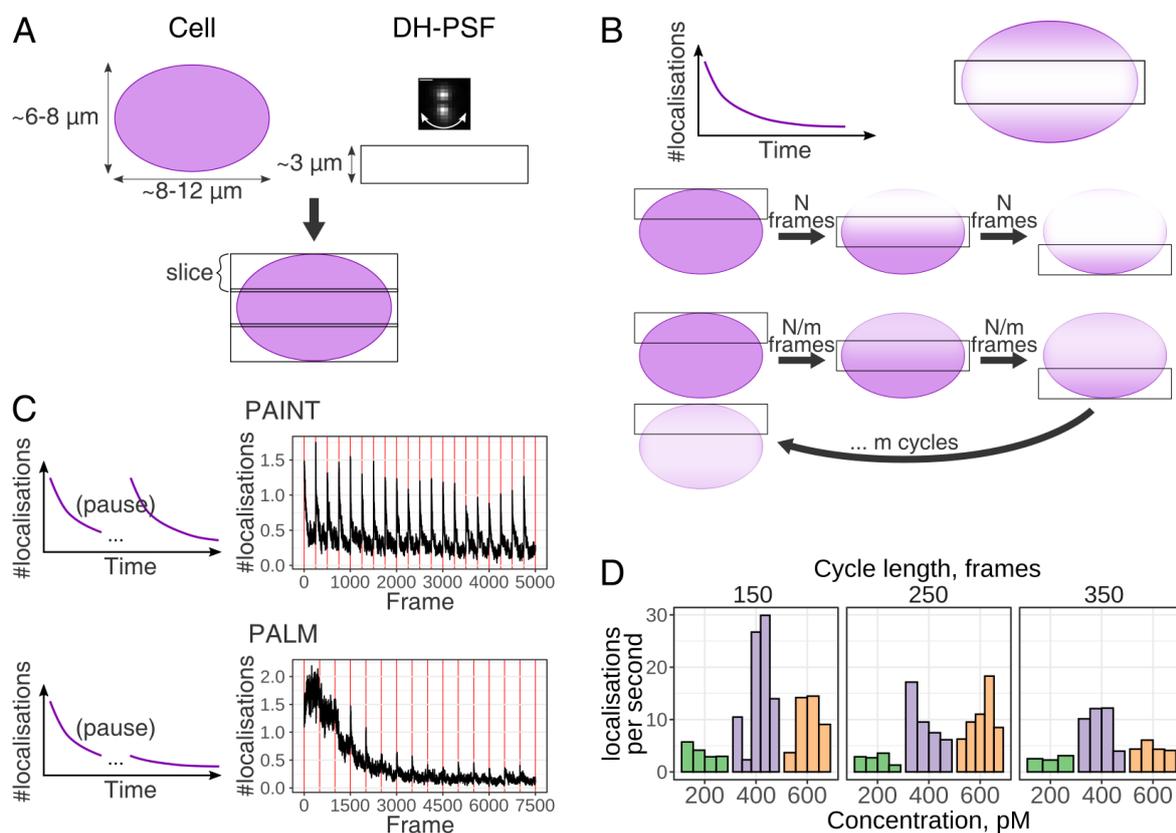
In parallel, optimisation of the imaging conditions was performed by Alzbeta Roeselova. 5-SiR-Hoechst was found to be dimmer than the Janelia Fluor dyes, and thus even with maximum excitation laser power a 10 ms exposure gave low-intensity localisations, which rarely passed the quality filters. Upon extending the exposure time to 40 ms, the number of localisations increased more than a 100-fold, and thus this frame rate was employed in subsequent experiments (Figure 4.7A).

## 4.5 Whole-nucleus imaging

The z-range of the DH-PSF microscopy with our phase masks is limited to about 4  $\mu\text{m}$ , and in practice only the middle  $\sim 3$   $\mu\text{m}$  are often used (see section 4.2). Therefore, to image roughly 6-8  $\mu\text{m}$ -thick mESCs, the data needs to be recorded at multiple positions in z, and then stitched together to produce a single image (Figure 4.8A). I will refer to localisations imaged at the same focal plane as a “slice”, and in this section the optimisation of recording and stitching of slices is discussed.

### 4.5.1 Imaging process

The microscope programme for capturing the image of the entire nucleus in 3D had to be designed with several considerations in mind. In the absence of a thin light sheet, imaging in one plane inevitably leads to bleaching of fluorophores in other planes. Therefore, imaging slices one after another would lead to different data densities between them. To avoid this, the bleaching effects are “blurred” by imaging in cycles with a relatively small number of frames in each (Figure 4.8B).



**Figure 4.8:** A. Several “slices” are needed to cover a typical cell, since the range of DH-PSF in  $z$  is about 3-4  $\mu\text{m}$ . B. Due to photoactivation and subsequent bleaching, the number of localisations drops exponentially during imaging (top; see also Figure 4.6A). When a slice is illuminated, some molecules in the neighbouring slices are also bleached (middle). To make sure this effect does not result in inhomogeneous data density between slices, imaging is performed in cycles (bottom). C. Cyclic acquisition leads to a significant replenishment of localisations in PAINT and a mild one in PALM (left). Example plots of the number of localisations per frame during SOX2-Halo whole-nucleus 3D DNA PAINT (250 frames per cycle) and PALM (500 frames per cycle) experiments demonstrate the bursty dynamics resulting from imaging in cycles (right). The traces were obtained by averaging across 12 slices and 6 or 5 cells, respectively. Red vertical lines indicate beginning of a new cycle. D. Bar plots of the number of localisations per second, captured in a whole-nucleus 3D DNA PAINT experiment, using different concentrations of 5-SiR-Hoechst and cycle lengths (see also Figure 4.7B). Each bar represents an individual cell.

However, imaging in cycles leads to a different dynamics of the number of localisations per frame, compared to continuous acquisition. Ordinarily in PALM, an exponential decay in number of localisations per frame is caused by the depletion of the pool of photoactivatable molecules through bleaching. A similar process occurs initially in PAINT, but eventually binding of fresh fluorophores from solution and bleaching come into equilibrium. During cyclic acquisition, the slices that are not imaged at a particular moment have a chance to replenish their pool of fluorophores (Figure 4.8C). Therefore, the sample preparation protocols (see section 4.4.2) and the imaging setup had to be optimised together.

### **Number of frames per cycle**

The number of frames per cycle was chosen based on the bleaching dynamics in a particular setting. The trade-off was the imaging time, since for every slice/cycle, several seconds are needed to close the shutters, move the stage, reopen the shutters, and acquire the images for drift correction.

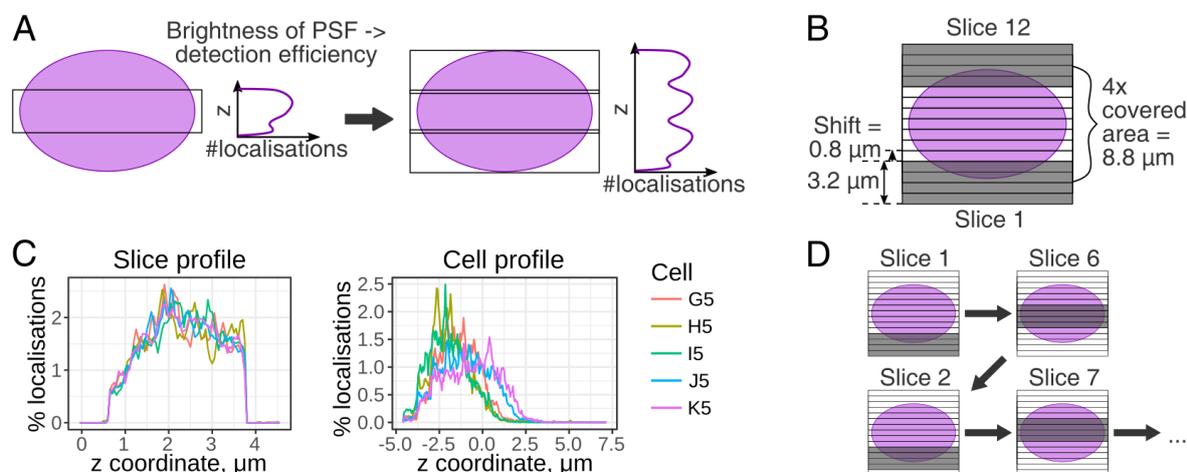
For protein PALM with PA-JF<sub>646</sub>, the effect of the cyclic acquisition on the dynamics of blinking was mild, with the spike in localisations restricted to the first few frames (Figure 4.8C). The burst may have been caused by autoactivation or exit from a dark state of fluorophores. Therefore, a cycle length of 500 frames (10 s) was chosen.

For DNA PAINT, where the bleached molecules bound to DNA can be replaced between the cycles, the oscillations were more dramatic (Figure 4.8C). At high concentrations of 5-SiR-Hoechst, the density of localisations in the first frames was too high for reliable analysis, but in the longer term the number of localisations per frame equilibrated at a higher value. It was thus not clear what concentrations of the dye would yield a maximum number of localisations in combination with a particular cycle length. To optimise these parameters, the cycle length and the 5-SiR-Hoechst concentration were varied together. The number of localisations acquired per second rather than per frame was computed to assess the efficiency of data collection (Figure 4.8D). Despite the time lost, shorter cycles appeared to result in a higher speed of data acquisition. However, bearing in mind high variability between cells, the difference was

not dramatic.

## Number and order of slices

In a real experimental setting, the DH-PSF is often brighter and easier to detect at some z-coordinates and dimmer at others, resulting in an uneven data density throughout the slice (Figure 4.9A). Thus, although the 6-8  $\mu\text{m}$  nucleus could be covered by just three slices, a larger number of overlapping slices was required to even out the density of the data. Furthermore, every point within the nucleus had to be covered by the same number of slices, i.e. the slice thickness needed to be a multiple of the displacement between the neighbouring slices. Imaging 12 slices, each cropped to 3.2  $\mu\text{m}$ , with a z-step of 0.8  $\mu\text{m}$ , was therefore chosen as a suitable strategy. This meant that the total thickness covered was equal to 12  $\mu\text{m}$ , and within the middle 8.8  $\mu\text{m}$  every point was covered by exactly four slices (Figure 4.9B). This strategy evened out the data density distribution throughout cells (Figure 4.9C).



**Figure 4.9:** A. Variation in the brightness of the DH-PSF in  $z$  results in an uneven density of data across a slice (left) and a cell, if too few slices are imaged (right). B. The imaging strategy adopted to even out the data coverage. 12 3.2  $\mu\text{m}$ -thick slices, shifted by 0.8  $\mu\text{m}$  (slices 1 and 12 highlighted), are imaged. Within 8.8  $\mu\text{m}$ , every point is “covered” by four different slices, resulting in uniform data density. C. Example profiles of data density across a slice from an experimental SOX2-Halo 3D PALM, generated by averaging the data from all 12 slices (left). The profiles of 5 different cells are similar, indicating that this profile is not due to biological variability. Concatenation of 12 slices results in reasonable data densities across each cell, reflecting the cell thickness (right). D. To maximise localisation replenishment, slices are imaged in an order that maximises the distance between the subsequent slices.

Moreover, because the off-target bleaching is most prominent in the slice neigh-

bouring the one currently being imaged, the acquisition order was designed such that the subsequent slices were as far apart as possible. For  $N$  slices, the imaging sequence within one cycle therefore was  $1, N/2 + 1, 2, N/2 + 2, \dots, N/2, N$ , if  $N$  was even, and  $1, (N + 1)/2 + 1, 2, (N + 1)/2 + 2, \dots, (N + 1)/2 - 1, N, (N + 1)/2$ , if  $N$  was odd. For the standard imaging of 12 slices, the data acquisition proceeded as  $1, 7, 2, 8, 3, 9, 4, 10, 5, 11, 6, 12$  (Figure 4.9D).

### 4.5.2 Accuracy of stitching: beads control

To join multiple slices into a single image, after the localisations are extracted and tracked, the relative offset is added to their coordinates.

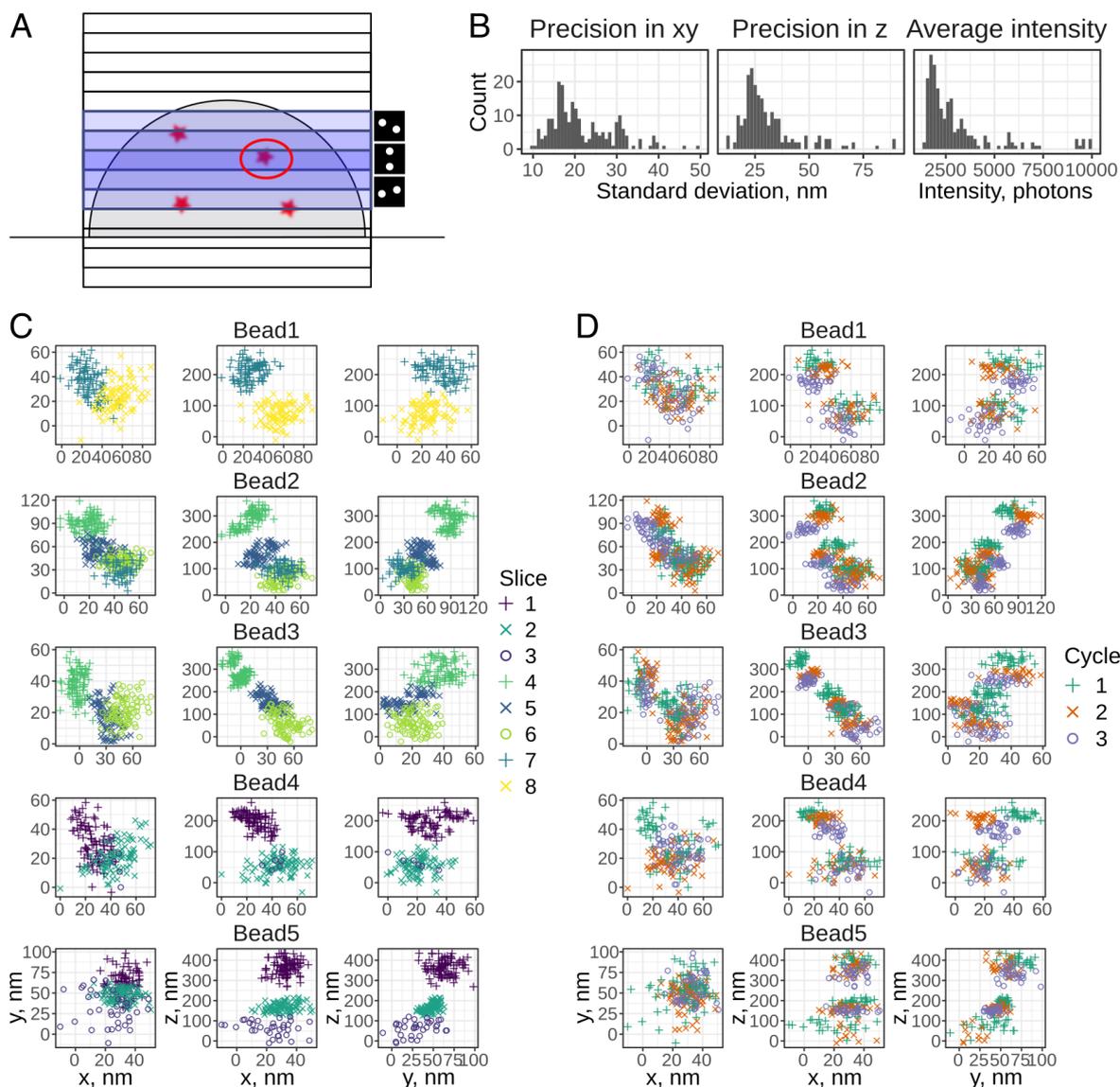
A control dataset with fluorescent beads was recorded and analysed to verify that the entire procedure generated valid results. Beads were suspended in agarose (Figure 4.10A) and imaged: 12 slices were recorded with a  $0.8 \mu\text{m}$  step, in 3 cycles of 30 frames each. Thus, the position of every bead could be determined from its images in different slices and compared. Furthermore, the localisation precision and the precision of the stage movement could also be measured.

#### Localisation precision

Firstly, using the long tracks obtained from fluorescent beads in this experiment, the localisation precision of the method could be calculated accurately. To this end, the standard deviation of the position (i.e. precision) and the average intensity of localisations within each track were determined (Figure 4.10B). On the average ( $\pm$  standard deviation) precision was  $21.4 \pm 7.7$  nm in  $xy$ , and  $31.7 \pm 13.3$  nm in  $z$ . This was about about two-fold better than that measured for individual PA-JF<sub>646</sub> molecules (Figure 4.6C) due to higher signal intensity.

#### Agreement between slices

Next, the coordinates of the same beads, obtained from different slices, were compared. The position estimates did not match well: in the five examples studied, the average



**Figure 4.10:** A. Schematic of the experiment. Fluorescent beads (red stars) are suspended in agarose and imaged, using the whole-nucleus DH-PSF SMLM protocol. Each bead's coordinates are then determined from its images in different slices: e.g. the bead circled appears in three different slices (highlighted in blue), having a different PSF in each of them (illustrated on the right). B. Histograms of localisation precision in  $xy$  (left) and in  $z$  (middle), defined as the standard deviation of the coordinates of the localisations belonging to the same track ( $>4$  positions). Histogram of average intensity for each track (right). C. Examples of bead localisations, coloured by the slice in which they appear. Localisations in each row correspond to the same bead (5 beads are shown in total), and are shown in three projections ( $xy$ ,  $xz$  and  $yz$ ). D. Same as C, but coloured by the cycle number in which the localisation was detected.

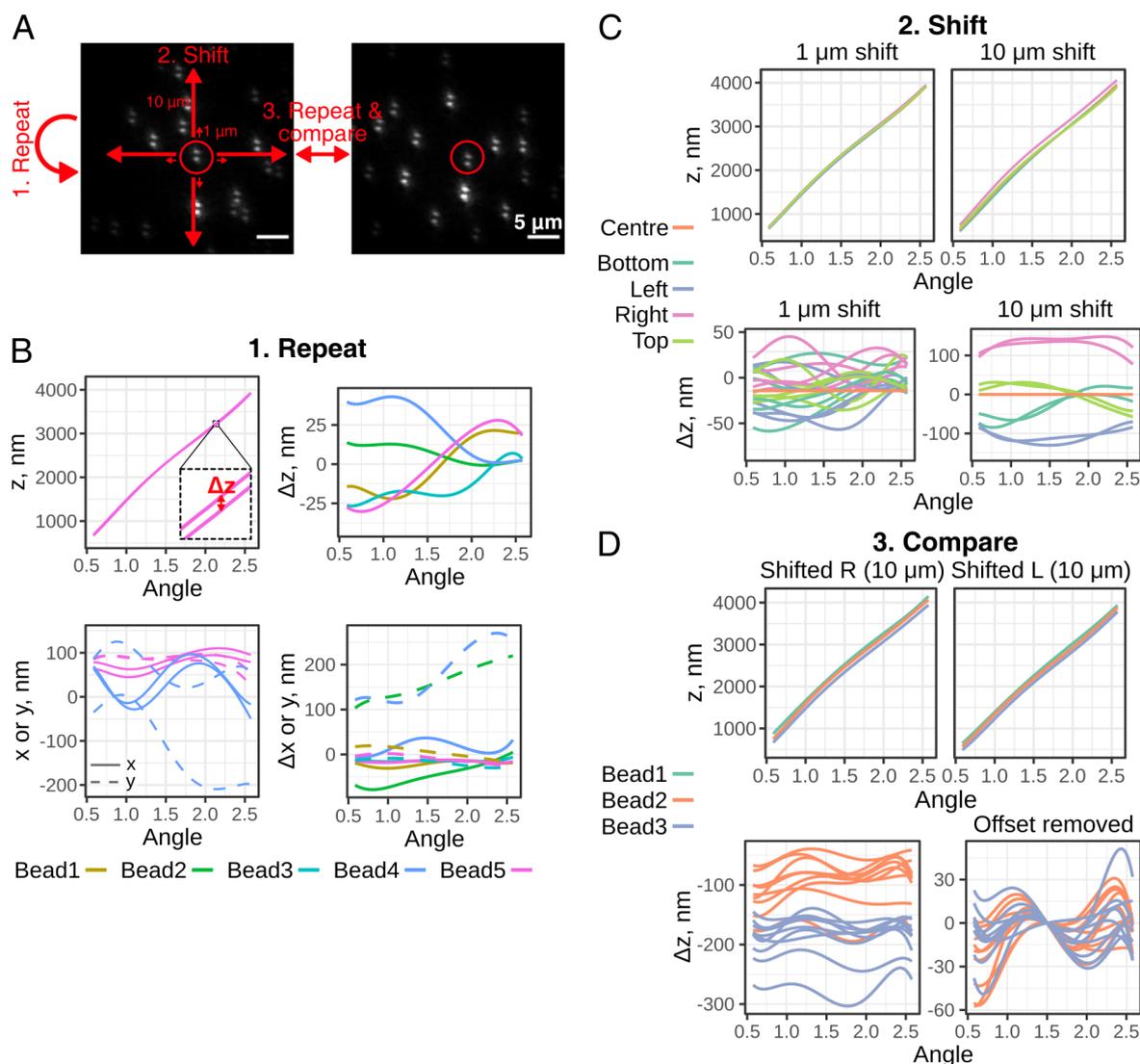
shift between slices was  $21.8 \pm 11.0$  nm in  $xy$  and  $107.5 \pm 55.1$  nm in  $z$ , which for  $z$  was substantially larger than the localisation precision (Figure 4.10C).

To test whether this could be explained by imprecise stage movement, localisations from different cycles were compared, since in each cycle, the stage had to return to the original position (Figure 4.10D). Some inaccuracy was present, e.g. for beads 2 and 3. However, the magnitude of the shifts was small (on average,  $13.1 \pm 5.6$  nm in  $xy$  and  $28.9 \pm 15.2$  nm in  $z$ ), and thus it was unlikely it was the source of the large discrepancies between slices.

### **Spatial variation in DH-PSF**

Another potential source of the discrepancies could be inaccurate mapping between the double helix pitch and the molecule's  $z$ -coordinate. Such an issue could be caused by random variation in the DH-PSF shape between different beads (e.g. if beads are deformed or poorly dispersed), some factor randomly distorting the calibration sequence (e.g. drift or jolts in stage movement), or spatial variation in the DH-PSF across the field of view due to spherical aberration.

To test these possibilities, calibration sequences were acquired repeatedly for several sets of beads (Figure 4.11A). For the same bead at the same position, the polynomial fits of  $z$ -coordinate vs angle were reproducible within about 25 nm, by sometimes drift in  $xy$  during acquisition was apparent, e.g. for bead 4 (Figure 4.11B). Recording the same bead in different regions of the field of view revealed reproducible systematic biases in the DH-PSF shape, which were more pronounced across larger distances (Figure 4.11C). Finally, calibrations from different beads placed into the same position within the field of view were generally similar (within about 30 nm), but displayed an offset, meaning that the calibration series started from slightly different  $z$ -positions (Figure 4.11D). The latter could be the result of unevenness of the coverslip over large distances, which should not impact experiments with cells. Therefore, overall the major factor in the inaccuracy of the DH-PSF method in mapping the  $z$ -coordinate to the PSF angle appeared to be the spatial variation of this mapping across the field of view.



**Figure 4.11:** A. Calibration sequences were acquired: 1. repeatedly for the same bead in the same position to measure reproducibility; 2. for the same bead, but shifted 10 or 1  $\mu\text{m}$  in every direction within the field of view; 3. for a different bead far away on the coverslip, placed roughly in the same position within the FOV. Shown are example images of beads. Scale bar, 5  $\mu\text{m}$ . B. Results of repeated measurements: examples of  $z$  vs DH-PSF angle (top left) or midpoint  $xy$  shift vs angle (bottom left) fitted to the replicate data from the same bead. The calculated differences in the fitting between replicates ( $\Delta z$  and  $\Delta x$  or  $\Delta y$ ) are shown for 5 different beads (right). C. Comparison between the same beads, shifted within the FOV: example polynomial fits of  $z$  vs angle (top) and  $\Delta z$  between the fits for the same bead upon shifting (bottom). D. Comparison between calibrations from 3 different beads in the same position in the FOV: example polynomial fits of  $z$  vs angle (top) in two positions and  $\Delta z$  for beads 2 and 3 with respect to bead 1 in all positions, with or without an offset correction (bottom).

## 4.6 Drift correction

Imaging for prolonged periods of time is often associated with substantial movement of the microscope setup or the sample. Therefore, characterising how much the system drifts during acquisition and correcting for it is crucial to achieve high resolution.

Two common methods for drift correction during long-term imaging involve parallel imaging of a fiducial marker or correction based on tracking a feature of the object itself. The former works by tracking the movement of a reference fluorescent bead throughout the experiment. The latter involves finding a well-defined feature in the super-resolution data, such as a filament or the edge of the nucleus, that can be detected using even a small subset of the data. The displacement of this feature over time is evaluated and used for correction.

Unfortunately, neither of these approaches were suitable for 3D DH-PSF SMLM. Since the cell is about 6-8  $\mu\text{m}$  thick, the beads located at the bottom of the well are too defocused when imaging the middle or the top of the cell. This could be remedied by suspending beads in agarose, but this approach was not compatible with the subsequent snHi-C steps. Alternatively, a wide-range phase mask, such as tetrapod (Shechtman et al. 2015), could be used in the second channel, but this would require further complication of the setup. On the other hand, proteins in the nucleus may not necessarily display any particular features that can be detected using a subset of the data.

Therefore, an approach using bright-field images of cells was developed, similar to Mennella et al. (2012). This method was pioneered by Dr Aleks Ponjavic, and was then thoroughly characterised, extended and modified by me.

### 4.6.1 Evaluation of drift

Firstly, to estimate the magnitude of drift in  $xy$  and  $z$  during a typical experiment, fluorescent beads were imaged for extended periods of time (Figure 4.12A). The microscope setup appeared to be rather stable, with a drift of  $<200$  nm observed in  $xy$  over 30 min of imaging. Almost no drift was observed in  $z$  owing to the Perfect Focus

system. Therefore, a method for drift correction only in xy was required.

### **4.6.2 Drift correction with bright-field image cross-correlation**

The drift correction method was based on capturing the bright-field images of cells in parallel with super-resolution imaging. To achieve this, a 530 nm monochromatic lamp was fitted over the stage, and the transmitted light was separated from fluorescence emission by a dichroic mirror and detected by a second camera (Figure 4.2). Test imaging showed that there was no leakage into the red channel, which could have compromised the quality of the super-resolution data.

The bright-field images were then matched using cross-correlation, and the calculated shift over time was used for correction (see section 6.6.3).

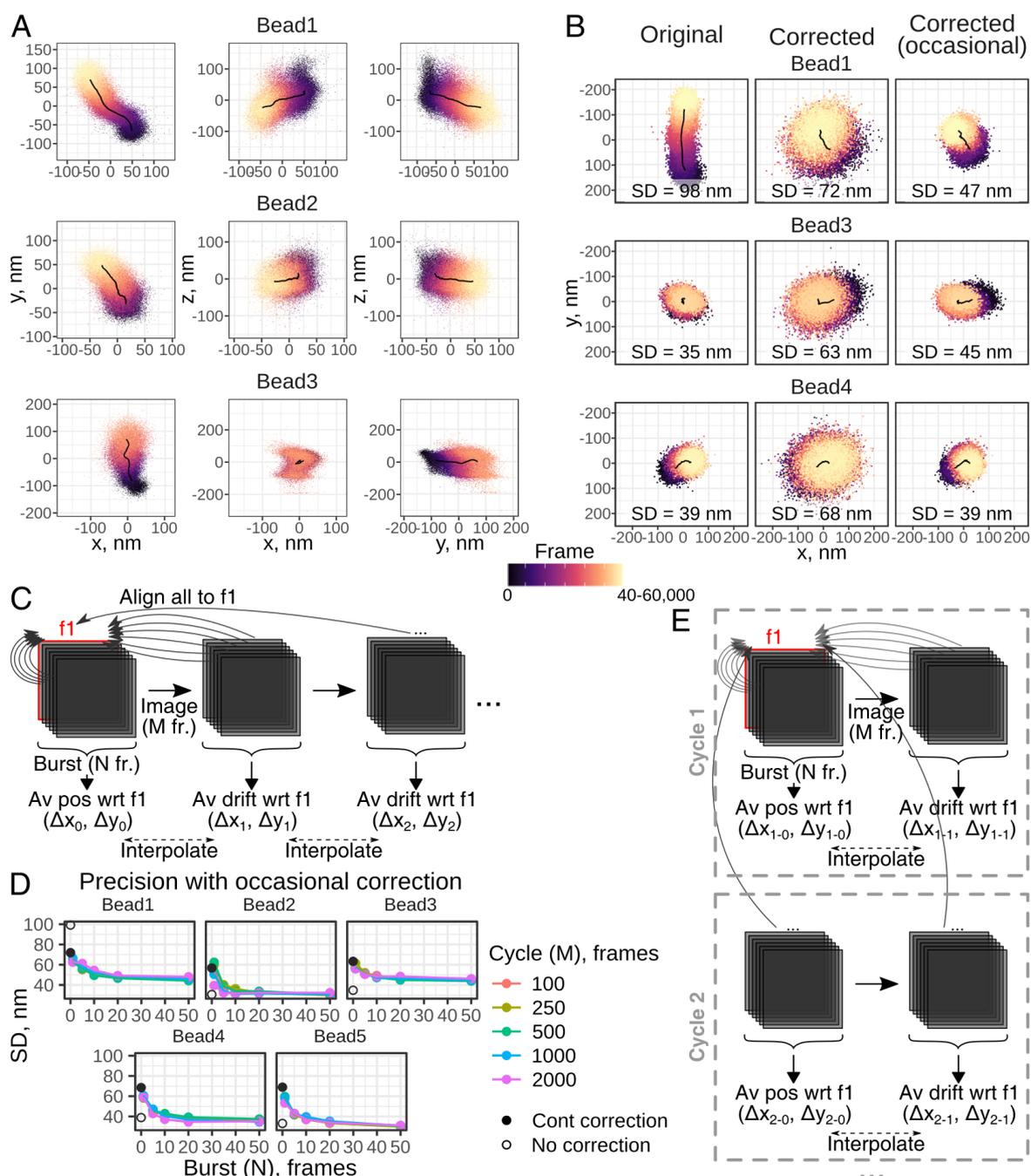
#### **Continuous drift correction**

To test this approach, beads were imaged in 2D alongside fixed mESCs. Furthermore, since super-resolution data is acquired in slices, it was important to check that this method worked with defocused cell images. Thus, fields of view with beads located at different heights (due to landing onto the surface of the coverslip or a cell) were tested. Unfortunately, due to the lower precision of cross-correlation compared to SMLM, drift correction worsened the precision in cases when the drift was not very substantial (Figure 4.12B).

#### **Occasional drift correction with averaging**

Since the stage drift was always slow and smooth, obtaining a corresponding bright-field image of cells for every frame of the super-resolution data was not necessary. An image could be taken occasionally, and movement interpolated in between. Moreover, such a strategy allowed capturing several images in a row and averaging the calculated drift to obtain a more precise value (Figure 4.12C).

This approach was tested on the same dataset, varying the frequency of sampling ( $M$ ) and the number of frames averaged ( $N$ ) (Figure 4.12D). Excitingly, occasional



**Figure 4.12:** A. Examples of beads, imaged with DH-PSF for 50,000-60,000 frames, displaying drift. The data is shown in three projections. The black line is a rolling average across 10,000 frames. B. Examples of beads added to a cell sample, imaged in 2D for 40,000 frames and drift-corrected using the bright-field method. The left column shows the original localisations, the middle presents the drift-corrected data, and the right column has data, drift-corrected using the occasional method with averaging ( $M = 500$ ,  $N = 20$ ). SD, standard deviation of the localisations. C. Illustration of the occasional drift correction with averaging. A burst of  $N$  bright-field frames is taken every  $M$  frames of the super-resolution imaging. All bright-field images are aligned to the first frame ( $f_1$ ) using cross-correlation. The shift for every burst is averaged, and the drift correction for  $M$  frames of the super-resolution series between two bursts is computed by linear interpolation.

*Legend continued on the next page.*

drift correction with averaging with  $N > 20$  and  $M$  in the range of several hundred significantly improved the precision of cross-correlation.

For Bead1, drift correction appeared to work very well, substantially improving the standard deviation of the bead position (Figure 4.12B, top and D, left). However, for other beads this was not so due to a problem with the test dataset: the movement of some of the cells was decoupled from the movement of the beads. Unfortunately, this issue prevented more thorough characterisation of drift correction, e.g. accurate estimation of the improvement in precision. However, the current findings strongly suggest that occasional drift correction with averaging performs well when there is substantial drift and introduces little additional imprecision. Examination of the raw data confirmed that even in cases when the bead moved in a different direction from the cell, the cross-correlation approach captured the direction of the cell drift accurately.

In the real setting, since super-resolution data comes of course from the same cell that is used for drift correction, such decoupling cannot occur. Therefore, the occasional drift correction with averaging appears to be a viable approach for 3D DH-PSF SMLM. Furthermore, it is well-suited for cyclic data acquisition throughout the nuclear volume, since the cycle length lies within the optimum range of parameter  $M$ . The imaging protocol was thus modified to take 20 bright-field images at the beginning and end of each cycle, and the algorithm was also adapted for this regime (Figure 4.12E).

**Figure 4.12:** D. Assessment of localisation precision (standard deviation of position) with varying parameters for occasional drift correction with averaging, compared to no correction or continuous correction. Beads 1, 3 and 4 are also shown in B. E. A schematic of occasional drift correction with averaging, adapted for imaging in cycles. A burst of bright-field images is recorded at the beginning and end of each cycle, and the drift during the acquisition is estimated by linear interpolation. Bright-field frames from subsequent cycles are also aligned to frame 1, allowing one to accurately superimpose data from different cycles.

## 4.7 Trial whole-nucleus 3D DH-PSF SMLM imaging of DNA, transcription factors and HP1

After implementing the described optimisations, the final imaging and Hi-C procedure consists of the following steps (Figure 4.1):

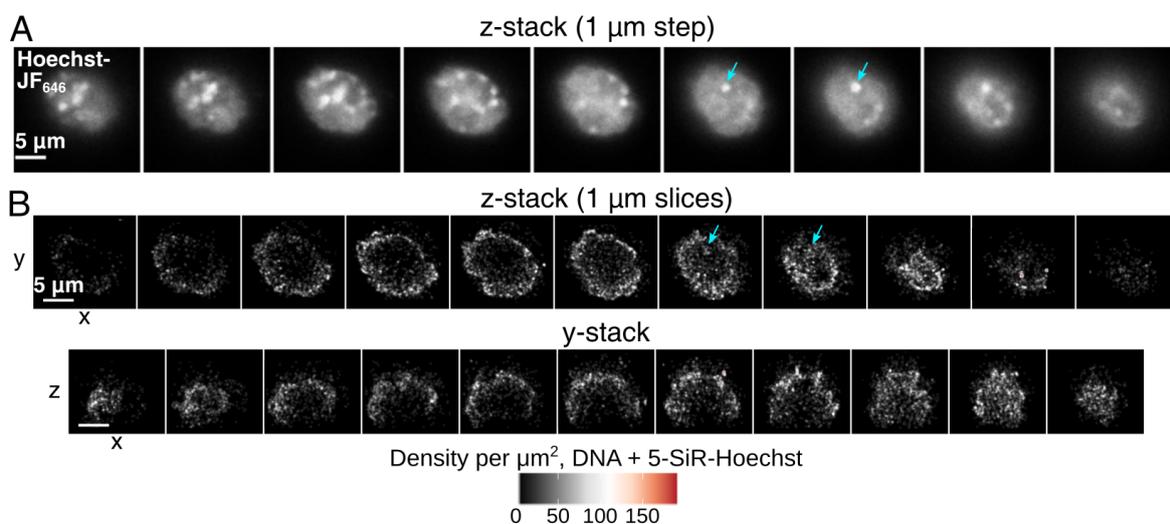
1. Labelling of cells and FACS-sorting into 384-well plates;
2. Alignment of the microscope and DH-PSF calibration sequence acquisition;
3. Locating the cells within wells, using a 10× air objective;
4. Whole-nucleus 3D DH-PSF SMLM (PALM or PAINT) imaging of up to 20 cells. Data was collected for up to 12 overlapping z-slices with a shift of 0.8  $\mu\text{m}$ , in 10-20 cycles, 250-500 frames each. At the beginning and end of each cycle, a burst of bright-field images was recorded for drift correction;
5. Performing the snHi-C protocol;
6. Analysis of SMLM data: localisation detection with GDSC PeakFit + DHPSFU, removal of duplicates, drift correction, rendering and statistics.
7. The procedure for combining information from snHi-C and microscopy for structure calculation is currently under development.

Since experiment and analysis optimisation was performed alongside the trial imaging and Hi-C experiments, some of the datasets were not acquired under ideal conditions, and several issues still remain. In this section, preliminary results are presented, sometimes alongside the corresponding genome structure calculated from the snhi-C data alone, and their drawbacks are discussed.

### 4.7.1 DNA PAINT

Preliminary 3D PAINT images of the entire nuclei showed large hollow spaces in the middle. To unequivocally determine whether they were real (e.g. nucleoli) or artificial, super-resolution microscopy was followed by addition of Hoechst-JF<sub>646</sub> and

imaging diffraction-limited z-stacks for comparison (Figure 4.13A). Unfortunately, creating a 3D density map of the data revealed that the vast majority of localisations originated from the outer shell of the nucleus (Figure 4.13B). This was recapitulated in the majority of cells from the same experiment, and also in replicates. Although some of the previous 2D images appeared to have a rather uniform density of localisations, it was potentially due to 5-SiR-Hoechst molecules located on top of the cells.



**Figure 4.13:** A. Diffraction-limited z-stack of mESC, labelled with Hoechst-JF<sub>646</sub>, recorded using HILO illumination and a z-step of 1  $\mu\text{m}$ . B. Whole-nucleus 3D PAINT images of DNA, rendered as density plots with a bandwidth of 0.5  $\mu\text{m}$ . Each row is a scan through the volume of the cell (along the z- or the y-axis), with one image representing a 1  $\mu\text{m}$ -thick slice. The cyan arrow points to a potentially coinciding focus in both images. Scale bars, 5  $\mu\text{m}$ .

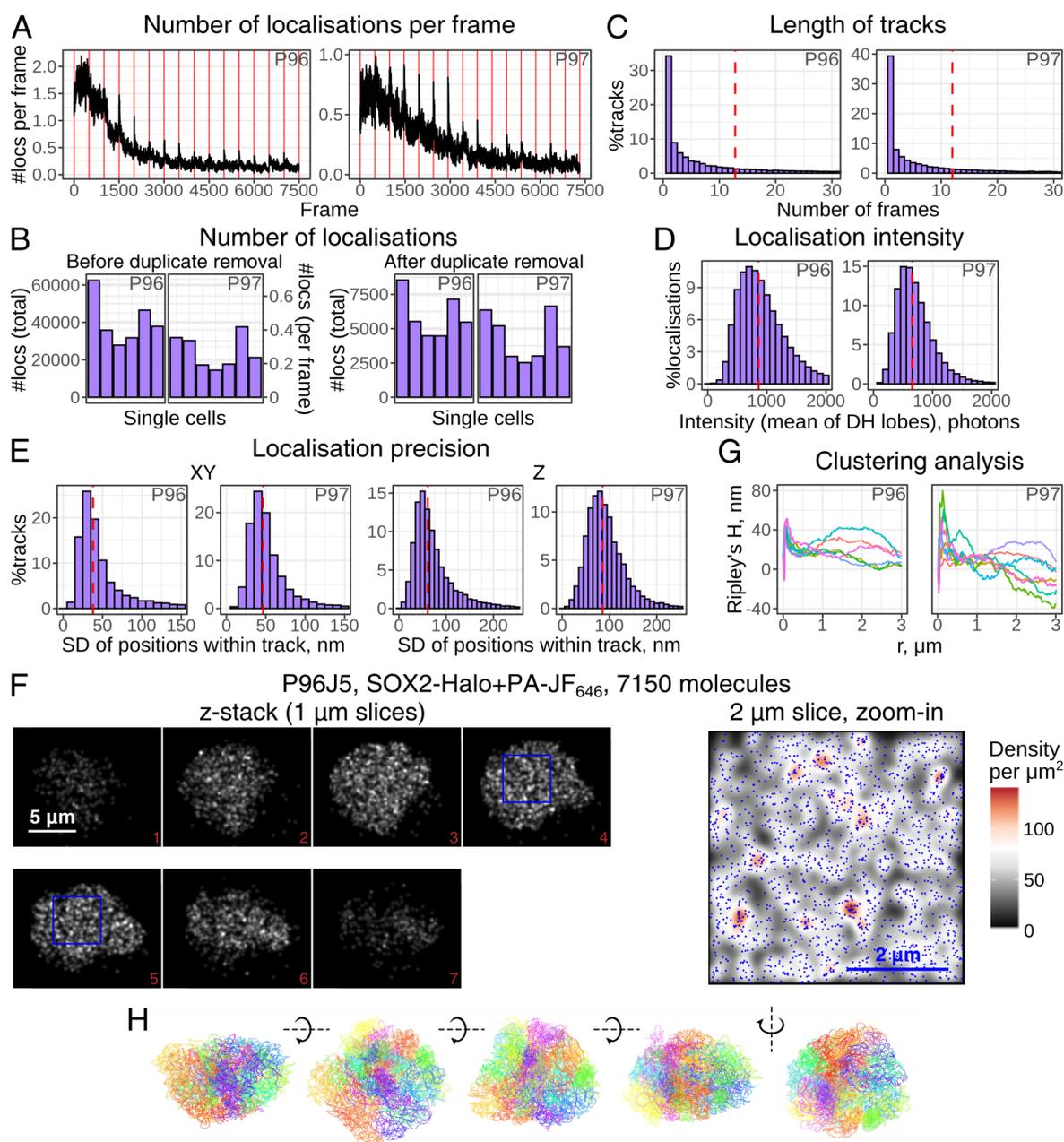
Labelling DNA with high concentrations 5-SiR-Hoechst results in good-quality wide-field images, and the observed phenomenon only arises at very low concentrations of the dye in the single-molecule regime. This suggests that at fast imaging rates, the fluorophore does not have time to penetrate deep into the nucleus before being bleached. This could be the result of 5-SiR-Hoechst potentially binding to the DNA close to the nuclear lamina straight away after getting into the nucleus, increasing in fluorescence and being bleached. Alternatively, its fluorogenicity might become activated in the wrong environment, just outside the nucleus rather than inside. In any case, the pure PAINT approach to imaging DNA did not work with 5-SiR-Hoechst.

## 4.7.2 SOX2 PALM

Two datasets of SOX2 PALM data were successfully collected, and snHi-C was performed on the same nuclei. As expected, the number of localisations decreased exponentially during acquisition (except for the short bursts following stage movement), and 7,500 frames were recorded per slice (Figure 4.14A). The total number of raw localisations varied between cells, and was lower in the second dataset (Figure 4.14B, left). However, removing multiple instances of the same fluorophore reduced the difference, with the final number of molecules varying between 4,500-8,000 and 2,500-6,500 for the two datasets (Figure 4.14B, right). Consistently, fluorophores tended to persist for a shorter time in the second dataset (Figure 4.14C), and the intensity of localisations was also lower (Figure 4.14D).

The multiple instances of the same fluorophore were used to estimate localisation precision (Figure 4.14E). Although not as high as potentially achievable, median values of 38 and 46 nm in *xy*, and 61 and 85 nm in *z* were good enough to make high-resolution maps. Furthermore, the precision was further enhanced by averaging of emitter position from several frames (Sage et al. 2019). The higher localisation error in the second dataset was consistent with the lower intensity.

SOX2 showed a relatively uniform distribution throughout the nucleus (Figure 4.14F). Interestingly, while the shape of the nucleus was clearly distinguishable, no nucleoli, manifested as large areas devoid of SOX2, could be seen in any of the images. This was consistent with confocal images of SOX2 in mESCs in 2i/LIF conditions obtained in the group (Devina Shah, unpublished), but disagreed with previous observations (Veneri et al. 2020). Analysis of clustering at different scales using Ripley’s method (Ripley 1977) showed heterogeneous results, but a consistent peak was present at  $r < 200$  nm, indicating clustering at a small scale (Figure 4.14G). However, this could be the result of multiple localisations of the same molecule that escaped the tracking procedure, and needs to be investigated further. Furthermore, the representativeness of the reconstructed images is currently unclear, and estimating the detection efficiency of the method is crucial to judge that. Higher detection efficiency and more robust



**Figure 4.14:** A. The number of detected localisations per frame throughout acquisition, averaged between all slices and cells, for plates P96 and P97. The red lines separate the 500-frame cycles and indicate stage movement between slices. B. Total number of localisations per cell before and after duplicate removal, and average number of raw localisations per frame for each cell. C. Histograms of the “track” (multiple occurrences of the same emitter) lengths (the red line indicates the mean). D. Localisation precision in xy and z, measured as the standard deviation of the coordinates of the same molecule detected  $>3$  times (the red lines indicate the medians). E. Histograms of intensities of raw localisations. In C, D and F, all cells and all slices were pooled. F. Example SOX2 DH-PSF PALM image of nucleus J5 from P96. Left, 1  $\mu\text{m}$ -thick slices through the nucleus (numbered in order), rendered as 2D density plots with a 500 nm bandwidth. Scale bar, 5  $\mu\text{m}$ . Right, a zoom-in of the area indicated with the blue square. Localisations from two slices were pooled and plotted as blue dots on top of a density map. Scale bar, 2  $\mu\text{m}$ .

*Legend continued on the next page.*

analysis methods are likely required for better assessment of clustering.

Genome structures were calculated for several of the imaged cells, including P96J5, as a proof-of-concept (Figure 4.14H). The algorithms to overlay 3D images and chromatin models are currently under development, although in this case the nuclear shape appears to be similar between the two methods.

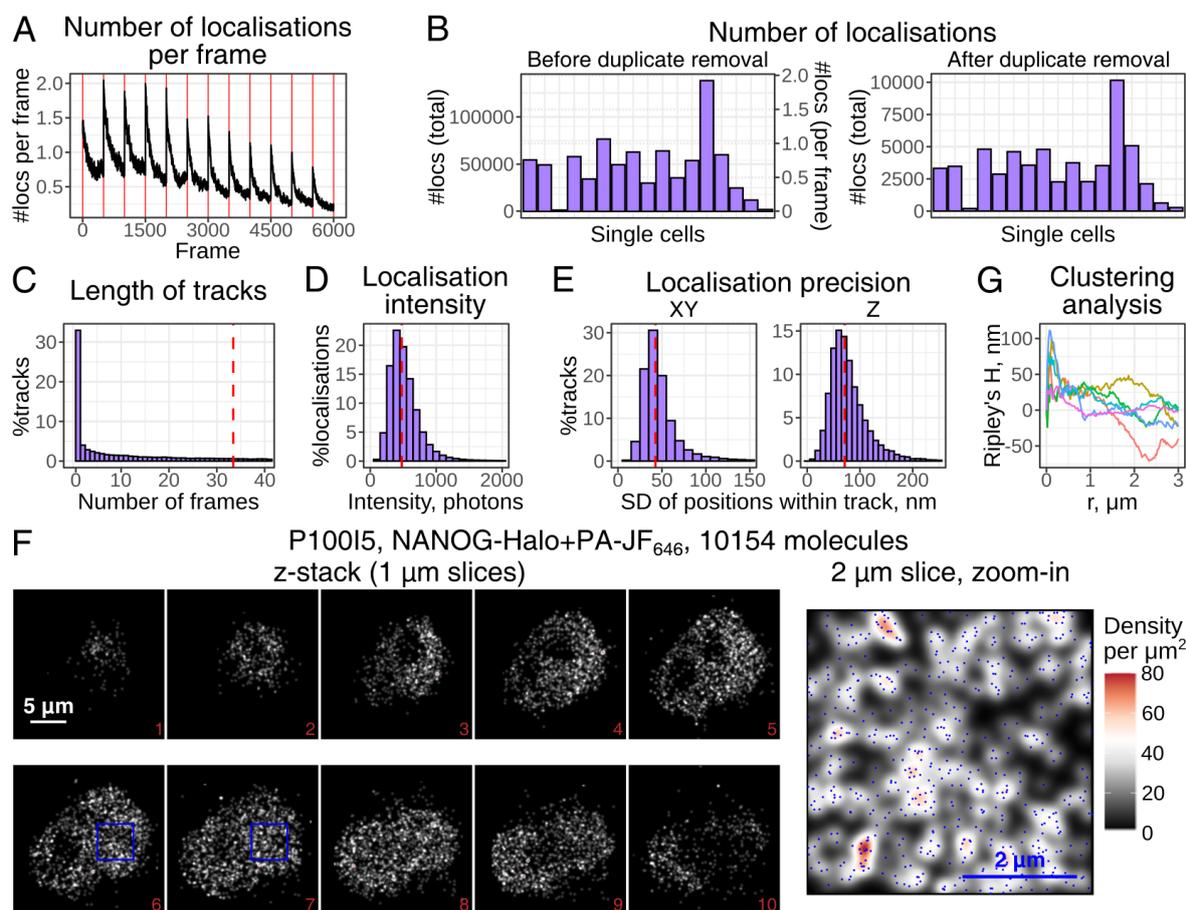
### 4.7.3 NANOG PALM

NANOG PALM was also successfully conducted, followed by the snHi-C protocol. Since NANOG is a less abundant protein than SOX2, and based on an initial assessment of the number of localisations per frame, 6,000 frames were recorded for each slice, while the other imaging settings were kept the same as for SOX2 PALM. However, the number of localisations per frame throughout the experiment displayed unusual dynamics: unlike in previous PALM trials, the peak following the stage movement was very high, and its effect was prolonged (Figure 4.15A). Potentially as a result of this, the number of localisations per frame decreased quite slowly. The total number of raw localisations was variable, but with the exception of two cells, it was higher than for SOX2 imaging, despite the shorter duration of acquisition (Figure 4.15B, left).

The fluorophores in this experiment tended to last even longer than in the SOX2 datasets (Figure 4.15C). Averaging of the multiple localisations of the same molecule thus resulted in a final number of localisations similar to the second SOX2 imaging session (Figure 4.15B, right). The localisation intensity also appeared to drop further compared to the previous experiment (Figure 4.15D), but the estimated precision, 42 nm in x and 71 nm in z, was still acceptable (Figure 4.15E).

Unlike SOX2, NANOG exclusion from nucleoli was obviously visible in the images (Figure 4.15F). The density of NANOG molecules was lower than that of SOX2, even for the P100I5 cell with 10,154 localisations. While this could be the result of

**Figure 4.14:** G. 3D adaptation of Ripley’s H function with edge correction at different radius  $r$ . Positive or negative values indicate clustering or dispersion, respectively. Each line corresponds to a single cell. H. Reconstruction of the genome fold of P96J5 at 1 Mbp resolution, where each chromosome is coloured differently.



**Figure 4.15:** A. The number of detected localisations per frame throughout acquisition, averaged between all slices and cells. The red lines separate the 500-frame cycles and indicate stage movement between slices. B. Total number of localisations per cell before and after duplicate removal, and average number of raw localisations per frame for each cell. C. Histograms of the “track” (multiple occurrences of the same emitter) lengths (the red line indicates the mean). D. Histogram of intensities of raw localisations. E. Localisation precision in xy and z, measured as before (the red lines indicate the medians). In C, D and E, all cells and all slices were pooled. F. Example NANOG DH-PSF PALM image of nucleus I5 from experiment P100. Left, 1  $\mu\text{m}$ -thick slices through the nucleus (numbered in order), rendered as 2D density plots with a 500 nm bandwidth. Scale bar, 5  $\mu\text{m}$ . Right, a zoom-in of the area indicated with the blue square. Localisations from two slices were pooled and plotted as blue dots on top of a density map. Scale bar, 2  $\mu\text{m}$ . G. 3D Ripley’s H function with edge correction at different radius  $r$ . Positive or negative values indicate clustering or dispersion, respectively. Each line corresponds to one cell.

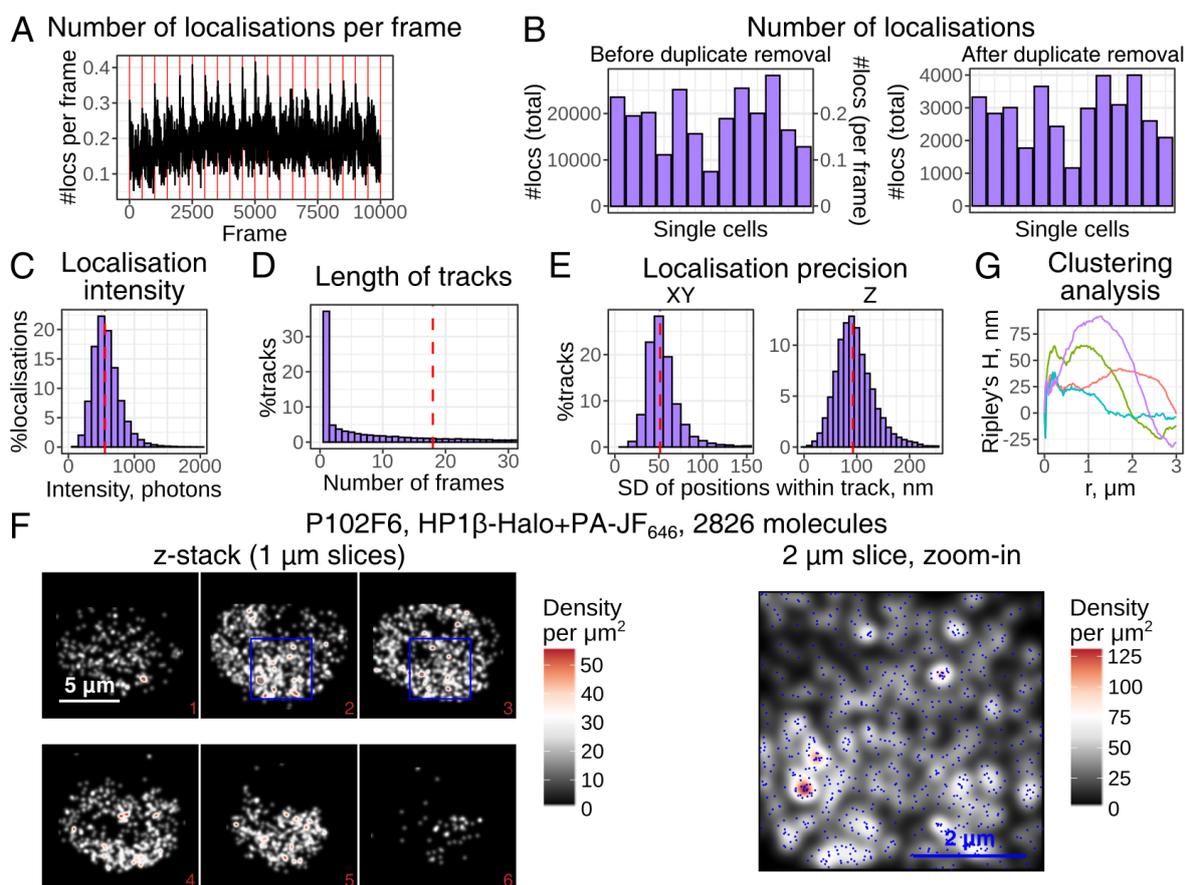
lower abundance of NANOG, it is also in line with lower signal intensity. Thus, an independent estimate of the fractions of detected and undetected molecules is needed to quantitatively compare images of different proteins. 3D Ripley's H functions for NANOG also displayed an initial peak at  $r < 300$  nm (Figure 4.15G).

The Hi-C libraries for this experiment have not yet been sequenced, and thus genome structures have not been generated.

#### 4.7.4 HP1 $\beta$ PALM

Finally, 3D PALM super-resolution images of HP1 $\beta$  were acquired. Due to lower apparent brightness of fluorophores, and following the results of laser intensity trials, the 638 nm excitation laser was used at full power, 2.5-fold stronger than in the previous experiments. Since HP1 $\beta$  is a very abundant protein, and the number of localisations per frame did not appear to drop significantly during acquisition (Figure 4.16A), 10,000 frames were recorded per slice. However, despite that, the raw number of localisations obtained was quite low (Figure 4.16B, left). Examination of the raw data revealed that although 2-3 molecules could typically be seen per frame by eye, they were not registered by the software due to the poor signal-to-noise ratio. Indeed, despite stronger excitation, the brightness of localisations was only slightly higher than before (Figure 4.16C), while the background was significantly higher due to the increased laser intensity. Therefore, it appeared that over time (the SOX2, NANOG and HP1 $\beta$  datasets were acquired in February, May, July and November 2021), the intensity of the PA-JF<sub>646</sub> fluorophore signal in our experiments gradually deteriorated.

Removing the duplicates resulted in 1,000-4,000 molecules per cell, which was lower than in the previous datasets despite higher HP1 $\beta$  abundance and longer imaging time (Figure 4.16B, right). Single fluorophores tended to persist for a rather long time, although comparison with previous experiments is hindered by the altered excitation laser intensity (Figure 4.16D). In agreement with a worse signal-to-noise ratio, the estimated localisation precision in this dataset was 51 nm in xy and 93 nm in z (Figure 4.16E).



**Figure 4.16:** A. The number of localisations per frame during acquisition, averaged between all slices and cells. Red lines separate the 500-frame cycles. B. Total number of localisations per cell before and after averaging of duplicates, and mean number of raw localisations per frame for each cell. C. Histogram of intensities of raw localisations. D. Histograms of the “track” lengths (the red line indicates the mean). E. Localisation precision in xy and z, measured as before (the red lines indicate the medians). In C, D and E, all cells and all slices were pooled. F. Example HP1 $\beta$  DH-PSF PALM image of nucleus F6 from plate P102. Left, 1  $\mu$ m-thick slices through the nucleus (numbered in order), rendered as 2D density plots with a 500 nm bandwidth. Scale bar, 5  $\mu$ m. Right, a zoom-in of the area indicated with the blue square. Localisations from two slices were pooled and plotted as blue dots on top of a density map. Scale bar, 2  $\mu$ m. G. 3D Ripley’s H function with edge correction. Positive or negative values indicate clustering or dispersion, respectively. Each line corresponds to a single cell.

Despite the low number of molecules, the shape of the nucleus and the nucleoli could be clearly seen in the images (Figure 4.16F). HP1 $\beta$  distribution also appeared to be less uniform than that of either SOX2 or NANOG, which was confirmed by Ripley's analysis (Figure 4.16G). Indeed, HP1 $\beta$  clustering on  $\sim 1$   $\mu\text{m}$  scale is consistent with formation of heterochromatic domains. However, due to the very low detection efficiency in this experiment, the foci could not be observed clearly in the images. The deterioration in localisation intensity and, as a consequence, low sampling rate in whole-nucleus 3D PALM is certainly a major factor that must be improved.

## 4.8 Discussion and future directions

In this chapter, the work leading up to 3D super-resolution imaging of DNA or proteins in the entire nuclear volume of mammalian cells has been described, as well as the proof-of-concept imaging of SOX2, NANOG and HP1 $\beta$  proteins combined with snHi-C.

Firstly, the labelling protocol and imaging conditions for either DNA PAINT or protein PALM microscopy were optimised. Secondly, three DH-PSF localisation detection software packages were characterised, and the best-performing algorithm was selected. Then, challenges associated with stitching data obtained by scanning the cells were addressed, and a suitable imaging procedure was developed. However, a control experiment identified inconsistencies in determining the coordinates of molecules using DH-PSF, which should be corrected in the future. Finally, a method for xy drift compensation using cross-correlation of bright-field images was developed and assessed. In parallel, the DH-PSF microscope built by Dr Aleks Ponjavic was adapted by Ziwei Zhang to include the above optimisations and perform SMLM of the entire nuclei in 384-well plates.

Volumetric DNA PAINT imaging revealed that at low concentrations of 5-SiR-Hoechst, the dye could not diffuse into the interior of the nuclei before being bleached, and thus the localisations were concentrated at the nuclear periphery. The approach to label and image DNA must therefore be modified. Protein PALM microscopy also indicated a technical issue that resulted in a drop in localisation intensity over time. A

more fundamental challenge for this method, however, is measuring and potentially enhancing the detection efficiency. The completeness and localisation precision estimates should also be combined to give a final approximation for image resolution.

Obtaining the 3D SMLM reconstructions of SOX2 distribution and the genome models for the same cells was a milestone, proving the possibility and feasibility of such an experiment. While this preliminary data is unlikely to provide biological insight, it opens up avenues for studying the link between chromatin-binding proteins and genome folding in the future.

### 4.8.1 DNA SMLM

Measuring the DNA density distribution throughout the nucleus could be a most useful experiment for improving genome structure calculations. This information in the form of a force field could instruct the model directly, while using protein images for this purpose will require additional ChIP-seq data.

Confocal imaging of DNA is the most straightforward way to obtain this information, and is pursued by other colleagues in the lab. However, super-resolution microscopy could yield much more fine-grained results, in particular in resolution in  $z$ , and thus potential ways of achieving this are discussed here.

Since high concentrations of 5-Sir-Hoechst label chromatin uniformly, and uniform labelling is also observed at the beginning of the PAINT experiment, the problem lies in the dynamics of dye binding and bleaching rather than in its steady-state distribution. Using an SMLM modality that relies less on binding kinetics could therefore circumvent this problem. However, abandoning the PAINT approach entirely would deprive us of the benefits of fluorophore replenishment, which is a big advantage when imaging huge structures like chromatin.

A combination of blinking and binding could help reap the benefits of both methods: the blinking behaviour would allow one to use higher concentrations of the dye, with the majority of fluorophores bound to DNA being “off”, and the bleached fluorophores would still in time be replaced by fresh ones from solution. This would require using

a photoactivatable or spontaneously blinking, ideally fluorogenic, DNA-binding dye with appropriate binding kinetics. Fortunately, multiple such organic fluorophores have been developed. 5-HMSiR-Hoechst increases in fluorescence intensity >400-fold upon binding DNA, and unlike 5-SiR-Hoechst, displays spontaneous blinking when bound to DNA due to the intramolecular transitions between the “closed” and “open” forms (Bucevičius et al. 2020). This dye has been used at high concentrations for DNA SMLM imaging in live cells, and is thus very promising. Spontaneously blinking HM-JF<sub>526</sub> has been successfully employed in SMLM immunofluorescence experiments (Zheng et al. 2019). Its non-blinking analogue, conjugated to Hoechst, was found to be fluorogenic (Zheng et al. 2019). Thus, HM-JF<sub>526</sub>-Hoechst could hypothetically also be a good candidate for DNA SMLM.

Should the direct DNA imaging prove challenging – e.g. by affecting the subsequent stages of snHi-C – SMLM of histone proteins could alternatively be employed to measure chromatin density throughout the nucleus. However, this approach appears less promising, due to the potentially much lower coverage and the difficulty and potential side effects of tagging histones.

## 4.8.2 Accuracy of DH-PSF

Control experiments matching the 3D coordinates of suspended beads, obtained from different slices, and subsequent work revealed spatial variation in the correspondence between the DH-PSF pitch and the z-coordinate. Furthermore, additional issues, such as inaccuracies in stage movement and occasional drift during recording of the calibration sequence were uncovered. Together, these factors can severely affect localisation precision, and thus need to be addressed.

### Spatial variation

The spatial variation in angle vs z-coordinate mapping could be the result of either the inherent optical factors, such as spheric or coma aberration, or the misalignment of some components of the optical pathway. Firstly, the microscope setup must be

checked to identify and fix any potential sources of the variation. Z. Zhang found that adjusting the phase mask tilt position in the optical path decreased the spatial heterogeneity. Furthermore, the objective axis was found not to be perfectly vertical, which needs to be adjusted by the manufacturer.

On the other hand, computational approaches can also be used to compensate for some equipment limitations. In a simple scenario, if the misalignment of microscope components leads to a planar tilt in the image, it should be possible to measure this tilt, using a beads sample, and to correct for it. In a more problematic case, the geometry might be more complex and/or change at different z-positions. A calibration sequence with multiple beads in the field of view could in that case be used to generate a polynomial field, describing the dependence of the z-coordinate on the DH-PSF angle and x- and y-coordinates of the emitter, instead of a monovariate function of z vs angle.

### **Drift and stage inaccuracy**

The change in midpoint xy coordinate of the double helix in the calibration sequence is a convolution of the intrinsic shift due to the DH-PSF setup (that is one of the parameters that must be determined using calibration) and random drift during acquisition. The latter presents a problem, since it masks the former. Surprisingly, drift of a similar magnitude to long-term imaging (1h) sometimes happens during the quick calibration sequence (2-3 minutes). This might be the result of more vigorous stage movement in z, and thus optimising the calibration procedure (e.g. moving slower or, conversely, faster) might improve it. Since drift does not happen every time, while the xy shift is reproducible, recording replicate calibration sequences and discarding the outliers can also mitigate the impact of drift. Ultimately, however, it would be best to be able to monitor drift explicitly – e.g. by imaging beads simultaneously in another channel without the phase mask.

The inaccuracy in stage movement in z is harder to address. While the Perfect Focus system keeps the stationary focus exceptionally well, it cannot compensate fully for the imprecise movement of the mechanical stage. Replacing the mechanical with a piezoelectric stage is probably the only real solution, but it is unclear whether it can

be done for the current setup. However, the inaccuracy is not large ( $28.9 \pm 15.2$  nm in z), and moving in small increments might help reducing stage backlash.

Once the hardware and software solutions to the current issues are developed, the experiment with suspended beads must be repeated to assess the effect of the modifications.

### 4.8.3 Image completeness and resolution

Increasing the number of molecules detected is perhaps the most important issue that must be addressed to improve the final image quality. For our purposes a very high completeness may not be necessary, but it must be sufficient to obtain an accurate and high-resolution representation of the density of molecules in particular regions.

#### Assessment of image completeness and resolution

Detection efficiency can be assessed directly, using particular structures known as SMLM standards. Their defined shape and composition allows segmentation in SMLM images and comparison of the number of detected localisations to the reference number of molecules. Such standards include, for example, well-characterised bacterial homo-oligomeric protein complexes (Finan et al. 2015) and DNA origami (Zanacchi et al. 2017).

However, these controls are particularly useful when they can be imaged at conditions closely resembling those for the protein of interest, i.e. inside cells, and in parallel with the PALM experiment rather than beforehand. To this end, the nuclear pore complex (NPC) is a very convenient reference. GFP-expressing mammalian cell lines have been developed with components of the NPC tagged with HaloTag, which can be mixed with the cells of interest and enable measurement of detection efficiency *in situ* (Thevathasan et al. 2019). In principle, an mESC line with similar features could also be generated.

The concept of resolution of an SMLM image is somewhat elusive, due to it being determined both by the completeness of the image and the localisation precision. As a

result, it may sometimes be hard to judge whether a particular SMLM image represents the underlying structure faithfully. One way to assess that is the Fourier ring correlation analysis (Nieuwenhuizen et al. 2013; Banterle et al. 2013). To this end, the SMLM dataset is separated into two subsets, and the correlation between the two at different scales is measured, using Fourier transform. This approach could be applied more easily to DNA or HP1 $\beta$  images that form well-defined structures in the nucleus, and could potentially also be used for NANOG and SOX2.

### **Increasing detection efficiency in PALM**

Potential sources of suboptimal detection efficiency include incomplete labelling, premature bleaching of molecules, and loss of localisations due to poor signal-to-noise ratio. These considerations have been addressed, as much as possible, in the trial experiments in this chapter. The decrease in localisation intensity between experiments throughout 9 months and the concomitant drop in the number and precision of localisations might be the result of a gradual deterioration of either the reagents in use – e.g. the PA-JF<sub>646</sub> dye – or the microscope hardware, e.g. the 638 nm excitation laser. Trial experiments with freshly resuspended PA-JF<sub>646</sub> and potentially a comparison to another 638 nm laser must be done to identify the root of the problem and to subsequently guarantee the best possible quality of the data.

In addition, the detection efficiency could be improved by employing thin light-sheet illumination, rather than HILO, which would drastically decrease the out-of-focus bleaching of molecules. Such a microscope is currently under development in the group.

Furthermore, the nature of PALM is that the rate of data collection drops in time as fluorophores are depleted. Thus, another potential way to increase the final number of localisations could be to gradually increase the activation laser intensity during acquisition.

#### 4.8.4 Clustering analysis

Clustering analysis is a useful instrument to extract meaningful information from SMLM reconstructions. Detecting a protein’s ability to self-associate and characterising the typical size and abundance of the clusters might give clues to its mechanism of action. Combining it with snHi-C-derived genome models will provide an additional layer of information, allowing to see which sequences the proteins tend to cluster at and how this relates to chromatin structure (A/B compartments, intermingling, TAD boundaries, etc).

#### Deduplication and controls

Although in the majority of cases, multiple instances of the same molecule are removed by the tracking procedure, some duplicates may escape this procedure if the molecule is not detected in  $>50$  frames. Furthermore, the current approach does not take into account molecules that might appear in different slices, since they are difficult to detect due to stitching inaccuracies. Thus, the first steps to address this problem are to increase the intensity of localisations, as described above, to improve the accuracy of DH-PSF and to adapt the tracking procedure to multiple slices. Next, existing methods for blinking correction (Hummer et al. 2016) can be used to check whether any pseudo-clusters remain and to correct for them. Alternatively, control experiments could be performed with SMLM standards to assess the prevalence of multiplets, or distribution of an inert probe such as the HaloTag could be analysed with Ripley’s method to generate a baseline distribution.

#### Cluster segmentation

The Ripley’s H function is a great first-line analysis, allowing one to determine whether there is any appreciable clustering on different spatial scales. However, the original function does not perform well at edges. Although edge correction techniques exist, they are hard to apply to the curved 3D surface of the nucleus. Furthermore, especially for NANOG and HP1 $\beta$ , it is unclear how to deal with void spaces within the nucleoli.

Therefore, further developments are needed to apply Ripley’s cluster analysis to the current and future whole-nucleus SMLM data in a rigorous way.

While Ripley’s H function provides a global statistical overview of clustering, it does not segment the clusters or provide information about their shape, location and heterogeneity directly. Segmentation, however, is necessary to determine the correlation between protein clustering and genome sequence and folding. Nearest-neighbour algorithm and hierarchical clustering, DBSCAN and Voronoi tessellation-based methods can all be applied to this end (Khater et al. 2020), and their performance will have to be assessed and adapted to the 3D SMLM data.

#### **4.8.5 Alternatives to 3D DH-PSF SMLM**

3D DH-PSF SMLM is a powerful technique that, due to its single-molecule nature, in addition to high resolution allows counting and cluster analysis of proteins. However, alternative methods could be considered, if the problems described above prove hard to resolve.

##### **3D tetrapod PSF SMLM**

In addition to double helix, other PSF engineering methods exist. In the current context, the tetrapod PSF approach could be beneficial due to its large z-range of up to 20  $\mu\text{m}$  (Shechtman et al. 2015). Employment of this method could therefore alleviate the need for scanning and stitching to obtain images of entire nuclei. However, other factors will have to be taken into consideration. Firstly, the increased z-range results in a decreased localisation precision. Secondly, the larger size of tetrapod PSF compared to double helix leads to lower signal intensity, which would further affect precision. Furthermore, a smaller number of molecules has to be imaged per frame to avoid overlaps. Finally, to cover the entire nucleus, the epifluorescence illumination regime might be required in place of HILO, which could result in higher photobleaching rates and background intensity.

### 3D SIM or STED

A different super-resolution technique could be adopted in place of SMLM. 3D SIM, although providing inferior resolution compared to SMLM, is however free of the detection efficiency concerns and much faster. Unfortunately, no 3D SIM instrument capable of imaging multi-well plates is currently available on site. STED, on the other hand, is comparable to SMLM in terms of resolution. Generally, due to extremely strong laser intensities, out-of-focus bleaching prevents imaging multiple planes within the same nucleus with consistent resolution and brightness. However, employment of exchangeable and photostable probes such as JF<sub>646</sub>-Hoechst or 5-SiR-Hoechst could help overcome this issue, particularly for DNA imaging (Spahn et al. 2018; Bucevičius et al. 2019). Although a suitable 3D STED instrument is also not currently available on site, this approach could be a viable alternative in the future.

#### 4.8.6 Image and snHi-C structure correlation

The initial challenge in correlating the snHi-C genome models and 3D images is to find their relative orientation. Since genome folding structures are determined by the number of Hi-C contacts retrieved for different regions of the genome, the shape of the final model may not be entirely accurate and might require stretching or compression in different directions. The presence of nucleoli, as well as the lack of sequencing coverage in the repetitive regions further complicate this task, although the latter issue is currently being addressed by obtaining a more complete and accurate reference genome, using PacBio sequencing.

Although the computational challenge of combining the snHi-C structures and microscopy images is beyond the scope of this thesis, this first step can be aided by imaging methods. In particular, identifying the positions of one or several loci in the nucleus could enable one to align the two. Since FISH is too invasive to be combined with Hi-C due to the formamide- and high temperature-induced denaturation steps, the recently developed dCas9 approach (Gu et al. 2018) is currently being explored in this context.

# Chapter 5

## Conclusion

The spatial organisation of chromatin changes drastically upon differentiation, and we are just starting to understand the importance of this (Dixon et al. 2015). Formative pluripotency is considered to be a pivotal point in early differentiation of mESCs (Smith 2017), and it also exhibits a very unusual chromatin conformation (see section 1.3.2). Therefore, understanding how this transient structure emerges and whether it serves any particular functions might give insights into the roles of chromatin in cell fate determination.

Partitioning of the genome into eu- and heterochromatin is crucial for genome stability, silencing of repetitive elements, transposons and endogenous retroviruses, and correct gene expression. Thus, it is important to understand the biological and physical features of heterochromatin, as well as the mechanisms governing its segregation. Genome compartmentalisation is weakened in the formative-state mESCs, which, on one hand, points to a potential role for heterochromatin during exit from the naive state; and on the other hand, provides a naturally dynamic system for studying heterochromatin organisation.

The work described in this thesis aimed to characterise heterochromatin in mESCs and during the onset of mESC differentiation. The main focus of this project was HP1 $\beta$  – a major heterochromatic protein, whose function is still not understood well despite decades of active research. The goals set at the beginning of this study included: i) an

initial investigation of HP1 and heterochromatic histone marks in naive-, formative- and primed-state mESCs (chapter 2); ii) application of single-molecule imaging to characterise the dynamic properties of HP1 $\beta$  and better understand the physical nature of heterochromatin (chapter 3); and iii) development of a novel method, combining 3D super-resolution imaging and snHi-C, that would enable studying the links between genome conformation and the distribution and activity of chromatin-binding proteins (chapter 4).

In chapter 2, comparing the spatial density distributions of HP1 $\beta$  and DNA in the same cells revealed that they were not perfectly correlated, and that HP1 $\beta$  often formed small puncta in euchromatin. At the onset of differentiation, HP1 $\beta$  expression increased, and a larger fraction of it localised to euchromatin. In contrast, in primed mESCs, HP1 $\beta$  became more enriched in the DNA-dense chromocentres, potentially partially replacing HP1 $\alpha$  in these regions. Importantly, HP1 $\beta$  has been shown to regulate gene expression during differentiation of mESCs (Mattout et al. 2015), which might be mediated by its dynamic targeting to different genomic regions. An attractive hypothesis is that HP1 $\beta$  might form small condensates around the genes destined for silencing, which subsequently coalesce with heterochromatin. Investigations into the dynamics of HP1 $\beta$  puncta, as well as identification of the sequences associated with them, are needed to test this model. Furthermore, exploring the relationship between DNA density, HP1 protein enrichment, and histone marks in the same cells might provide further clues to the principles of heterochromatin organisation and reorganisation.

Experiments described in chapter 2 also showed that the nuclei of formative-state mESCs are flatter, consistent with chromatin decondensation and nuclear softening observed previously in similar conditions (Chalut et al. 2012; Pagliara et al. 2014). Interestingly, HP1 $\beta$  displayed slightly slower diffusion at this stage, potentially caused by the altered environment in the nucleus (chapter 3). It remains to be seen whether these transient changes in the mechanical properties of nuclei during the exit from naive pluripotency might play any functional roles.

The main focus of chapter 3 was characterisation of HP1 $\beta$  diffusion in eu- and heterochromatin, and in particular search for properties that support or contradict

the LLPS or cross-linked polymer globule models of heterochromatin organisation. Firstly, the concentration of HP1 $\beta$  not bound to chromatin was found to be higher in heterochromatin. Furthermore, unbound HP1 $\beta$  displayed stronger confinement in heterochromatin, but, surprisingly, its diffusion coefficient was not reduced. No evidence was found for the HP1 $\beta$  molecules “bouncing off” the boundary of heterochromatic foci – in contrast, a higher tendency for motion reversal was observed deeper inside heterochromatin. Overall, these observations turned out to be hard to interpret, since they could potentially be compatible with either model under certain conditions. Moreover, potential features of non-equilibrium behaviour were detected at the border of the foci, further complicating the picture. Thus, an appropriate biophysical framework must be developed to reach a deeper understanding of heterochromatin organisation.

Chapter 4 describes the development of 3D whole-nucleus SMLM imaging method for nuclear proteins, such as SOX2, NANOG and HP1 $\beta$ , as well as DNA, that could be combined with snHi-C and genome fold reconstruction. Several proof-of-concept datasets have been generated, proving the feasibility of this approach. Currently, the main limitations for realising the full potential of the technique include low sampling rate of the images, aberrations affecting the accuracy of DH-PSF mapping, and problems with DNA labelling. However, these issues can be solved in due course. The next challenge will then be the overlaying of the 3D images and the snHi-C structures.

The combination of super-resolution imaging and snHi-C would enable researchers obtain sequence-specific information about the genome structure alongside the precise protein locations in the same cell. This, for example, could allow one to identify the genomic sequences located in the euchromatic HP1 $\beta$  speckles. Many biological questions beyond the scope of this work could also be addressed, such as characterisation of the genes co-clustering within transcription factories or polycomb bodies. Importantly, the single-molecule nature of the experiment provides an opportunity to quantify protein clustering and correlate the size, shape, density and other features of the clusters with the associated genomic sequences.

In conclusion, the chromatin field in the recent decades has moved very far, uncovering the principles of genome folding from the level of nucleosomes all the way to

compartments and chromosome territories. Furthermore, understanding of the functional significance of the different layers of genome organisation in different contexts is also growing. However, some basic questions still remain unanswered, such as how the partitioning of the genome into eu- and heterochromatin is achieved. I believe that answers to these questions are certainly within our reach, and that they will be obtained in the near future by combining knowledge from different fields.

# Chapter 6

## Materials and Methods

### 6.1 Materials

#### 6.1.1 Cell lines

The cell lines used in this thesis are listed in Table 6.1.

**Table 6.1:** List of cell lines. JG, Joost Gribnau group, Erasmus MC, Rotterdam, line described in Rivera-Mulia et al. (2018); MW, Meike Wiese, MPI Immunology and Epigenetics, Freiburg; MB, Marc Bühler lab, Friedrich Miescher Institute, Basel, lines described in Ostapczuk et al. (2018); AFS, Anne Ferguson-Smith lab, University of Cambridge, lines described in Strogantsev et al. (2015); BDH, Brian Hendrich lab, Cambridge Stem Cell Institute.

Name	Background	Modification	From
Hybrid	129/CAST ♀	-	JG
HP1 $\beta$ -Halo(C)	129/CAST ♀	<i>Cbx1</i> <sup>HaloTag/HaloTag</sup> C-terminal	MW
HP1 $\beta$ -Halo(N)	129/CAST ♀	<i>Cbx1</i> <sup>HaloTag/HaloTag</sup> N-terminal	MW
mES159	mES159	<i>Rosa26</i> <sup>Cre-ERT2/BirA-V5</sup>	MB
HP1 $\gamma$ KO	mES159	<i>Cbx1</i> <sup>FlagAvi/FlagAvi</sup> <i>Cbx3</i> <sup>3/-</sup> <i>Rosa26</i> <sup>BirA-V5/-</sup>	MB
HP1 $\beta$ KO, HP1 $\gamma$ FL	mES159	<i>Cbx1</i> <sup>-/-</sup> <i>Cbx3</i> <sup>loxP/loxP</sup> <i>Rosa26</i> <sup>Cre-ERT2/BirA-V5</sup>	MB
HP1 $\beta$ KO, HP1 $\alpha,\gamma$ FL	mES159	<i>Cbx1</i> <sup>-/-</sup> <i>Cbx3</i> <sup>loxP/loxP</sup> <i>Cbx5</i> <sup>loxP/loxP</sup> <i>Rosa26</i> <sup>Cre-ERT2/BirA-V5</sup>	MB
CB9 hybrid	Black6/CAST ♀	-	AFS
BC8 hybrid	CAST/Black6 ♂♀	-	AFS
NANOG-Halo	Black6/CAST ♀	<i>NANOG</i> <sup>HaloTag/+</sup>	BDH
SOX2-Halo	CAST/Black6 ♂	<i>SOX2</i> <sup>HaloTag/+</sup>	BDH

## 6.1.2 Tissue culture media and supplements

The reagents used for growing mESC cells in culture are listed in Table 6.2. The recipe for the N2B27 medium, after Ying et al. (2003), is given in Table 6.3, that for 2i/LIF in Table 6.4 (Ying et al. 2008). To prepare serum/LIF, an aliquot of the base medium described in Table 6.5 was supplemented with 10 ng/ $\mu$ l LIF and 100 I.U./ml/100  $\mu$ g/ml penicillin/streptomycin and kept for <1 week.

**Table 6.2:** List of media and supplements used in cell culture. CSCI TC facility, Cambridge Stem Cell Institute tissue culture facility.

Reagent	Supplier and cat. no.
Dulbecco's PBS	Sigma-Aldrich D8537-500ML
Gibco Neurobasal (no phenol red)	Thermo Fisher 12348017
Gibco DMEM/F-12 (no phenol red)	Thermo Fisher 11039047
Gibco GMEM	Thermo Fisher 11710035
N2 (200 $\times$ )	CSCI TC facility
Gibco B-27	Thermo Fisher 17504044
L-glutamine	CSCI TC facility
Sodium pyruvate	CSCI TC facility
Non-essential amino acids (100 $\times$ )	CSCI TC facility
Gibco 2-mercaptoethanol	Thermo Fisher 31350010
Penicillin/streptomycin	CSCI TC facility
CHIR99021	CSCI TC facility
PD0325901	CSCI TC facility
mLIF	CSCI TC facility
Fetal bovine serum 10 $\times$	CSCI TC facility
Accutase	CSCI TC facility
0.2% Gelatin in PBS	CSCI TC facility
Poly-L-ornithine solution (0.01%)	Sigma-Aldrich P4957-50ML
Laminin	Sigma-Aldrich L2020-1MG
4-hydroxytamoxifen	CSCI TC facility

## 6.1.3 Antibodies

The commercial primary and secondary antibodies used in this work are listed in Table 6.6. The antibodies conjugated to fluorophores in-house are listed in Table 6.7.

**Table 6.3:** Recipe for 1L of N2B27 medium (no phenol red), kept for <1 month at 4°.

Reagent	Amount/Concentration
Neurobasal	500 ml
DMEM/F-12	500 ml
N2	5 ml
B-27	10 ml
L-glutamine	2 mM
Non-essential amino acids	10 ml
2-mercaptoethanol	50 $\mu$ M

**Table 6.4:** Recipe for 50 ml of 2i/LIF medium. Kept for <1 week at 4°.

Reagent	Amount/Concentration
N2B27	50 ml
Penicillin/streptomycin	100 I.U./ml/100 $\mu$ g/ml
PD0325901	1 $\mu$ M
CHIR99021	3 $\mu$ M
mLIF	10 ng/ $\mu$ l

**Table 6.5:** Recipe for 500 ml of serum-containing base medium, kept for <1 month at 4°.

Reagent	Amount/Concentration
GMEM	500 ml
Fetal bovine serum	50 ml
L-glutamine	2 mM
Sodium pyruvate	1 mM
Non-essential amino acids	5 ml
2-mercaptoethanol	50 $\mu$ M

**Table 6.6:** List of antibodies and the dilutions used in different types of experiments. IF, immunofluorescence, WB, western blot, HRP, horseradish peroxidase.

Antibody	Supplier and cat. no.	Dilution
<b>Primary antibodies</b>		
Anti-HP1 $\beta$ Rb pAb	Abcam ab10478	1:200 IF
Anti-HP1 $\beta$ Ms mAb	Invitrogen MA3-053	1:2000 WB
Anti-HP1 $\alpha$ Rb mAb	Abcam ab109028	1:250 IF
Anti-H3K9me3 Rb pAb	Active Motif 39062	1:500 IF
Anti-H3K27me3 Rb pAb	Sigma-Aldrich 07-449	1:500 IF
Anti-RNAPII NTD Rb mAb	Cell Signalling 14958S	1:2000 WB
<b>Secondary antibodies</b>		
Dk anti-Rabbit IgG	Jackson Immunoresearch	NA
Sh anti-Mouse IgG + HRP	GE Healthcare NA931V	1:2000 WB
Dk anti-Rabbit IgG + HRP	GE Healthcare NA934V	1:2000 WB

**Table 6.7:** List of antibodies, conjugated to fluorophores in-house, and the dilutions used in different types of experiments.

Initial antibody	Fluorophore	Dilution
Anti-HP1 $\beta$ Rb pAb	Cy5	1:20 IF, 1:10 IF (flow cyt)
Dk anti-Rabbit IgG	Cy3	1:75 IF
Dk anti-Rabbit IgG	Cy5	1:75 IF

### 6.1.4 Fluorophores

The fluorophores used for microscopy are listed in Table 6.8.

**Table 6.8:** The fluorophores used for live- or fixed-cell staining in this work.

Fluorophore	Conjugated to	Source
Janelia Fluor 549 (JF <sub>549</sub> )	HaloTag ligand	L. Lavis, Janelia
Janelia Fluor 646 (JF <sub>646</sub> )	HaloTag ligand	L. Lavis, Janelia
Photoactivatable JF <sub>549</sub> (PA-JF <sub>549</sub> )	HaloTag ligand	L. Lavis, Janelia
Photoactivatable JF <sub>646</sub> (PA-JF <sub>646</sub> )	HaloTag ligand	L. Lavis, Janelia
Hoechst33342	-	Invitrogen H1399
Silicon Rhodamine (SiR)	Hoechst (5')	G. Lukinavičius, MPI-BPC

### 6.1.5 Primers and TaqMan probes

The primers used for genotyping of the *Cbx1* locus in HP1 $\beta$ -Halo(N) cells can be found in Table 6.9, and the primers and TaqMan probes used for RT-qPCR are listed in Table 6.10.

**Table 6.9:** Primers used for genotyping (F, forward, R, reverse).

Name	Sequence
F1	CACCTTGCCCTTGACAACCTCG
F2	CTTTGACGCGCTCTGGATTG
F3	GGAGTTGAGCAGAGCATTGG
F4	CCAAAAGCCTGCCTAACTGC
R1	TATCCAGCCAGGAAATGTGCC
R2	TATCCAGCCAGGAAATGTGCC
R3	TCTGGTTTGTCCGATTTGCC
R4	GATCCTCTTGCCTCCCTCTT

**Table 6.10:** RT-qPCR primers and probes.

Gene	Forward and Reverse Primers
<i>Prdm14</i>	CAGCACCAAGGGCAGATATT, GCGTGTACTTCAGGTGCTTA
<i>Nefl</i>	AATGCAGACATTAGCGCCAT, TGAGAGTAGCCGCTGGTTAT
<i>Nefm</i>	AAATGAACTTCGGGGAACCAA, CGGTGATGCTTCCTGAAAATG
Gene	TaqMan Probe ID
<i>Nes</i>	Mm00450205.m1
<i>Dppa3</i>	Mm01184198.g1
<i>Pou5f1</i>	Mm03053917.g1
<i>Zfp42</i>	Mm03053975.g1
<i>NANOG</i>	Mm02019550.s1
<i>Otx2</i>	Mm00446859.m1
<i>Fgf5</i>	Mm03053745.s1
<i>Gapdh</i>	Mm99999915.g1
<i>Atp5a1</i>	Mm00431960.m1
<i>Ppia</i>	Mm02342430.g1

## 6.2 Characterisation of cell lines

### 6.2.1 Cell culture

Mouse embryonic stem cells were routinely grown under 2i/LIF conditions (Table 6.4) (Ying et al. 2008) on tissue culture flasks, dishes or multi-well plates. They were typically passaged every second day with a split of 1:8-1:12, using accutase. Testing for mycoplasma contamination was performed approximately once-twice a year, and all results were negative.

#### Differentiation protocol

To differentiate cells into the formative or primed states, they were first grown in 2i medium without mLIF for 24h. Then the medium was changed to N2B27 (with addition of penicillin/streptomycin), and cells were grown for further 24h or 48h. To avoid overgrowing, the mESCs that were going to be differentiated for 48h in N2B27 were plated at a two-times lower density than usual.

## Deletion of *loxP*-flanked genes

To induce Cre-mediated excision of *Cbx3* and *Cbx5* in HP1 $\beta$  KO, HP1 $\gamma$  FL or HP1 $\beta$  KO, HP1 $\alpha,\gamma$  FL mESCs, cells were transferred into 2i/LIF medium, containing 400 nM 4-hydroxytamoxifen, 48h prior to the start of differentiation protocol. Cells lacking all HP1 proteins grew slower and showed increased cell death, but enough cells stayed viable to conduct experiments.

### 6.2.2 Genotyping

For genotyping of the *Cbx1* locus in HP1 $\beta$ -Halo(N) mESCs, genomic DNA (gDNA) was prepared from  $\sim 10^5$  harvested mESCs. 25  $\mu$ l of lysis buffer (75 mM Tris-HCl pH8.8, 20 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 3 mM MgCl<sub>2</sub>, 0.45% NP-40 (Sigma I8896), 0.45% Tween20 (Sigma P1379) and 0.2 mg/ml proteinase K (NEB P8107S)) were added to each pellet, incubated at 55°C for 1 h and then at 95°C for 10 min.

PCR reactions were performed, using the Platinum SuperFi Master Mix (ThermoFisher 12358010), 0.5  $\mu$ l of gDNA template or water as a negative control and 0.5  $\mu$ M of either primer, in 10  $\mu$ l total volume. The reaction products were mixed with a loading buffer, run on a 1% agarose gel and visualised using ethidium bromide.

### 6.2.3 Western blotting

#### Nuclear protein extraction

To prepare the extract for western blot, cells from a T75 culture flask were harvested by scraping in cold phosphate buffer saline (PBS), centrifuged, and the pellet was resuspended in buffer A (10 mM HEPES pH7.9, 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, 0.5 mM DTT (Thermo Fisher R0861), and protease inhibitors: 1 mM PMSF (Sigma 93482), 1  $\mu$ g/ml aprotinin (Sigma 10820), 10  $\mu$ g/ml leupeptin (Sigma L2884), 10  $\mu$ g/ml pepstatin (Sigma P5318) and 1 mM Na<sub>3</sub>VO<sub>4</sub> (Sigma 5086050004)). After a 15 min incubation on ice, 31.25  $\mu$ l of 10% NP-40 was added, the tube was vortexed and centrifuged at 10,000 rpm for 1 min at 4°C. The pellet was washed once with buffer A, and 60  $\mu$ l

of buffer C (10 mM HEPES pH7.9, 12.5% glycerol, 0.75 mM MgCl<sub>2</sub>, 10 mM KCl, 0.1 mM EDTA, 400 mM NaCl), 0.5 mM DTT and protease inhibitors) were added. Next, the sample was shaken for 1h at 4°C and centrifuged at 14,000 rpm for 10 min at 4°C. The supernatant was collected, and the protein concentration was estimated using NanoDrop.

### **Gel electrophoresis and blot**

15 µg of protein were mixed with water and 2× Laemmli buffer concentrate (100mM Tris-HCl pH6.8, 4% SDS, 0.2% w/v bromophenol blue, 20% glycerol, 100mM DTT) to a final volume of 16 µl, boiled at 100°C for 10 min and run through a polyacrylamide gel (5% stacking, 15% running). The proteins were transferred onto a 0.45 µm nitrocellulose membrane, using standard procedures, and presence of protein was confirmed with Ponceau.

Blocking was performed in 5% milk (Merck 70166) in PBST (PBS + 0.1% Tween20) for 1 h at room temperature. The membrane was then incubated with primary antibody in the same buffer for 1h at room temperature or overnight at 4°C and washed 3 times in PBST, 15 min each. Labelling with HRP-conjugated secondary antibody proceeded in the same way, followed by 5 PBST washes.

The blot was visualised on film, using an HRP detection kit (Fisher Scientific 10340125).

### **6.2.4 Alkaline phosphatase assay**

For this experiment, mESCs were transferred into serum/LIF medium 7-9 days before scoring. 2 days before, 30000 cells were plated onto wells of a 6-well plate. Cells were rinsed with PBST (PBS + 0.5% Tween20), fixed with 4% formaldehyde (prepared from 16% formaldehyde, ThermoFisher 28906, and kept at 4°C for <1 week), rinsed again with PBST and stained for alkaline phosphatase activity, using the Sigma 86R-1KT kit (without counterstaining with haematoxylin). Approximately 100 colonies were scored using a bright-field light microscope, according to their staining (clear pink staining,

no staining, or mixed/unclear). Scoring was performed blind. The experiment was conducted twice for each cell line.

### 6.2.5 RT-qPCR

Reverse transcription-quantitative PCR (RT-qPCR) experiments were performed to assess RNA levels of various genes. In a typical experiment, cells were grown on separate wells of a 6-well plate as biological replicates. RNA was extracted with TRI reagent (Sigma T9424) using a standard protocol. Contaminant DNA was degraded with RNase-free DNase (Qiagen 79254) and 100-500 ng of the resultant purified RNA was used for RT. The standard two-step RT protocol was performed, using SuperScriptIV (ThermoFisher 18090010). No RT controls were produced analogously, but without addition of the SuperScriptIV polymerase. 1  $\mu$ l of template cDNA was then used for each 5  $\mu$ l qPCR reaction, using either TaqMan probes (Life Technologies) or SYBR Green mix (Life Technologies 4385614) and custom-made primers. Each qPCR reaction was done in triplicate, and for each probe or primer pair a no-template control was also performed.

#### RNA levels of HP1 $\beta$ -controlled genes

RT-qPCR experiments were performed to compare expression of the genes listed below between the HP1 $\beta$ -Halo(N) and control hybrid cells:

1. *Prdm14* and *Dppa3* – decrease upon HP1 $\beta$  knock-out (Mattout et al. 2015);
2. *Nes*, *Nefl* and *Nefm* – rise upon HP1 $\beta$  knock-out (Mattout et al. 2015);
3. *Pou5f1* (Oct4) and *Zfp42* (Rex1) – common pluripotency markers.

RNA levels of three housekeeping genes *Gapdh*, *Atp5a1* and *Ppia* were used for normalization, and the results indicated similar trends regardless of the reference gene used. Three biological repeats were examined, with three technical replicates within each. The data was analyzed in R (R Core Team 2017) with double delta Ct analysis (here and further, the `dplyr`, `tidyr` and `ggplot2` packages (Wickham et al. 2019; Wickham

2016) were commonly used for data carpentry and plotting). Statistical significance of the difference from the wild-type level of expression was tested by applying Student's t-test both to each biological repeat individually, using pooled variance from all genes within the same biological sample, and to sets of biological replicates.

### **Differentiation analysis**

To assess progression through the differentiation protocol, changes in RNA levels of marker genes *NANOG*, *Zfp42*, *Otx2* and *Fgf5* were measured alongside housekeeping gene controls, using TaqMan probes. The experimental procedure was analogous to the one described above, typically with three biological repeats per condition, each including three technical replicates. The data was analyzed in R.

### **6.2.6 Flow cytometry and cell sorting**

Flow cytometry experiments were performed on the CytoFLEX instrument at the Cambridge University Flow Cytometry Facility. Samples for comparison were always recorded on the same day, using the same setup and gains. About  $1.5\text{-}2.5 \times 10^6$  cells were typically prepared, and at least 10,000 events were recorded per sample. The measurements included the forward and side scatter (FSC and SSC) signal height and area, used to select intact single cells; blue fluorescence intensity for cell cycle analysis, using Hoechst; and yellow or red fluorescence intensity for JF<sub>549</sub>, JF<sub>646</sub> or Cy5.

### **Live cell staining and measurement**

Live HP1 $\beta$ -Halo(N) mESCs were first labelled with 250 nM of either JF<sub>549</sub> or JF<sub>646</sub> in growth medium for 15 min at 37°C, and three 15-20 min washes in medium were performed to remove excess dye. Short washes in PBS were done in between the steps.

Harvested cells were passed through a cell strainer and kept at 37°C. 10  $\mu\text{g}/\text{ml}$  of Hoechst33342 were added 25-30 min prior to measurements, since this time was found to result in optimal fluorescence signal and minimal cell death.

In the initial round, hybrid cells without HaloTag were used as a negative control

to assess non-specific fluorescence. This was found to be several orders of magnitude lower than the specific signal. In addition, HP1 $\beta$ -Halo(N) mESCs labelled only with Hoechst or unlabelled also did not have any appreciable background fluorescence.

## **Immunofluorescence**

Hybrid mESCs were harvested and fixed in suspension for 10 min with 2% formaldehyde. After quenching with 100 mM of glycine and washing with PBS, cells were permeabilised with 0.5% TritonX-100 (VWR 28817.295) in PBS. 0.1% TritonX-100 was used for washing, followed by 1h incubation with Cy5-labelled anti-HP1 $\beta$  antibody, diluted in 0.5% bovine serum albumin (BSA) (Sigma A7906), 0.1% TritonX-100 in PBS, at room temperature. The cells were washed twice with the same buffer, resuspended in PBS with 10  $\mu$ g/ml of Hoechst and incubated for 1.5h.

The negative control cells, treated in the same way as the sample, except the antibody not being added during the labelling step, gave rise to very low unspecific signal.

## **Data analysis**

The flow cytometry data was analysed in R (R Core Team 2017), using the flowCore (Ellis et al. 2019), flowWorkspace (Finak and Jiang 2020) and ggcyto (Van et al. 2018) packages. Gates on the SSC-area vs FSC-area and FSC-height vs FSC-area plots were drawn manually to select single cells. Thresholds for fluorescence in yellow or red channels were also used to exclude any unlabelled cells (incidence rate  $\ll 1\%$ ), informed by the appropriate negative controls. Quantification of the percentages of populations was also performed by manual drawing of gates on the Hoechst vs dye fluorescence plots.

## **FACS**

Fluorescence-activated cell sorting (FACS) was conducted on BD Aria III sorter at the Facility. To separate HP1 $\beta$ -Halo(N) high- and low-HP1 $\beta$ -expressing mESCs, HP1 $\beta$  in  $>10^7$  cells was labelled with JF<sub>646</sub> as described above. To minimise cytotoxicity, the

Hoechst dye was not used. Single cells were picked using the SSC-area vs FSC-area and FSC-height vs FSC-area plots. Stringent gates for the two populations were set on the SSC-area (as a proxy for cell size and thus cell cycle stage) vs JF<sub>646</sub> fluorescence plot, aided by unlabelled and Hoechst-labelled controls. The high- and low-expressing populations were sorted, and a few cells from each tube were analysed to confirm the purity of the sort. Since two days later, the sorted high-expressing population was found to still contain a proportion of low-expressing cells, the process was repeated to enrich for the high-expressing population further.

## 6.3 Confocal imaging

Confocal imaging of fixed cells was performed to study the distribution of HP1 $\beta$ , HP1 $\alpha$ , H3K9me3 and H3K27me3, alongside the DNA density pattern, in mESCs. Whenever possible, the samples under comparison were prepared and imaged together, unless otherwise specified. Two biological repeats were collected for HP1 $\beta$ , HP1 $\alpha$  and H3K9me3, but only one for H3K27me3. The data shown in chapter 2 are from replicate 1, and the observations were similar in replicate 2.

### 6.3.1 Sample preparation

Cells for any type microscopy experiment (except imaging combined with Hi-C) were grown on glass-bottom 35 mm dishes (MatTek P35G-1.0-14-C), coated with poly-L-ornithine (>30 min at 37°C) and laminin (1:100, >4h at 37°C). To decrease the occurrence of cells growing in multiple layers and thus obtain clearer images, mESCs were plated at two-times lower density than usual (and further diluted if differentiated into the primed state).

To visualise DNA, after other staining procedures, cells were incubated with 1  $\mu\text{g}/\text{ml}$  Hoechst33342 for 30 mins at room temperature, and rinsed with PBS.

## **Staining with JF dyes**

Cells were stained with Janelia Fluor dyes as described in section 6.2.6 and fixed with 4% formaldehyde for 10 min at room temperature. Next, cells were treated with 0.1% Triton-X100 in PBS for 4 min, and 1% BSA in PBS for 1h, with short PBS washes in between. Samples were stored and imaged in PBS.

## **Immunofluorescence**

For immunofluorescence, cells were fixed with 4% formaldehyde as above, permeabilised with 1-1.5% TritonX-100 for 10 min at room temperature and blocked with 10% BSA for 1h. The sample was then incubated with primary antibody diluted in 10% BSA for 4-5h at room temperature, or overnight at 4°C. Three 10 min washes were performed with 0.1% TritonX-100 and 0.2% BSA in PBS. If necessary, labelling with the secondary antibody diluted in 10% BSA proceeded for 1h. After repeating the washing steps and a quick wash in PBS, mESCs were re-fixed in 3% formaldehyde for 10 mins. Finally, the sample was rinsed with PBS three times and with water twice, and stored or imaged in PBS.

### **6.3.2 Microscopy**

Confocal imaging was performed on Leica SP8 microscope at Cambridge Advanced Imaging Centre, typically with scan speed 400 Hz, line averaging of 4 and pixel size 90×90 nm. When imaging Hoechst with JF<sub>549</sub> or Cy3, sequential scanning was used due to partial overlap of the spectra, while simultaneous scanning was applied for Hoechst and JF<sub>646</sub> or Cy5. Z-stacks with 0.5 µm step were collected, covering the entire height of the cells. Exactly the same settings were used when collecting data for comparison. Whenever possible, samples in different conditions were imaged in one session, or spread over multiple sessions without bias (i.e. equal numbers of images were taken for each sample/condition during one session), unless otherwise specified.

### 6.3.3 Image analysis

Analysis of confocal images was performed in Fiji (Schindelin et al. 2012). Stacks were flattened by maximum projection, since this type of projection highlights heterochromatic foci, hides nucleoli, and is not affected by the intensity of pixels above/below the cell. Alternatively, a slice through the middle of cells was sometimes selected manually. Since the image background was very low, no background correction was performed. The intensity values of images were not adjusted, unless otherwise specified.

#### Segmentation of nuclei and heterochromatic foci

HP1 $\beta$  and DNA patterns were analysed quantitatively, using maximum projections. For that, the nuclei of mESCs were first segmented in the HP1 $\beta$  channel, using the Huang2 automatic global threshold and watershed. The results were examined and corrected manually.

HP1 $\beta$  foci were then segmented with the MidGrey automatic local threshold with 100 px radius. To select more pronounced foci and exclude very small intensity spikes, “Fill holes”, “Erode” and “Dilate” commands were applied to the binary mask. The selected foci were demarcated heterochromatin, while the regions outside of them, but within the nuclei, comprised euchromatin. This procedure was somewhat biased to miss foci in 2i/LIF mESCs and pick false foci in 24h and 48h images, but the bias was not large and, if anything, weakened the observed trends. This segmentation was then used to measure the area of nuclei and of foci, and the brightness of signal in eu- and heterochromatin as a proxy for HP1 $\beta$  concentration. The data was then plotted and analysed in R (R Core Team 2017).

The same segmentation of nuclei was reused in DNA staining intensity analysis, but heterochromatin was redefined by applying the same algorithm to the Hoechst channel. Since some nuclei clearly displayed very weak DNA staining, those were manually excluded from the analysis.

## DNA and HP1 $\beta$ staining intensity correlation

Pearson's correlation coefficient between the pixel-wise intensity values in the HP1 $\beta$  and DNA channels was calculated for each segmented nucleus, using the Coloc 2 plugin. Since small discrepancies can occur between two fluorescence channels in z, comparing the maximum intensity projections in this case was especially appropriate.

## 6.4 SptPALM of HP1 $\beta$

### 6.4.1 Sample preparation

HP1 $\beta$ -Halo(N) mESCs were used in sptPALM experiments, and in the 2D datasets shown in chapter 3, the population with correct levels of HP1 $\beta$  expression was enriched for. HP1 $\beta$  in those cells was labelled, using the photoactivatable dyes PA-JF<sub>549</sub> or PA-JF<sub>646</sub>. The optimised protocol included a 15 min incubation with 250 nM of the dye, followed by 6 15-20 min washes with growth medium, all at 37°C. Short PBS rinses were performed between the washes.

In two-colour experiments, the fluorophore used for wide-field imaging had to have a shorter excitation wavelength compared to the photoactivatable dye to avoid bleaching. Thus, PA-JF<sub>646</sub> has a longer excitation wavelength than PA-JF<sub>549</sub> was used for sptPALM in dual-colour experiments. Labelling of HP1 $\beta$  for wide-field imaging was achieved by adding 25 nM of JF<sub>549</sub> during the third wash. On the other hand, DNA staining with 1  $\mu$ g/ml of Hoechst33342 was performed immediately before imaging to minimise cytotoxicity. Cells were incubated with the dye 15 min at 37°C, followed by two PBS rinses and addition of fresh medium.

### 6.4.2 Microscopy

#### Microscope setup

A custom-built microscope, designed by Dr A. Ponjavic, was used in this study. The microscope body is an inverted fluorescence microscope (Nikon Eclipse Ti2 with a Perfect

Focus system). The 405 nm, 561 nm and 638 nm laser beams (100 mW MLD 405-100, Cobolt; 500 mW Jive-500 DPSS, Cobolt; 360 mW L2C laser combiner, Oxxius) with excitation filters (FF01-403/5-25; FF01-561/4-25; FF01-637/7-25, Semrock) were expanded and collimated using Galilean beam expanders and then combined using dichroic mirrors. Neutral density filters in front of the fibre heads were employed to regulate laser intensities.

A  $60\times$  1.49 NA oil immersion objective (2D experiments), or  $60\times$  1.27 NA water immersion objective (DH-PSF experiments) (CFI Apochromat TIRF and IR series, Nikon) were used to focus the light onto the sample and to collect the fluorescence emission. The laser intensities entering the sample were measured to be approximately  $0.2\text{ kW/cm}^2$  (405 nm),  $5\text{ kW/cm}^2$  (561 nm) and  $5\text{ kW/cm}^2$  (638 nm), when no neutral density filters were used. To further boost the excitation power for super-resolution imaging, the yellow and red laser beams were shrunk, using the beam expanders, resulting in illumination intensities of  $20\text{ kW/cm}^2$  and  $40\text{ kW/cm}^2$ , respectively.

An UltraFlat ( $<0.25\text{ }\lambda/\text{inch}$ ) quad-band dichroic mirror (ZT405/488/561/640rpc-UF3, Chroma) was used to separate excitation and emission. A Köhler lens was used to focus the light onto the back focal plane of the objective, resulting in widefield excitation, but highly inclined and laminated optical sheet (HILO) illumination was typically used to minimise out-of-focus background. For dual-colour experiments, the emitted light was separated by another dichroic mirror (FF624-Di01, Semrock). For DH-PSF microscopy, a double helix phase mask designed for an appropriate wavelength (Double Helix) was fitted into the emission path. After passing through emission filters (87752, Edmund Optics + BLP02-561R-25, Semrock; 67038, Edmund Optics + BLP01-635R-25), the light was focused onto EMCCD cameras (Evolve 512 Delta, Photometrics; EM gain 250) using tube lenses. The final pixel size was 271 nm.

A part of the microscope body, including the stage, was encased in an incubator to maintain  $37^\circ\text{C}$  temperature during live-cell experiments.

## Imaging conditions

In SPT, molecules should be tracked for as long as possible to accurately calculate the diffusion parameters. However, a strong excitation laser is still needed to make localisation with decent precision possible. Therefore, the minimum 638 nm laser power needed to get sharp PSFs from individual PA-JF<sub>646</sub> molecules with 5 ms exposure was empirically found to be approximately 4 kW/cm<sup>2</sup>. Furthermore, the 405 nm activation laser intensity, resulting in an optimal density of molecules per frame, was determined to be ~0.00008 kW/cm<sup>2</sup>. In these conditions, in 3D or 2D experiments, respectively, <10 or <20 PSFs were normally present in a field of view at any one time, the intensity of localisations was about 200-400 or 50-250 photons, and a substantial number of tracks lasted for >10 frames (Table 3.1).

For two-colour imaging, 561 nm laser at ~0.002 kW/cm<sup>2</sup> intensity was used to image HP1 $\beta$  labeled with JF<sub>549</sub>, while DNA stained with Hoechst33342 was excited with the low-power 405 nm laser also used for photoactivation of PA-JF<sub>646</sub>.

For long-exposure imaging, the autoactivation rate of molecules was sufficient, and no activation was required. Moreover, autoactivation of the dyes resulted in a bright signal initially, and thus pre-bleaching with maximum laser intensity for around a minute was required to reach the single-molecule mode. The 638 nm excitation laser was used at ~0.5 kW/cm<sup>2</sup> intensity. These settings resulted in a good density and brightness of localisations (500-1500 photons).

### 6.4.3 Controls

#### Acquiring the immobile control data

Two types of controls were recorded to estimate the lower limits of the technique: fixed mESCs with labelled HP1 $\beta$ , or the dye immobilised on a coverslip. To prepare samples for the latter, 150 nM of the PA-JF<sub>549</sub> or PA-JF<sub>646</sub> in PBS were incubated on glass coverslips coated with poly-L-lysine (Sigma P4707) for 5 min. The solution was then replaced with fresh PBS.

Two replicates of the fixed-cell experiment and one replicate of the immobilised

PA-JF<sub>549</sub> imaging were conducted under the same conditions as 3D sptPALM, with 10000 frames recorded at each of 20 different positions on the dish/coverslip. The data was analysed in the same way as the short-exposure live-cell experiments, and since all trajectories must be confined in these datasets, all points were used for comparison with the experiment. A single replicate was performed under 2D sptPALM conditions with immobilised PA-JF<sub>646</sub>, which is shown in chapter 3. The comparison of the 3D sptPALM controls and the 2D sptPALM controls with the experimental data, however, all yielded the same conclusions, indicating their robustness.

### Measurement of the bleaching rate

To assess the contribution of photobleaching to the dissociation kinetics investigated in the long-exposure SPT, two approaches for measurement of the bleaching rate were considered: measurement of the bulk rate in live mESCs, or determination of bleaching on the single-molecule level by SPT in fixed mESCs. Preliminary experiments showed that the bleaching kinetics in fixed cells was different from that in live cells, and thus the former approach was chosen.

For this, live cells prepared for long-exposure imaging of HP1 $\beta$  by labelling with PA-JF<sub>549</sub> (3D) or PA-JF<sub>646</sub> (2D) were used. After a short (5-10 s) pre-bleaching, 4-5 positions with 5-10 cells each were imaged for 400 frames under the sptPALM conditions. Several square regions with 5-7  $\mu\text{m}$  side were then selected in each series in Fiji, and their intensity profiles over time were measured. Each curve was then fitted to a simple exponential decay equation, with the exponent indicating the bleaching coefficient  $k_b$ . The mean of the measurements was then used to correct the residence time values.

## 6.4.4 Primary image analysis

### Localisation and tracking

Localisations were extracted from the data, using either the easy-DHPSF (Lew et al. 2013) or GDSC PeakFit software (Herbert 2014). The parameters were tuned

manually to detect a maximum number of true localisations without picking much noise (Table 6.11), but they were kept lenient due to the subsequent filtering at the tracking stage.

**Table 6.11:** Parameters used in GDSC PeakFit for fitting 2D short- and long-exposure SPT data. The parameters not shown were left at default values.

Parameter	Value (short exp)	Value (long exp)
Camera type	EMCCD	
Read noise	0	7.1622
Camera bias	400	
Gain	37.7	35.0
Quantum efficiency	0.95	
Smoothing	0.40	0.32
Search width	1	0.9
Border width	1	0.9
Fitting width	3	2.15
Neighbour height	0.3	0.3
Residuals threshold	1	0.8
Duplicate distance	0.2	0.4
Shift factor	1.75	1.2
Min photons	50	10
Min & max width	0.2 & 3	0.4 & 1.7
Precision	50	

Next, localisations originating from the same molecule were connected into tracks, using a script written by Dr Wayne Boucher (<https://github.com/wb104/trajectory-analysis>). The localisations were connected into the same track if they fulfilled the following criteria.

- They belonged either to adjacent frames, or were separated by no more than two frames. The latter adjustment enabled harvesting more information from the imperfect tracking data by not interrupting the trajectory if just a single localisation was not detected.
- The points were no more than  $1000 \times \sqrt{\text{separation in frames}}$  nm apart. The scaling reflected the fact that the distance traveled by a particle is proportional to the square root of time.

Next, trajectories containing fewer than three points were filtered out. The number of trajectories obtained for the datasets is shown in Tables 3.1 and 3.2.

### **Wide-field images and segmentation**

Due to the short exposure and low laser intensities used to visualise HP1 $\beta$  or DNA distributions in parallel with 5 ms sptPALM experiment, the signal-to-noise ratio of these images was low. Thus, the DNA and HP1 $\beta$  images were binned every 100 or 500 frames, respectively. In addition, the structured camera noise accumulated over this period was subtracted, using recorded sequences of dark counts. Binary masks were produced, outlining entire nuclei or heterochromatic foci. Although multiple automatic algorithms were tested, due to low signal-to-noise ratio and uneven illumination, the foci had to be selected by hand.

The nuclear outlines were used to exclude any noise originating from outside the cells. On the basis of the masks, trajectories were assigned as eu- or heterochromatic, if all of their points fell within one of the compartments, using a script in R.

Since picking trajectories entirely enclosed in small heterochromatic foci could have lead to bias, heterochromatic regions were compared with equivalent regions in euchromatin rather than with the entire nuclear space. The binary masks for these “pseudo-foci” were produced by randomly shifting the segmented heterochromatic foci within one cell into euchromatic areas, and the trajectories that were entirely within these regions were used for analysis. For comparison between the differentiation stages, however, all the euchromatic trajectories were used.

To calculate the intensity ratios in section 3.5.1, for each cell, the mean signal in the heterochromatic foci was divided by the mean signal in the euchromatic “pseudo-foci”.

Although in the long-exposure sptPALM, the wide-field images were of too low quality to distinguish the heterochromatic foci, the nuclei could be segmented out, which was used to filter out noise.

## 6.4.5 Short-exposure SPT data analysis

### The 4P algorithm

The code used for data analysis is implemented in Matlab and described in Basu et al. (2020). In brief, it performs the following steps:

1. Filter out all the trajectories shorter than  $n$  points ( $n = 11$  in this study);
2. Calculate the four parameters for each window along each trajectory. The window for each point is defined as the point  $\pm m$  points ( $m = 5$  used here) – i.e. the first and last points in a trajectory have a window of length  $m + 1$  associated with them, while for a point in the middle it is of length  $2m + 1$ ;
3. Apply a Gaussian mixture model (GMM) classifier to the calculated features;
4. Rejoin the separate points/windows into trajectories. Now, each point in each trajectory has a class associated with it, confined or unconfined. Next, divide trajectories into subtrajectories – stretches with all points belonging to the same class. For example, a trajectory with 11 points, belonging to classes *uuuuccccccu*, will be split into three subtrajectories *uuuu*, *cccccc* and *u*;
5. Filter out all subtrajectories shorter than  $l$  ( $l = 8$  in this study); and
6. Recalculate the four diffusion parameters for each subtrajectory rather than for each window. Record these values separately for each class.

### 4P algorithm adaptation for missing points

The filtering and classification adjustments are described in section 3.2.3, while below the modified formulae for diffusion parameter calculation are given.

1. For apparent diffusion coefficient, the correction for the varying time step was brought into the mean calculation in the following way:

$$D = \frac{MSD}{2d\Delta t} = \frac{\sum_{i=1}^l SD_i}{2ld\Delta t} = \frac{\sum_{i=1}^l \frac{SD_i}{2d\Delta t_i}}{l}, \quad (6.1)$$

where  $(M)SD$  is (mean) squared displacement and  $l$  is the number of steps in a trajectory (not counting the missing points).

2. The anomalous exponent was not affected much by the presence of gaps, probably because fitting was dominated by the correct points of the  $MSD$  vs time lag curve. Therefore, the calculation of  $\alpha$  was left as is.
3. The length of confinement is the standard deviation of the particle's position in a certain amount of time, and it is proportional to  $\sqrt{t}$ . Thus,  $Lc$  was normalised by the absolute length of observation to make values for different sliding windows and subtrajectories, which have different lengths and potentially gaps, comparable:

$$Lc_{norm} = STD(A)\sqrt{\frac{1}{l_{abs}}}, \quad (6.2)$$

where  $l_{abs}$  is the real number of steps in the trajectory/subtrajectory/window, including the missing points. In other words, the time period for  $Lc_{norm}$  was scaled to  $\Delta t$ .

4. The drift calculation is implemented in the 4P script on a step-by-step basis as follows:

$$V = \sqrt{\overline{\left(\frac{\Delta x_i}{\Delta t}\right)^2} + \overline{\left(\frac{\Delta y_i}{\Delta t}\right)^2} + \overline{\left(\frac{\Delta z_i}{\Delta t}\right)^2}}, \quad (6.3)$$

where  $\Delta x_i$  is the displacement in the x dimension at step  $i$ , and  $\overline{\left(\frac{\Delta x_i}{\Delta t}\right)}$  is the mean of displacements divided by the time step across the whole trajectory. This formula is equivalent to equation 3.4 only when there are no missing points. For trajectories with gaps, the calculation was changed to:

$$V = \sqrt{\left(\frac{\Delta X}{l_{abs}\Delta t}\right)^2 + \left(\frac{\Delta Y}{l_{abs}\Delta t}\right)^2 + \left(\frac{\Delta Z}{l_{abs}\Delta t}\right)^2}, \quad (6.4)$$

where  $\Delta X$ ,  $\Delta Y$  and  $\Delta Z$  are the total displacements in each dimension,  $\sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2} = \Delta L$ .

## Simulations of SPT data

Short-exposure SPT data was simulated, using a Matlab script published in Basu et al. (2020). In brief, a two-state model of diffusion was used, where the unconfined-state molecules displayed Brownian motion, while the molecules in the confined state were attracted to their initial position by a potential well. The diffusion coefficients of the two populations and the switching rates were varied in the simulations, but the depth of the potential well was kept constant.

To compare the performance of the original script and the modified version on trajectories with missing points, six simulations were conducted (Table 6.12). Each contained 3000 trajectories with a minimum length of 20 points and an average length of 30. An increasing proportion of data points (1-30%) were removed randomly to mimic trajectories with missing points, and the resulting datasets were analysed using the two versions of the 4P code.

63 simulations were used to assess the classification accuracy of the algorithm: 21 simulations with diffusion coefficient combinations outlined in Table 6.13, each with  $\tau_a = \tau_d = 1, 10$  or 100 s.

**Table 6.12:** A summary of the parameters, used to simulate trajectories for assessing the 4P algorithm performance on data with gaps.

	$\Delta t$	$D_{conf}$ and $D_{unconf}$	$\tau_a$ and $\tau_d$
<b>1</b>	5 ms	1.5; 3 $\mu\text{m}^2/\text{s}$	10000; 10000 s
<b>2</b>	5 ms	1.5; 3 $\mu\text{m}^2/\text{s}$	0.1; 0.1 s
<b>3</b>	5 ms	1.5; 3 $\mu\text{m}^2/\text{s}$	0.02; 0.02 s
<b>4</b>	5 ms	0.7; 1.5 $\mu\text{m}^2/\text{s}$	10000; 10000 s
<b>5</b>	5 ms	0.7; 1.5 $\mu\text{m}^2/\text{s}$	0.1; 0.1 s
<b>6</b>	20 ms	0.7; 1.5 $\mu\text{m}^2/\text{s}$	0.1; 0.1 s

## Jump distance analysis

Jump distance analysis involves comparing the distributions of the displacements of molecules within a time interval, which in this work was either 5 or 500 ms. The trajectories used in this analysis were subject to quality filtering, i.e. they had a minimum length of 8 and not too many missing points. Only the single-frame jumps

**Table 6.13:** A summary of the combinations of the diffusion coefficients used in the simulations, employed to test the 4P algorithm performance against the ground truth.

		$D_{conf}, \mu\text{m}^2/\text{s}$					
		0.001	0.01	0.1	0.5	0.7	1
$D_{unconf}, \mu\text{m}^2/\text{s}$	0.2	✓	✓				
	0.5	✓	✓	✓			
	1	✓	✓	✓	✓		
	1.5	✓	✓	✓	✓	✓	
	2	✓	✓	✓	✓	✓	✓

were considered, i.e. any gaps were omitted. The individual displacements comprising a trajectory were treated as independent data points.

Next, the following model was fit to the cumulative distributions (Weimann et al. 2013):

$$P(r^2, \Delta t) = 1 - \sum_{j=1}^m A_j e^{-\frac{r^2}{2dD_j\Delta t}}, \quad (6.5)$$

where  $P(r^2, \Delta t)$  is the probability that after a time interval  $\Delta t$ , a particle is found within a shell of radius  $r$  from its initial position;  $m$  is the number of species  $j$ , each characterised by its diffusion coefficient  $D_j$  and the proportion of the total number of particles that it constitutes  $A_j$  ( $\sum_{j=1}^m A_j = 1$ ); and  $d$  is the dimensionality ( $d = 2$  for 2D experiments). Fitting with  $m = 1, 2, 3$  populations was attempted. The least-squares parameters of the non-linear model were estimated by the nl2sol algorithm (R package Port) with initial values set to  $D_1 = 1 \mu\text{m}^2/\text{s}$  ( $m = 1$ );  $D_1 = 0.1 \mu\text{m}^2/\text{s}$ ,  $D_2 = 1 \mu\text{m}^2/\text{s}$ ,  $A_1 = A_2 = 0.5$  ( $m = 2$ ); or  $D_1 = 0.1 \mu\text{m}^2/\text{s}$ ,  $D_2 = 1 \mu\text{m}^2/\text{s}$ ,  $D_3 = 20 \mu\text{m}^2/\text{s}$ ,  $A_1 = A_2 = A_3 = 0.33$  ( $m = 3$ ) (using different initial values did not affect the result). The parameter estimates were constrained between 0 and  $100 \mu\text{m}^2/\text{s}$  for all diffusion coefficients, and 0 and 1 for the population proportions. The resulting models were scored using the Bayesian information criterion (BIC), which balances the goodness-of-fit with a penalty for a high number of variables in the model to avoid overfitting (Schwarz 1978).

A modified formula was used to take into account the localisation precision  $\sigma$

(Hansen et al. 2018):

$$P(r^2, \Delta t) = 1 - \sum_{j=1}^m A_j e^{-\frac{r^2}{2d(D_j \Delta t + \sigma^2)}}. \quad (6.6)$$

The precision was first estimated beforehand as  $\sigma = 50$  nm by measuring the median standard deviation of particle localisation in the immobilised dye control, which also approximately coincided with the PeakFit precision threshold used in detecting localisations. However, since the precision is combined with the  $D_j \Delta t$  term, it generally does not affect the shape of the curve that fits the data the best (i.e. the best-fit  $m$ ,  $A_j$  or  $D_j \Delta t + \sigma^2$  parameters), but only alters the apparent diffusion coefficient estimation. The only situation when this is not so is when  $D_j \rightarrow 0$  for some species. In those cases, the predicted number of populations did not change, however, the fit was worse, suggesting that this precision value was likely an overestimation.

Treating  $\sigma$  as a parameter for fitting, rather than pre-defining it, yielded similar results. In addition, the fitting procedure failed more often. Overall, since the comparisons in this work are relative, and there is no reason to believe that the precision changed between datasets, which were recorded on the same day, using the more complex model makes no real difference. Thus, these calculations were omitted from the main results.

#### 4P analysis

The 4P analysis of the experimental data was conducted with trajectories with a minimum length ( $n$  in the algorithm description) of 11, with a sliding window size ( $m$ )  $\pm 5$  points, and with a length threshold ( $l$ ) of 8 for the classified subtrajectories. The GMM classification was performed on each dataset and each compartment (eu/heterochromatin) separately, rather than a single model being applied to all data.

The distributions of diffusion parameters for all data, as well as for the confined data compared with control in Figure 3.10, were comprised of values calculated for individual points within trajectories (i.e. in the window around this point). Individual points/windows were also used in the calculations of the confined/unconfined popu-

lation proportions and in the correlation plots. In contrast, distributions of diffusion parameters for confined or unconfined populations consisted of values computed on subtrajectories rather than points (see steps 5 and 6 in the algorithm description) for increased accuracy and enrichment for the desired population (Basu et al. 2020).

### **Correlation between diffusion parameters and DNA or HP1 $\beta$ density**

The DNA and HP1 $\beta$  wide-field image timelapse series were averaged, since the density of the SPT data did not allow to perform time-resolved analysis. Furthermore, for simplicity, the raw intensity was divided by the mean of all DNA-labelling or all HP1 $\beta$ -labelling datasets.

Each point in the SPT trajectories that went through the 4P algorithm was mapped to the corresponding 271 nm $\times$ 271 nm pixel in the wide-field images of DNA or HP1 $\beta$  of the same cell. Pixels containing <10 different trajectories were filtered out to ensure that the observed effects were common between several molecules. After that, the individual points within trajectories were considered independent. A similar procedure was performed separately for points classified as confined or unconfined by the 4P algorithm. Since the number of points within each class was obviously lower than the total, fewer pixels passed the threshold of 10 trajectories.

For every pixel, a) the percentage of confined molecules, b) the median apparent diffusion coefficient, anomalous exponent, length of confinement, and drift, and c) the mean absolute value of the angle between trajectory steps were calculated. The values computed for each pixel are shown on the scatter plots against the brightness of that pixel in the wide-field channel.

Pearson correlation coefficient was then calculated for all, confined and unconfined points in each dataset (2i/LIF, 24h and 48h, each either with DNA or HP1 $\beta$  labelled), alongside the p-value indicating whether it was significantly different from 0. A linear regression was also computed, and its slope was used to assess the magnitude of the linkage between the diffusion parameters and DNA or HP1 $\beta$  density. Note that since the wide-field brightness values are not comparable between datasets, the comparisons can only be made within datasets. Thus, the slope was only used to show the change

in trends for all data vs confined or unconfined. It was also normalised to simplify the plots.

A similar analysis could be performed for individual points in trajectories, rather than the values averaged per pixel. However, firstly, the pixel brightness is an average characteristic of the pixel, and it is better practice to compare like with like. Secondly, SPT data is inherently probabilistic and noisy, and the experimental error adds up to it. The averaging procedure helps to lower the variation in the data and to calculate values characterising the environment in a particular region. Furthermore, correlation measurements are especially sensitive to noise.

The analysis was performed in R, with the EBImage package used to read in microscopy images (Pau et al. 2010).

### **Analysis of features on the inter-compartment boundary**

In this analysis, the points (not subtrajectories) classified as unconfined by the 4P algorithm were used, and the individual displacements within trajectories were considered separately. When a localisation was situated in one compartment in one frame, and in another compartment in the next, the molecule was considered to have crossed the inter-compartment boundary. The distribution of the length of the inter-compartment displacements, either into or out of heterochromatin, was compared to that of the particles entering or exiting the euchromatic “pseudo-foci”. The number of the inter-compartment displacements was also divided by the total number of displacements in the entire dataset to estimate the probabilities of these events.

Maps of molecular drift were constructed by dividing points into 2D bins/pixels and calculating the vector sum of displacements in each bin. This was performed for a variety of bin widths from 50 to 500 nm. However, presence of long trajectories and low density of independent molecules lead to artifacts in maps with <250 nm resolution. To investigate general trends on the boundaries of heterochromatic domains, the boundaries were segmented by selecting white (heterochromatic) pixels that had at least one black (euchromatic) and one white neighbour in the binary masks. In both x and y, the direction of the euchromatic neighbour was considered “outwards” from the

focus, and the direction of the heterochromatic neighbour was considered “inwards”. The potential inwards or outwards bias of the average vectors (in the 271 nm×271 nm maps, matching the binary mask pixel size) on the border was then investigated, but found to be insignificant.

To analyse trajectory step angles at the heterochromatin domain boundaries, the boundary was defined as above. The distribution of angles for points belonging to the boundary pixels was then compared to that for particles deeper in heterochromatin.

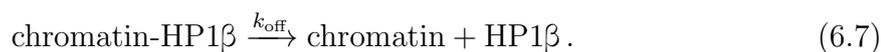
### 6.4.6 Long-exposure SPT data analysis

The jump distance and the 4P analyses were performed similarly to the short-exposure experiments.

#### Residence time analysis

The residence times of individual HP1 $\beta$  molecules on chromatin were defined as their trajectory lengths (>2 frames), and their distributions were plotted to compare dissociation kinetics in different conditions.

The characteristic residence time  $\tau_{\text{res}}$  and its reciprocal dissociation rate constant  $k_{\text{off}} = 1/\tau_{\text{res}}$  describe the probabilistic Poisson process of dissociation:



In the context of single-molecule data, the normalised distribution of the individual residence times  $f(t)$  follows:

$$f(t) = e^{-k_{\text{off}}t} = e^{\frac{-t}{\tau_{\text{res}}}}. \quad (6.8)$$

If several modes of binding are possible, multiple dissociation processes with different kinetics take place. Thus, the overall kinetics is described by their weighted sum:

$$f(t) = \sum_{i=1}^N A_i e^{-k_{\text{off},i}t} = \sum_{i=1}^N A_i e^{\frac{-t}{\tau_{\text{res},i}}}, \quad (6.9)$$

where  $N$  is the number of different processes and  $A_i$  are the relative abundances of these processes, with  $\sum_{i=1}^N A_i = 1$ . A script in python written by Dr Wayne Boucher was used to fit the observed distribution to a one-, two- or three-component dissociation model (<https://github.com/wb104/trajectory-analysis>).

The additional contribution of photobleaching can be accounted for as follows:

$$f(t) = \sum_{i=1}^N A_i e^{-(k_{\text{off},i} + k_b)t}, \quad (6.10)$$

where  $k_b$  is the bleaching rate. Thus, the obtained effective dissociation rates  $\tau_{\text{eff},i}$  were corrected by dividing by  $1 - k_b \tau_{\text{eff},i}$ .

## 6.5 FLIP

### 6.5.1 Experimental protocol

Live HP1 $\beta$ -Halo(N) cells with normal HP1 $\beta$  expression were grown on laminin-coated dishes and labelled with JF<sub>549</sub>, as described above. A Leica SP8 laser-scanning confocal microscope at CAIC was used to perform the experiment. 100% of the 561 nm laser and 50% of the 514 nm laser intensity with FRAP boosters were used for bleaching, while acquisition was conducted with 561 nm laser at 0.3% intensity. Imaging was performed at 400 Hz scan frequency and without line averaging to obtain maximum time resolution. The pixel size was 90 nm  $\times$  90 nm.

Cells containing at least two HP1 $\beta$  foci were chosen for imaging, and the fields of view also always contained a portion of a neighbouring cell, used as a control. The bleaching spot with an area of 1  $\mu\text{m}^2$  was selected in euchromatin, outside of HP1 $\beta$  foci and nucleoli. Each cell was initially imaged for 10 frames. Next, targeted bleaching was performed for one frame, and one frame was acquired under normal illumination. This cycle was repeated 200 times, resulting in a substantial drop in fluorescence intensity. The acquisition time for every frame was 650 ms, so the data was acquired with 1.3 s resolution for 260 s in total. Five cells were recorded for every condition.

## Data analysis

For every cell, three euchromatic and three heterochromatic regions (or two, if there were only two foci) were selected,  $1\ \mu\text{m}^2$  each. To accurately measure the loss in fluorescence in heterochromatin, the movement of foci was tracked semi-manually, using the Fiji plugin Time Lapse. To avoid potential bias, the regions for measurement were chosen such that the distances from the bleached area to the euchromatic and heterochromatic spots were comparable. Another  $1\ \mu\text{m}^2$  spot was selected in a neighbouring cell as a control. In addition, the fluorescence intensity drop in the bleached area was also measured.

The individual decay curves were normalised by their maximum intensity values. Mean curves and their standard deviation for all bleached, euchromatic, heterochromatic and control regions were calculated and plotted for each differentiation state. In order to test the statistical significance of differences between collections of timelapse curves, one must fit an appropriate model to each curve and compare the resultant parameters. However, the data was extremely noisy, and fitting of individual curves was very imprecise. Thus, statistical testing was not performed.

## 6.6 3D DH-PSF SMLM of fixed cells, combined with snHi-C

### 6.6.1 Sample preparation

Different sample preparation and imaging protocols were used to produce super-resolved images of proteins or DNA. For protein PALM, mESC lines expressing HaloTag fusion proteins were used: NANOG-Halo, SOX2-Halo, or HP1 $\beta$ -Halo(N); while the CB9 line was typically imaged in DNA PAINT experiments. Below, the optimised sample preparation protocols are described, while any deviations from them are marked in the relevant sections of chapter 4.

## Labelling

For PALM imaging, cells were labelled with PA-JF<sub>646</sub> in the same way as for sptPALM. They were then harvested and fixed in suspension with 2% formaldehyde for 10 minutes. Following quenching with 100 mM glycine and washing with PBS, cells were incubated with 0.1% TritonX-100 for 4 min and washed with PBS again. Cells were resuspended in ice-cold freshly prepared nuclear extraction buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% NP-40 and protease inhibitor (Roche 11873580001)) and incubated on ice for 30 min, with gentle mixing by inversion performed every 10 min. The cells were once again washed in PBS and resuspended in NEBuffer 3 (NEB B7003S) for sorting (see below). After FACS, 12.5 µl of NEBuffer 3 were added to each well, the plate was sealed with film and imaged. The snHi-C protocol was then continued as in Stevens et al. (2017).

MESCs for DNA PAINT imaging were harvested and fixed, and resuspended in NEBuffer 3 before FACS. 12.5 µl of the MEA buffer (100 mM Tris-HCl pH 8.5, 50 mM 2-mercaptoethylamine (MEA; Sigma M9768), 10 mM NaCl), prepared fresh and containing 400 pM of 5-SiR-Hoechst dye, were added to every well. Nuclear extraction and the rest of the snHi-C procedure were performed after microscopy.

For trial experiments, such as optimisations of the laser intensities and labelling protocols, the cell suspension was placed onto glass coverslips for imaging rather than FACS-sorted.

## FACS

Individual fixed mESCs were sorted into wells of a 384-well optical cyclic olefine-bottom plate (Perkin Elmer 6057302), using the BD Aria III or Beckman Coulter MoFlo Astrios instruments at the Cambridge University Flow Cytometry Facility. Only the inner area of the plate was used (excluding outer 5 rows and columns), since the microscope objective could not reach the edges of the plate. Using the SSC-area vs FSC-area and FSC-height vs FSC-area plots, single non-clumped cells were selected, while gates were set on the SSC-area parameter to enrich for mESCs in G1 phase by their size.

## 6.6.2 Microscopy

### Microscope setup

The microscope employed for the 3D DH-PSF SMLM experiments was described in section 6.4.2. A 10× air objective (CFI Plan Fluor, Nikon) was used to find the position of individual nuclei or cells in wells, while the 60× water immersion objective and double helix phase mask for far-red emission were used for SMLM. In addition, a 530 nm monochromatic lamp was fitted above the stage to provide illumination for bright-field drift correction.

### Calibration sequence

Before each experiment, a calibration sequence was acquired, using fluorescent beads (TetraSpeck Microspheres 0.1 μm, Invitrogen T7279) on the bottom of a 384-well plate identical to the one used in experiments. 121 positions along the z axis were imaged, 40 nm apart, centered about the best-focus point. 10 frames were acquired for each position to even out any potential random noise.

### Locating cells

An automatic procedure was developed to acquire and stitch low-magnification bright-field images of the entire plate. Cells or nuclei were then picked out manually, and their coordinates recorded. They were then adjusted after switching to the 60× objective lens.

### Imaging conditions and procedure

For protein PALM with PA-JF<sub>646</sub>, 638 nm and 405 nm lasers were typically used at 16 kW/cm<sup>2</sup> and 0.0002-0.0006 kW/cm<sup>2</sup> intensities, respectively, unless otherwise specified, while to image 5-SiR-Hoechst, full power of the 638 nm excitation laser (40 kW/cm<sup>2</sup>) was required. All acquisition was performed in the HILO regime. 20 ms exposure time was used for PALM, and 40 ms for DNA PAINT.

Z-stacks with overlapping slices were used to cover the entire nucleus in 3D. Typically, 12 slices (sometimes 10-13) were imaged, with a relative shift of 0.8  $\mu\text{m}$  in  $z$ . The stage was programmed to move between slices in a staggered order, e.g. for 12 slices, the order was 1, 7, 2, 8, 3, 9, 4, 10, 5, 11, 6, 12. A 1 s pause was introduced after each movement to ensure the stage was stable before starting the acquisition. Imaging was performed in 10-20 cycles, with 500 (protein PALM) or 250 (DNA PAINT) frames collected for each slice in each cycle.

At the beginning and end of each cycle, 20 bright-field images were acquired for drift correction, using the monochromatic green lamp. The transmitted light was deflected to a second camera by a dichroic mirror.

3D SMLM data for each cell was acquired sequentially, and about 15 cells were usually imaged per session.

### 6.6.3 Data analysis

#### Localisation detection

The 3D coordinates of each molecule were determined, using the information from the calibration series. At the trial stage, easy-DHPSF (Lew et al. 2013) and SMAP2018 (Li et al. 2018; Ries 2020) were applied to the data, according to the authors' instructions. The two-step fitting procedure that was selected as the most reliable involved initial detection of Gaussian-shaped peaks with GDSC PeakFit plugin (Herbert 2014) and subsequent 3D coordinate computation and quality filtering with DHPSFU (<https://github.com/TheLaueLab/DHPSFU>). The parameters used in the trials, as well as the final parameter set, are shown in Tables 6.14 and 6.15.

Only localisations with the DH-PSF pitch within a certain range were accepted to ensure correct mapping. The range always corresponded to 2.4  $\mu\text{m}$  in  $z$ , but the exact start and end angles were selected anew for every dataset, using the calibration data. The 2.4  $\mu\text{m}$  thickness was chosen to match the 0.8  $\mu\text{m}$  step between the slices, so that any area inside the cell (except the very top or bottom) was covered by exactly four slices.

**Table 6.14:** Parameters used in GDSC PeakFit for fitting 3D SMLM data. All other parameters were kept at default values. Parameter set 1 was selected as the best-performing during trials (underlined).

Parameter	<u>Set1</u>	Set2
Camera type	EMCCD	
Read noise	7.1622	
Camera bias	400	
Gain	35.0	
Quantum efficiency	0.95	
Smoothing	0.32	0.4
Search width	0.9	1.0
Border width	0.9	1.0
Fitting width	2.15	3.0
Neighbour height	0.3	0.3
Residuals threshold	1.0	1.0
Duplicate distance	0.5	0.2
Shift factor	1.2	1.75
Min photons	10	200
Min & max width	0.4 & 1.7	0.2 & 3.0
Precision	30, <u>40</u> , 50, 70	70

**Table 6.15:** Parameters used in DHPSFU trials. Set 1 (underlined) was chosen for fitting 3D SMLM data.

Parameter	<u>Set1</u>	Set2	Set3	Set4	Set5	Set6
Initial distance filter (px)	[8 3]	[8 3]	[6 3.5]	[6 3.5]	[8 3]	[6 3.5]
Max distance deviation	0.2	0.1	0.2	0.1	0.2	0.1
Max intensity ratio deviation	1	1	1	1	0.5	0.5

## Removal of duplicates

The tracking code (<https://github.com/wb104/trajectory-analysis>) employed to join points into trajectories in live SPT data was also applied to the SMLM localisations to average the coordinate of the same molecule determined in several consecutive frames. Although the majority of fluorophores were bleached within one or a few frames, some lasted for  $>50$ , and thus the maximum frame gap for two localisations to be considered the same molecule was increased to 50. 500 nm were used as the maximum distance, and this radius was not adjusted based on the frame gap, since molecules were not diffusing.

## Drift correction

The drift correction was performed by cross-correlating the bright-field images of cells, using a script in Matlab (<https://github.com/TheLaueLab/DFT-drift-correction>). Discrete Fourier transform was used to compute the shift between the images, and subpixel precision was reached using a fast algorithm for local upsampling (Guizar-Sicairos et al. 2008).

The basic method, developed by Dr Aleks Ponjavic, was extended to allow occasional drift correction with averaging. Bright-field snapshots were taken only occasionally, rather than continuously, and the drift between them was inferred by interpolation. Acquiring multiple images in a row increased the precision of the procedure, since the computed shift was averaged between them, reducing random inaccuracies in cross-correlation.

## Stitching and rendering

The lists of localisations from different slices were combined, adding the appropriate shifts to the z-coordinates. The different features of the datasets were calculated and plotted in R. To reconstruct the images, the localisations were split into 1  $\mu\text{m}$ -thick slices in x, y or z dimensions. Note that these slices were different from the acquisition slices: they were non-overlapping, and could include localisations from different

acquisition slices. An image for each slice was then rendered as a density plot, using 2D kernel density estimation with a 500 nm bandwidth. This bandwidth was selected empirically, since 2D histograms with 500 nm×500 nm bin tended to have about 8-10 molecules per bin. The density was multiplied by the total number of molecules in the slice to transform the units into molecules per unit area.

## Cluster analysis

Ripley’s K function (Ripley 1977) is the density-normalised expectation of the number of points within distance  $r$  from a randomly chosen point. For  $n$  points, the K function is calculated as

$$K(r) = \frac{1}{\lambda} \left( \frac{1}{n} \sum_{i=1}^n N_{p_i}(r) \right), \quad (6.11)$$

where  $p_i$  is the  $i$ th point in the dataset,  $N_{p_i}(r)$  is the number of neighbours of point  $p_i$  within radius  $r$ , and  $\lambda$  is the spatial density of points.

The H function (Ehrlich et al. 2004) is a normalisation of Ripley’s K, which takes positive or negative values when the points display clustering or dispersion, respectively, and equals zero when the distribution is totally random. In 3D, it is

$$H(r) = \left( \frac{3}{4} \frac{K(r)}{\pi} \right)^{1/3} - r. \quad (6.12)$$

In this work, Ripley’s K functions were calculated in 3D for each cell, using the spatstat package in R (Baddeley and Turner 2005), and then converted into the H functions.

Ripley’s functions estimated for real datasets suffer from the edge effect, because points at the edge have fewer neighbours and thus contribute a smaller  $N_{p_i}(r)$  term into the calculation. Here, the effect was corrected, using the isotropic method (Ripley 1977; Baddeley et al. 1993). To perform edge correction, one must specify the region in which the points are distributed. Since fitting complex meshes to 3D data was beyond the scope of this work, box-shaped regions  $4 \times 4 \times 2 \mu\text{m}^3$  in the middle of cells (avoiding nucleoli) were selected for analysis.

#### 6.6.4 Control experiments with beads

To assess the accuracy of the method, two controls with fluorescent Tetraspeck beads were performed. In the first one, beads were suspended in drops of 2% agarose on coverslips, creating a 3D sample. The beads were then imaged with the DH-PSF microscope in 12 slices, repeated 3 times, with 25 frames per slice per cycle. 10 positions were recorded.

In the second experiment, beads were placed onto a glass coverslip. Three fields of view were selected, each of them containing a bead roughly in the same position in the middle. A calibration sequence was recorded twice for each position. Then, the stage was systematically moved 1  $\mu\text{m}$  or 10  $\mu\text{m}$  up, down, left or right from the central position. This allowed assessment of the reproducibility of the z-to-angle mapping with a) the same bead at the same position within the field of view, b) different beads at the same position, or c) the same bead at different positions within the field of view.

# References

- Akhtar, Asifa, and Susan M. Gasser. 2007. “The nuclear envelope and transcriptional control”. *Nature Reviews Genetics* 8, no. 7 (): 507–517. doi:10.1038/nrg2122.
- Alberti, Simon. 2017. “Phase separation in biology”. *Current Biology* 27, no. 20 (): R1097–R1102. doi:10.1016/j.cub.2017.08.069.
- Aucott, Rebecca, Jörn Bullwinkel, Yang Yu, Wei Shi, Mustafa Billur, Jeremy P. Brown, Ursula Menzel, Dimitris Kioussis, Guozheng Wang, Ingrid Reisert, Jörg Weimer, Raj K. Pandita, Girdhar G. Sharma, Tej K. Pandita, Reinald Fundele, and Prim B. Singh. 2008. “HP1-*extbeta* is required for development of the cerebral neocortex and neuromuscular junctions”. *Journal of Cell Biology* 183, no. 4 (): 597–606. doi:10.1083/jcb.200804041.
- Ayyanathan, Kasirajan, Mark S. Lechner, Peter Bell, Gerd G. Maul, David C. Schultz, Yoshihiko Yamada, Kazuhiro Tanaka, Kiyoyuki Torigoe, and Frank J. Rauscher. 2003. “Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation”. *Genes & Development* 17, no. 15 (): 1855–1869. doi:10.1101/gad.1102803.
- Baddeley, A. J., R. A. Moeved, C. V. Howard, and A. Boyde. 1993. “Analysis of a Three-Dimensional Point Pattern with Replication”. *Applied Statistics* 42 (4): 641. doi:10.2307/2986181.
- Baddeley, Adrian, and Rolf Turner. 2005. “spatstat: AnRPackage for Analyzing Spatial Point Patterns”. *Journal of Statistical Software* 12 (6). doi:10.18637/jss.v012.i06.
- Balzarotti, Francisco, Yvan Eilers, Klaus C. Gwosch, Arvid H. Gynnå, Volker Westphal, Fernando D. Stefani, Johan Elf, and Stefan W. Hell. 2016. “Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes”. *Science* 355, no. 6325 (): 606–612. doi:10.1126/science.aak9913.
- Banani, Salman F., Hyun O. Lee, Anthony A. Hyman, and Michael K. Rosen. 2017. “Biomolecular condensates: organizers of cellular biochemistry”. *Nature Reviews Molecular Cell Biology* 18, no. 5 (): 285–298. doi:10.1038/nrm.2017.7.
- Bancaud, Aurélien, Sébastien Huet, Nathalie Daigle, Julien Mozziconacci, Joël Beaudouin, and Jan Ellenberg. 2009. “Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin”. *The EMBO Journal* 28, no. 24 (): 3785–3798. doi:10.1038/emboj.2009.340.
- Bannister, Andrew J., Philip Zegerman, Janet F. Partridge, Eric A. Miska, Jean O. Thomas, Robin C. Allshire, and Tony Kouzarides. 2001. “Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain”. *Nature* 410, no. 6824 (): 120–124. doi:10.1038/35065138.
- Banterle, Niccolò, Khanh Huy Bui, Edward A. Lemke, and Martin Beck. 2013. “Fourier ring correlation as a resolution criterion for super-resolution microscopy”. *Journal of Structural Biology* 183, no. 3 (): 363–367. doi:10.1016/j.jsb.2013.05.004.
- Barth, R., K. Bystricky, and H. A. Shaban. 2020. “Coupling chromatin structure and dynamics by live super-resolution imaging”. *Science Advances* 6, no. 27 (). doi:10.1126/sciadv.aaz2196.

- Bártová, Eva, Jana Krejčí, Andrea Harničarová, and Stanislav Kozubek. 2008. “Differentiation of human embryonic stem cells induces condensation of chromosome territories and formation of heterochromatin protein 1 foci”. *Differentiation* 76 (1): 24–32. doi:10.1111/j.1432-0436.2007.00192.x.
- Basu, Srinjan, et al. 2020. “Live-cell 3D single-molecule tracking reveals how NuRD modulates enhancer dynamics” (). doi:10.1101/2020.04.03.003178.
- Beddington, R.S., and E.J. Robertson. 1989. “An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo”. *Development* 105, no. 4 (): 733–737. doi:10.1242/dev.105.4.733.
- Bednar, Jan, et al. 2017. “Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1”. *Molecular Cell* 66, no. 3 (): 384–397.e8. doi:10.1016/j.molcel.2017.04.012.
- Belly, Henry De, Aki Stubb, Ayaka Yanagida, Céline Labouesse, Philip H. Jones, Ewa K. Paluch, and Kevin J. Chalut. 2021. “Membrane Tension Gates ERK-Mediated Regulation of Pluripotent Cell Fate”. *Cell Stem Cell* 28, no. 2 (): 273–284.e6. doi:10.1016/j.stem.2020.10.018.
- Bergert, Martin, Sergio Lembo, Sumana Sharma, Luigi Russo, Danica Milovanović, Kristjan H. Grestarsson, Mandy Börmel, Pierre A. Neveu, Jamie A. Hackett, Evangelia Petsalaki, and Alba Diz-Muñoz. 2021. “Cell Surface Mechanics Gate Embryonic Stem Cell Differentiation”. *Cell Stem Cell* 28, no. 2 (): 209–216.e4. doi:10.1016/j.stem.2020.10.017.
- Betschinger, Joerg, Jennifer Nichols, Sabine Dietmann, Philip D. Corrin, Patrick J. Paddison, and Austin Smith. 2013. “Exit from Pluripotency Is Gated by Intracellular Redistribution of the bHLH Transcription Factor Tfe3”. *Cell* 153, no. 2 (): 335–347. doi:10.1016/j.cell.2013.03.012.
- Bo, Stefano, Lars Hubatsch, Jonathan Bauermann, Christoph A. Weber, and Frank Julicher. 2021. “Stochastic dynamics of single molecules across phase boundaries” (). doi:10.1101/2021.04.27.441464.
- Bonev, Boyan, Netta Mendelson Cohen, Quentin Szabo, Lauriane Fritsch, Giorgio L. Papadopoulos, Yaniv Lubling, Xiaole Xu, Xiaodan Lv, Jean-Philippe Hugnot, Amos Tanay, and Giacomo Cavalli. 2017. “Multiscale 3D Genome Rewiring during Mouse Neural Development”. *Cell* 171, no. 3 (): 557–572.e24. doi:10.1016/j.cell.2017.09.043.
- Boroviak, Thorsten, Remco Loos, Paul Bertone, Austin Smith, and Jennifer Nichols. 2014. “The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification”. *Nature Cell Biology* 16, no. 6 (): 513–525. doi:10.1038/ncb2965.
- Bosch-Presegué, Laia, Helena Raurell-Vila, Joshua K. Thackray, Jessica González, Carmen Casal, Noriko Kane-Goldsmith, Miguel Vizoso, Jeremy P. Brown, Antonio Gómez, Juan Ausió, Timo Zimmermann, Manel Esteller, Gunnar Schotta, Prim B. Singh, Lourdes Serrano, and Alejandro Vaquero. 2017. “Mammalian HP1 Isoforms Have Specific Roles in Heterochromatin Structure and Organization”. *Cell Reports* 21, no. 8 (): 2048–2057. doi:10.1016/j.celrep.2017.10.092.
- Bradley, Allan, Martin Evans, Matthew H. Kaufman, and Elizabeth Robertson. 1984. “Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines”. *Nature* 309, no. 5965 (): 255–256. doi:10.1038/309255a0.
- Brasher, Sally V., Brian O. Smith, Rasmus H. Fogh, Daniel Nietlispach, Abarna Thiru, Peter R. Nielsen, R. William Broadhurst, Linda J. Ball, Natalia V. Murzina, and Ernest D. Laue. 2000. “The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer”. *The EMBO Journal* 19, no. 7 (): 1587–1597. doi:10.1093/emboj/19.7.1587.
- Brons, I. Gabrielle M., Lucy E. Smithers, Matthew W. B. Trotter, Peter Rugg-Gunn, Bowen Sun, Susana M. Chuva de Sousa Lopes, Sarah K. Howlett, Amanda Clarkson, Lars Ahrlund-Richter, Roger A. Pedersen, and Ludovic Vallier. 2007. “Derivation of pluripotent epiblast stem cells from mammalian embryos”. *Nature* 448, no. 7150 (): 191–195. doi:10.1038/nature05950.

- Bryan, Louise C., Daniel R. Weilandt, Andreas L. Bachmann, Sinan Kilic, Carolin C. Lechner, Pascal D. Odermatt, Georg E. Fantner, Sandrine Georgeon, Oliver Hantschel, Vassily Hatzimanikatis, and Beat Fierz. 2017. “Single-molecule kinetic analysis of HP1-chromatin binding reveals a dynamic network of histone modification and DNA interactions”. *Nucleic Acids Research* 45, no. 18 (): 10504–10517. doi:10.1093/nar/gkx697.
- Bucevičius, Jonas, Tanja Gilat, and Gražvydas Lukinavičius. 2020. “Far-red switching DNA probes for live cell nanoscopy”. *Chemical Communications* 56 (94): 14797–14800. doi:10.1039/d0cc06759h.
- Bucevičius, Jonas, Jan Keller-Findeisen, Tanja Gilat, Stefan W. Hell, and Gražvydas Lukinavičius. 2019. “Rhodamine–Hoechst positional isomers for highly efficient staining of heterochromatin”. *Chemical Communications* 10 (7): 1962–1970. doi:10.1039/c8sc05082a.
- Bunch, Heeyoun, Xiaofeng Zheng, Adam Burkholder, Simon T Dillon, Shmulik Motola, Gabriel Birrane, Christopher C Ebmeier, Stuart Levine, David Fargo, Guang Hu, Dylan J Taatjes, and Stuart K Calderwood. 2014. “TRIM28 regulates RNA polymerase II promoter-proximal pausing and pause release”. *Nature Structural & Molecular Biology* 21, no. 10 (): 876–883. doi:10.1038/nsmb.2878.
- Cammas, Florence, Marielle Herzog, Thierry Lerouge, Pierre Chambon, and Régine Losson. 2004. “Association of the transcriptional corepressor TIF1b with heterochromatin protein 1 ( HP1 ): an essential role for progression through differentiation”. *Genes & development* 18 (18): 2147–2160. doi:10.1101/gad.302904.TIF1.
- Cammas, Florence, Agnes Janoshazi, Thierry Lerouge, and Régine Losson. 2007. “Dynamic and selective interactions of the transcriptional corepressor TIF1 $\beta$  with the heterochromatin protein HP1 isoforms during cell differentiation”. *Differentiation* 75 (7): 627–637. doi:10.1111/j.1432-0436.2007.00166.x.
- Canzio, Daniele, Adam Larson, and Geeta J. Narlikar. 2014. “Mechanisms of functional promiscuity by HP1 proteins”. *Trends in Cell Biology* 24, no. 6 (): 377–386. doi:10.1016/j.tcb.2014.01.002.
- Canzio, Daniele, Maofu Liao, Nariman Naber, Edward Pate, Adam Larson, Shenping Wu, Diana B. Marina, Jennifer F. Garcia, Hiten D. Madhani, Roger Cooke, Peter Schuck, Yifan Cheng, and Geeta J. Narlikar. 2013. “A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly”. *Nature* 496, no. 7445 (): 377–381. doi:10.1038/nature12032.
- Carr, Alexander R., Aleks Ponjavic, Srinjan Basu, James McColl, Ana Mafalda Santos, Simon Davis, Ernest D. Laue, David Klenerman, and Steven F. Lee. 2017. “Three-Dimensional Super-Resolution in Eukaryotic Cells Using the Double-Helix Point Spread Function”. *Biophysical Journal* 112, no. 7 (): 1444–1454. doi:10.1016/j.bpj.2017.02.023.
- Chalut, Kevin J., Markus Höpfler, Franziska Lautenschläger, Lars Boyde, Chii Jou Chan, Andrew Ekpenyong, Alfonso Martinez-Arias, and Jochen Guck. 2012. “Chromatin Decondensation and Nuclear Softening Accompany Nanog Downregulation in Embryonic Stem Cells”. *Biophysical Journal* 103, no. 10 (): 2060–2070. doi:10.1016/j.bpj.2012.10.015.
- Cheutin, Thierry, Adrian J McNairn, Thomas Jenuwein, David M Gilbert, Prim B Singh, and Tom Misteli. 2003. “Maintenance of Stable Heterochromatin Domains by Dynamic HP1 Binding”. *Science* 299, no. 5607 (): 721–725. doi:10.1126/science.1078572.
- Chin, Hang Gyeong, Pierre-Olivier Estève, Mihika Pradhan, Jack Benner, Debasis Patnaik, Michael F. Carey, and Sriharsa Pradhan. 2007. “Automethylation of G9a and its implication in wider substrate specificity and HP1 binding”. *Nucleic Acids Research* 35, no. 21 (): 7313–7323. doi:10.1093/nar/gkm726.
- Chittock, Emily C., Sebastian Latwiel, Thomas C.R. Miller, and Christoph W. Müller. 2017. “Molecular architecture of polycomb repressive complexes”. *Biochemical Society Transactions* 45, no. 1 (): 193–205. doi:10.1042/bst20160173.

- Chowdhury, Farhan, Yanzhen Li, Yeh-Chuin Poh, Tamaki Yokohama-Tamaki, Ning Wang, and Tet-suya S. Tanaka. 2010. “Soft Substrates Promote Homogeneous Self-Renewal of Embryonic Stem Cells via Downregulating Cell-Matrix Traction”. Ed. by Zhongjun Zhou. *PLoS ONE* 5, no. 12 (): e15655. doi:10.1371/journal.pone.0015655.
- Collombet, Samuel, Isabell Rall, Claire Dugast-Darzacq, Alec Heckert, Aliaksandr Halavatyi, Agnes Le Saux, Gina Dailey, Xavier Darzacq, and Edith Heard. 2021. “RNA polymerase II depletion from the inactive X chromosome territory is not mediated by physical compartmentalization” (). doi:10.1101/2021.03.26.437188.
- Cremer, T., and M. Cremer. 2010. “Chromosome Territories”. *Cold Spring Harbor Perspectives in Biology* 2, no. 3 (): a003889–a003889. doi:10.1101/cshperspect.a003889.
- Daujat, Sylvain, Ulrike Zeissler, Tanja Waldmann, Nicole Happel, and Robert Schneider. 2005. “HP1 Binds Specifically to Lys26-methylated Histone H1.4, whereas Simultaneous Ser27 Phosphorylation Blocks HP1 Binding”. *Journal of Biological Chemistry* 280, no. 45 (): 38090–38095. doi:10.1074/jbc.c500229200.
- Dialynas, G. K., S. Terjung, J. P. Brown, R. L. Aucott, B. Baron-Luhr, P. B. Singh, and S. D. Georgatos. 2007. “Plasticity of HP1 proteins in mammalian cells”. *Journal of Cell Science* 120 (19): 3415–3424. doi:10.1242/jcs.012914.
- Dixon, Jesse R., Inkyung Jung, Siddarth Selvaraj, Yin Shen, Jessica E. Antosiewicz-Bourget, Ah Young Lee, Zhen Ye, Audrey Kim, Nisha Rajagopal, Wei Xie, Yarui Diao, Jing Liang, Huimin Zhao, Victor V. Lobanenko, Joseph R. Ecker, James A. Thomson, and Bing Ren. 2015. “Chromatin architecture reorganization during stem cell differentiation”. *Nature* 518, no. 7539 (): 331–336. doi:10.1038/nature14222.
- Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. 2012. “Topological domains in mammalian genomes identified by analysis of chromatin interactions”. *Nature* 485, no. 7398 (): 376–380. doi:10.1038/nature11082.
- Ehrlich, Marcelo, Werner Boll, Antoine van Oijen, Ramesh Hariharan, Kartik Chandran, Max L. Nibert, and Tomas Kirchhausen. 2004. “Endocytosis by Random Initiation and Stabilization of Clathrin-Coated Pits”. *Cell* 118, no. 5 (): 591–605. doi:10.1016/j.cell.2004.08.017.
- Eissenberg, J. C., T. C. James, D. M. Foster-Hartnett, T. Hartnett, V. Ngan, and S. C. Elgin. 1990. “Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*.” *Proceedings of the National Academy of Sciences* 87, no. 24 (): 9923–9927. doi:10.1073/pnas.87.24.9923.
- Ekwall, Karl, Jean-Paul Javerzat, Axel Lorentz, Henning Schmidt, Gwen Cranston, and Robin Allshire. 1995. “The Chromodomain Protein Swi6: A Key Component at Fission Yeast Centromeres”. *Science* 269, no. 5229 (): 1429–1431. doi:10.1126/science.7660126.
- Elgin, S. C. R., and G. Reuter. 2013. “Position-Effect Variegation, Heterochromatin Formation, and Gene Silencing in *Drosophila*”. *Cold Spring Harbor Perspectives in Biology* 5, no. 8 (): a017780–a017780. doi:10.1101/cshperspect.a017780.
- Ellis, B, Perry Haaland, Florian Hahne, Nathan Le Meur, Nishant Gopalakrishnan, Josef Spidlen, Mike Jiang, and Greg Finak. 2019. *flowCore: Basic structures for flow cytometry data*. R package version 1.52.1.
- Erdel, Fabian, Anne Rademacher, Rifka Vlijm, Jana Tünnermann, Lukas Frank, Robin Weinmann, Elisabeth Schweigert, Klaus Yserentant, Johan Hummert, Caroline Bauer, Sabrina Schumacher, Ahmad Al Alwash, Christophe Normand, Dirk-Peter Herten, Johann Engelhardt, and Karsten Rippe. 2020. “Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of HP1-Driven Liquid-Liquid Phase Separation”. *Molecular Cell* 78, no. 2 (): 236–249.e7. doi:10.1016/j.molcel.2020.02.005.

- Erdmann, Roman S., Stephanie Wood Baguley, Jennifer H. Richens, Rebecca F. Wissner, Zhiqun Xi, Edward S. Allgeyer, Sheng Zhong, Alexander D. Thompson, Nicholas Lowe, Richard Butler, Joerg Bewersdorf, James E. Rothman, Daniel St Johnston, Alanna Schepartz, and Derek Toomre. 2019. “Labeling Strategies Matter for Super-Resolution Microscopy: A Comparison between HaloTags and SNAP-tags”. *Cell Chemical Biology* 26, no. 4 (): 584–592.e6. doi:10.1016/j.chembio.2019.01.003.
- Etheridge, Thomas J., Rémi L. Boulineau, Alex Herbert, Adam T. Watson, Yasukazu Daigaku, Jem Tucker, Sophie George, Peter Jönsson, Matthieu Palayret, David Lando, Ernest Laue, Mark A. Osborne, David Klenerman, Steven F. Lee, and Antony M. Carr. 2014. “Quantification of DNA-associated proteins inside eukaryotic cells using single-molecule localization microscopy”. *Nucleic Acids Research* 42, no. 19 (): e146–e146. doi:10.1093/nar/gku726.
- Evans, M. J., and M. H. Kaufman. 1981. “Establishment in culture of pluripotential cells from mouse embryos”. *Nature* 292, no. 5819 (): 154–156. doi:10.1038/292154a0.
- Fakan, Stanislav, and Roel van Driel. 2007. “The perichromatin region: A functional compartment in the nucleus that determines large-scale chromatin folding”. *Seminars in Cell & Developmental Biology* 18, no. 5 (): 676–681. doi:10.1016/j.semcdb.2007.08.010.
- Fan, Jun Y., Danny Rangasamy, Karolin Luger, and David J. Tremethick. 2004. “H2A.Z Alters the Nucleosome Surface to Promote HP1 $\alpha$ -Mediated Chromatin Fiber Folding”. *Molecular Cell* 16, no. 4 (): 655–661. doi:10.1016/j.molcel.2004.10.023.
- Fanti, Laura, Giovanna Giovinazzo, Maria Berloco, and Sergio Pimpinelli. 1998. “The Heterochromatin Protein 1 Prevents Telomere Fusions in Drosophila”. *Molecular Cell* 2, no. 5 (): 527–538. doi:10.1016/s1097-2765(00)80152-5.
- Festenstein, Richard, Stamatis N Pagakis, Kyoko Hiragami, Debbie Lyon, Alain Verreault, Belaid Sekkali, and Dimitris Kioussis. 2003. “Modulation of Heterochromatin Protein 1 Dynamics in Primary Mammalian Cells”. *Science* 299, no. 5607 (): 719–721. doi:10.1126/science.1078694.
- Finak, Greg, and Mike Jiang. 2020. *flowWorkspace: Infrastructure for representing and interacting with gated and ungated cytometry data sets*. R package version 3.34.1.
- Finan, Kieran, Anika Raulf, and Mike Heilemann. 2015. “A Set of Homo-Oligomeric Standards Allows Accurate Protein Counting”. *Angewandte Chemie International Edition* 54, no. 41 (): 12049–12052. doi:10.1002/anie.201505664.
- Fischle, Wolfgang, Boo Shan Tseng, Holger L. Dormann, Beatrix M. Ueberheide, Benjamin A. Garcia, Jeffrey Shabanowitz, Donald F. Hunt, Hironori Funabiki, and C. David Allis. 2005. “Regulation of HP1–chromatin binding by histone H3 methylation and phosphorylation”. *Nature* 438, no. 7071 (): 1116–1122. doi:10.1038/nature04219.
- Fischle, Wolfgang, Yanming Wang, Steven A. Jacobs, Youngchang Kim, C. David Allis, and Sepideh Khorasanizadeh. 2003. “Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains”. *Genes & Development* 17, no. 15 (): 1870–1881. doi:10.1101/gad.1110503.
- Fuks, F. 2003. “The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase”. *Nucleic Acids Research* 31, no. 9 (): 2305–2312. doi:10.1093/nar/gkg332.
- Gebhardt, J Christof M, David M Suter, Rahul Roy, Ziqing W Zhao, Alec R Chapman, Srinjan Basu, Tom Maniatis, and X Sunney Xie. 2013. “Single-molecule imaging of transcription factor binding to DNA in live mammalian cells”. *Nature Methods* 10, no. 5 (): 421–426. doi:10.1038/nmeth.2411.
- Geula, Shay, et al. 2015. “m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation”. *Science* 347, no. 6225 (): 1002–1006. doi:10.1126/science.1261417.
- Gibson, Bryan A., Lynda K. Doolittle, Maximillian W.G. Schneider, Liv E. Jensen, Nathan Gamarra, Lisa Henry, Daniel W. Gerlich, Sy Redding, and Michael K. Rosen. 2019. “Organization of Chromatin by Intrinsic and Regulated Phase Separation”. *Cell* 179, no. 2 (): 470–484.e21. doi:10.1016/j.cell.2019.08.037.

- Gong, Wanjun, Pintu Das, Soham Samanta, Jia Xiong, Wenhui Pan, Zhenyu Gu, Jianguo Zhang, Junle Qu, and Zhigang Yang. 2019. “Redefining the photo-stability of common fluorophores with triplet state quenchers: mechanistic insights and recent updates”. *Chemical Communications* 55 (60): 8695–8704. doi:10.1039/c9cc02616a.
- Grimm, Jonathan B, Brian P English, Heejun Choi, Anand K Muthusamy, Brian P Mehl, Peng Dong, Timothy A Brown, Jennifer Lippincott-Schwartz, Zhe Liu, Timothée Lionnet, and Luke D Lavis. 2016. “Bright photoactivatable fluorophores for single-molecule imaging”. 13, no. 12 (): 985–988. doi:10.1038/nmeth.4034.
- Gu, Bo, Tomek Swigut, Andrew Spencley, Matthew R. Bauer, Mingyu Chung, Tobias Meyer, and Joanna Wysocka. 2018. “Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements”. *Science* 359, no. 6379 (): 1050–1055. doi:10.1126/science.aao3136.
- Guizar-Sicairos, Manuel, Samuel T. Thurman, and James R. Fienup. 2008. “Efficient subpixel image registration algorithms”. *Optics Letters* 33, no. 2 (): 156. doi:10.1364/ol.33.000156.
- Guo, Ge, Jian Yang, Jennifer Nichols, John Simon Hall, Isobel Eyres, William Mansfield, and Austin Smith. 2009. “Klf4 reverts developmentally programmed restriction of ground state pluripotency”. *Development* 136 (7): 1063–1069. doi:10.1242/dev.030957. eprint: <https://dev.biologists.org/content/136/7/1063.full.pdf>.
- Hackett, Jamie A., and M. Azim Surani. 2014. “Regulatory Principles of Pluripotency: From the Ground State Up”. *Cell Stem Cell* 15, no. 4 (): 416–430. doi:10.1016/j.stem.2014.09.015.
- Hansen, Anders S, Maxime Woringer, Jonathan B Grimm, Luke D Lavis, Robert Tjian, and Xavier Darzacq. 2018. “Robust model-based analysis of single-particle tracking experiments with Spot-On”. *eLife* 7 (). doi:10.7554/eLife.33125.
- Hansen, Jeffrey C., Kazuhiro Maeshima, and Michael J. Hendzel. 2021. “The solid and liquid states of chromatin”. *Epigenetics & Chromatin* 14, no. 1 (). doi:10.1186/s13072-021-00424-5.
- Harp, Joel M., B. Leif Hanson, David E. Timm, and Gerard J. Bunick. 2000. “Asymmetries in the nucleosome core particle at 2.5Å resolution”. *Acta Crystallographica Section D Biological Crystallography* 56, no. 12 (): 1513–1534. doi:10.1107/s0907444900011847.
- Hediger, Florence, and Susan M Gasser. 2006. “Heterochromatin protein 1: don’t judge the book by its cover!” *Current Opinion in Genetics & Development* 16, no. 2 (): 143–150. doi:10.1016/j.gde.2006.02.013.
- Heltberg, Mathias Luidor, Judith Miné-Hattab, Angela Taddei, Aleksandra M. Walczak, and Thierry Mora. 2021. “Physical observables to determine the nature of membrane-less cellular sub-compartments” (). doi:10.1101/2021.04.01.438041.
- Henninger, Jonathan E., Ozgur Oksuz, Krishna Shrinivas, Ido Sagi, Gary LeRoy, Ming M. Zheng, J. Owen Andrews, Alicia V. Zamudio, Charalampos Lazaris, Nancy M. Hannett, Tong Ihn Lee, Phillip A. Sharp, Ibrahim I. Cissé, Arup K. Chakraborty, and Richard A. Young. 2021. “RNA-Mediated Feedback Control of Transcriptional Condensates”. *Cell* 184, no. 1 (): 207–225.e24. doi:10.1016/j.cell.2020.11.030.
- Herbert, Alex. 2014. “Single Molecule Light Microscopy ImageJ Plugins”.
- Hiragami-Hamada, K., K. Shinmyozu, D. Hamada, Y. Tatsu, K. Uegaki, S. Fujiwara, and J.-i. Nakayama. 2011. “N-Terminal Phosphorylation of HP1 Promotes Its Chromatin Binding”. *Molecular and Cellular Biology* 31, no. 6 (): 1186–1200. doi:10.1128/mcb.01012-10.
- Hiragami-Hamada, Kyoko, Szabolcs Soeroes, Miroslav Nikolov, Bryan Wilkins, Sarah Kreuz, Carol Chen, Inti A. De La Rosa-Velázquez, Hans Michael Zenn, Nils Kost, Wiebke Pohl, Aleksandar Chernev, Dirk Schwarzer, Thomas Jenuwein, Matthew Lorincz, Bastian Zimmermann, Peter Jomo Walla, Heinz Neumann, Tuncay Baubec, Henning Urlaub, and Wolfgang Fischle. 2016. “Dynamic and flexible H3K9me3 bridging via HP1 $\beta$  dimerization establishes a plastic state of condensed chromatin”. *Nature Communications* 7, no. 1 (). doi:10.1038/ncomms11310.

- Hiragami, K., and R. Festenstein. 2005. “Heterochromatin protein 1: a pervasive controlling influence”. *Cellular and Molecular Life Sciences* 62, no. 23 (): 2711–2726. doi:10.1007/s00018-005-5287-9.
- Hummer, Gerhard, Franziska Fricke, and Mike Heilemann. 2016. “Model-independent counting of molecules in single-molecule localization microscopy”. Ed. by Jennifer Lippincott-Schwartz. *Molecular Biology of the Cell* 27, no. 22 (): 3637–3644. doi:10.1091/mbc.e16-07-0525.
- Jacobs, Steven A., and Sepideh Khorasanizadeh. 2002. “Structure of HP1 Chromodomain Bound to a Lysine 9-Methylated Histone H3 Tail”. *Science* 295, no. 5562 (): 2080–2083. doi:10.1126/science.1069473.
- James, T C, and S C Elgin. 1986. “Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene”. *Molecular and Cellular Biology* 6, no. 11 (): 3862–3872. doi:10.1128/mcb.6.11.3862-3872.1986.
- Kaaij, Lucas J.T., Fabio Mohn, Robin H. van der Weide, Elzo de Wit, and Marc Bühler. 2019. “The ChAHP Complex Counteracts Chromatin Looping at CTCF Sites that Emerged from SINE Expansions in Mouse”. *Cell* 178, no. 6 (): 1437–1451.e14. doi:10.1016/j.cell.2019.08.007.
- Kalkan, Tüzer, Nelly Olova, Mila Roode, Carla Mulas, Heather J. Lee, Isabelle Nett, Hendrik Marks, Rachael Walker, Hendrik G. Stunnenberg, Kathryn S. Lilley, Jennifer Nichols, Wolf Reik, Paul Bertone, and Austin Smith. 2017. “Tracking the embryonic stem cell transition from ground state pluripotency”. *Development* 144 (7): 1221–1234. doi:10.1242/dev.142711.
- Keenen, Madeline M, David Brown, Lucy D Brennan, Roman Renger, Harrison Khoo, Christopher R Carlson, Bo Huang, Stephan W Grill, Geeta J Narlikar, and Sy Redding. 2021. “HP1 proteins compact DNA into mechanically and positionally stable phase separated domains”. Ed. by Sebastian Deindl and Kevin Struhl. *eLife* 10 (): e64563. doi:10.7554/eLife.64563.
- Keller, Gordon. 2005. “Embryonic stem cell differentiation: emergence of a new era in biology and medicine”. *Genes & Development* 19, no. 10 (): 1129–1155. doi:10.1101/gad.1303605.
- Khater, Ismail M., Ivan Robert Nabi, and Ghassan Hamarneh. 2020. “A Review of Super-Resolution Single-Molecule Localization Microscopy Cluster Analysis and Quantification Methods”. *Patterns* 1, no. 3 (): 100038. doi:10.1016/j.patter.2020.100038.
- Kilic, Sinan, Andreas L. Bachmann, Louise C. Bryan, and Beat Fierz. 2015. “Multivalency governs HP1 $\alpha$  association dynamics with the silent chromatin state”. *Nature Communications* 6, no. 1 (). doi:10.1038/ncomms8313.
- Kimura, Hiroshi, and Peter R. Cook. 2001. “Kinetics of Core Histones in Living Human Cells”. *Journal of Cell Biology* 153, no. 7 (): 1341–1354. doi:10.1083/jcb.153.7.1341.
- Kinoshita, Masaki, Michael Barber, William Mansfield, Yingzhi Cui, Daniel Spindlow, Giuliano Giuseppe Stirparo, Sabine Dietmann, Jennifer Nichols, and Austin Smith. 2021. “Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency”. *Cell Stem Cell* 28, no. 3 (): 453–471.e8. doi:10.1016/j.stem.2020.11.005.
- Kojima, Yoji, Keren Kaufman-Francis, Joshua B. Studdert, Kirsten A. Steiner, Melinda D. Power, David A.F. Loebel, Vanessa Jones, Angelyn Hor, Gustavo de Alencastro, Grant J. Logan, Erdahl T. Teber, Oliver H. Tam, Michael D. Stutz, Ian E. Alexander, Hilda A. Pickett, and Patrick P.L. Tam. 2014. “The Transcriptional and Functional Properties of Mouse Epiblast Stem Cells Resemble the Anterior Primitive Streak”. *Cell Stem Cell* 14, no. 1 (): 107–120. doi:10.1016/j.stem.2013.09.014.
- Kumar, Amarjeet, and Hidetoshi Kono. 2020. “Heterochromatin protein 1 (HP1): interactions with itself and chromatin components”. *Biophysical Reviews* 12, no. 2 (): 387–400. doi:10.1007/s12551-020-00663-y.
- Kumar, Roshan M., Patrick Cahan, Alex K. Shalek, Rahul Satija, A. Jay Daley, Keyser, Hu Li, Jin Zhang, Keith Pardee, David Gennert, John J. Trombetta, Thomas C. Ferrante, Aviv Regev, George Q. Daley, and James J. Collins. 2014. “Deconstructing transcriptional heterogeneity in pluripotent stem cells”. *Nature* 516, no. 7529 (): 56–61. doi:10.1038/nature13920.

- Lachner, Monika, Dónal O’Carroll, Stephen Rea, Karl Mechtler, and Thomas Jenuwein. 2001. “Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins”. *Nature* 410, no. 6824 (): 116–120. doi:10.1038/35065132.
- Lando, David, Srinjan Basu, Tim J Stevens, Andy Riddell, Kai J Wohlfahrt, Yang Cao, Wayne Boucher, Martin Leeb, Liam P Atkinson, Steven F Lee, Brian Hendrich, Dave Klenerman, and Ernest D Laue. 2018. “Combining fluorescence imaging with Hi-C to study 3D genome architecture of the same single cell”. *Nature Protocols* 13, no. 5 (): 1034–1061. doi:10.1038/nprot.2018.017.
- Larson, Adam G., Daniel Elnatan, Madeline M. Keenen, Michael J. Trnka, Jonathan B. Johnston, Alma L. Burlingame, David A. Agard, Sy Redding, and Geeta J. Narlikar. 2017. “Liquid droplet formation by HP1 $\alpha$  suggests a role for phase separation in heterochromatin”. *Nature* 547, no. 7662 (): 236–240. doi:10.1038/nature22822.
- Larson, Adam G., and Geeta J. Narlikar. 2018. “The Role of Phase Separation in Heterochromatin Formation, Function, and Regulation”. *Biochemistry* 57, no. 17 (): 2540–2548. doi:10.1021/acs.biochem.8b00401.
- Lavigne, Marc, Ragnhild Eskeland, Saliha Azebi, Violaine Saint-André, Suk Min Jang, Eric Batsché, Hua-Ying Fan, Robert E. Kingston, Axel Imhof, and Christian Muchardt. 2009. “Interaction of HP1 and Brg1/Brm with the Globular Domain of Histone H3 Is Required for HP1-Mediated Repression”. Ed. by Asifa Akhtar. *PLoS Genetics* 5, no. 12 (): e1000769. doi:10.1371/journal.pgen.1000769.
- Lechner, Mark S., Gillian E. Begg, David W. Speicher, and Frank J. Rauscher. 2000. “Molecular Determinants for Targeting Heterochromatin Protein 1-Mediated Gene Silencing: Direct Chromoshadow Domain–KAP-1 Corepressor Interaction Is Essential”. *Molecular and Cellular Biology* 20, no. 17 (): 6449–6465. doi:10.1128/mcb.20.17.6449–6465.2000.
- LeRoy, Gary, John T. Weston, Barry M. Zee, Nicolas L. Young, Mariana D. Plazas-Mayorca, and Benjamin A. Garcia. 2009. “Heterochromatin Protein 1 Is Extensively Decorated with Histone Code-like Post-translational Modifications”. *Molecular & Cellular Proteomics* 8 (11): 2432–2442. doi:10.1074/mcp.M900160-MCP200.
- Lever, Melody A., John P. H. Th’ng, Xuejun Sun, and Michael J. Hendzel. 2000. “Rapid exchange of histone H1.1 on chromatin in living human cells”. *Nature* 408, no. 6814 (): 873–876. doi:10.1038/35048603.
- Lew, Matthew D., Alexander R. S. von Diezmann, and W. E. Moerner. 2013. “Easy-DHPSF open-source software for three-dimensional localization of single molecules with precision beyond the optical diffraction limit”. *Protocol Exchange* (). doi:10.1038/protex.2013.026.
- Li, Meng Amy, Paulo P Amaral, Priscilla Cheung, Jan H Bergmann, Masaki Kinoshita, Tüzer Kalkan, Meryem Ralser, Sam Robson, Ferdinand von Meyenn, Maike Paramor, Fengtang Yang, Caifu Chen, Jennifer Nichols, David L Spector, Tony Kouzarides, Lin He, and Austin Smith. 2017. “A lncRNA fine tunes the dynamics of a cell state transition involving Lin28, let-7 and de novo DNA methylation”. *eLife* 6 (). doi:10.7554/eLife.23468.
- Li, Yiming, Markus Mund, Philipp Hoess, Joran Deschamps, Ulf Matti, Bianca Nijmeijer, Vilma Jimenez Sabinina, Jan Ellenberg, Ingmar Schoen, and Jonas Ries. 2018. “Real-time 3D single-molecule localization using experimental point spread functions”. *Nature Methods* 15, no. 5 (): 367–369. doi:10.1038/nmeth.4661.
- Lippincott-Schwartz, Jennifer, Erik Snapp, and Anne Kenworthy. 2001. “Studying protein dynamics in living cells”. *Nature Reviews Molecular Cell Biology* 2, no. 6 (): 444–456. doi:10.1038/35073068.
- Liu, Huadong, Marek Galka, Aimee Iberg, Zezhou Wang, Lei Li, Courtney Voss, Xinfeng Jiang, Gilles Lajoie, Zhiping Huang, Mark T. Bedford, and Shawn S. C. Li. 2010. “Systematic Identification of Methyllysine-Driven Interactions for Histone and Nonhistone Targets”. *Journal of Proteome Research* 9, no. 11 (): 5827–5836. doi:10.1021/pr100597b.

- Liu, Huadong, Marek Galka, Eiichiro Mori, Xuguang Liu, Yu-fen Lin, Ran Wei, Paula Pittock, Courtney Voss, Gurpreet Dhama, Xing Li, Masaaki Miyaji, Gilles Lajoie, Benjamin Chen, and Shawn Shun-Cheng Li. 2013. “A Method for Systematic Mapping of Protein Lysine Methylation Identifies Functions for HP1 $\text{extbeta}$  in DNA Damage Response”. *Molecular Cell* 50, no. 5 (): 723–735. doi:10.1016/j.molcel.2013.04.025.
- Liu, Yanli, Su Qin, Ming Lei, Wolfram Tempel, Yuzhe Zhang, Peter Loppnau, Yanjun Li, and Jinrong Min. 2017. “Peptide recognition by heterochromatin protein 1 (HP1) chromoshadow domains revisited: Plasticity in the pseudosymmetric histone binding site of human HP1”. *Journal of Biological Chemistry* 292, no. 14 (): 5655–5664. doi:10.1074/jbc.m116.768374.
- Liu, Zhe, Wesley R Legant, Bi-Chang Chen, Li Li, Jonathan B Grimm, Luke D Lavis, Eric Betzig, and Robert Tjian. 2014. “3D imaging of Sox2 enhancer clusters in embryonic stem cells”. *eLife* 3 (). doi:10.7554/eLife.04236.
- Lomberk, Gwen, Debora Bensi, Martín E. Fernandez-Zapico, and Raul Urrutia. 2006. “Evidence for the existence of an HP1-mediated subcode within the histone code”. *Nature Cell Biology* 8, no. 4 (): 407–415. doi:10.1038/ncb1383.
- Maeshima, Kazuhiro, Ryan Rogge, Sachiko Tamura, Yasumasa Joti, Takaaki Hikima, Heather Szerlong, Christine Krause, Jake Herman, Erik Seidel, Jennifer DeLuca, Tetsuya Ishikawa, and Jeffrey C Hansen. 2016. “Nucleosomal arrays self-assemble into supramolecular globular structures lacking 30-nm fibers”. *The EMBO Journal* 35, no. 10 (): 1115–1132. doi:10.15252/embj.201592660.
- Maison, Christèle, Delphine Bailly, Antoine H.F.M. Peters, Jean-Pierre Quivy, Danièle Roche, Angela Taddei, Monika Lachner, Thomas Jenuwein, and Geneviève Almouzni. 2002. “Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component”. *Nature Genetics* 30, no. 3 (): 329–334. doi:10.1038/ng843.
- Manley, Suliana, Jennifer M. Gillette, and Jennifer Lippincott-Schwartz. 2010. “Single-Particle Tracking Photoactivated Localization Microscopy for Mapping Single-Molecule Dynamics”. In *Methods in Enzymology*, 109–120. Elsevier. doi:10.1016/s0076-6879(10)75005-9.
- Martin, G. R. 1981. “Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.” *Proceedings of the National Academy of Sciences* 78, no. 12 (): 7634–7638. doi:10.1073/pnas.78.12.7634.
- Mateescu, Bogdan, Brigitte Bourachot, Christophe Rachez, Vasily Ogryzko, and Christian Muchardt. 2008. “Regulation of an inducible promoter by an HP1 $\text{extbeta}$ –HP1 $\text{extgamma}$  switch”. *EMBO reports* 9, no. 3 (): 267–272. doi:10.1038/embor.2008.1.
- Mateo, Leslie J., Sedona E. Murphy, Antonina Hafner, Isaac S. Cinquini, Carly A. Walker, and Alistair N. Boettiger. 2019. “Visualizing DNA folding and RNA in embryos at single-cell resolution”. *Nature* 568, no. 7750 (): 49–54. doi:10.1038/s41586-019-1035-4.
- Mateos-Langerak, Julio, Maartje C. Brink, Martijn S. Luijsterburg, Ineke van der Kraan, Roel van Driel, and Pernette J. Verschure. 2007. “Pericentromeric Heterochromatin Domains Are Maintained without Accumulation of HP1”. Ed. by Wendy Bickmore. *Molecular Biology of the Cell* 18, no. 4 (): 1464–1471. doi:10.1091/mbc.e06-01-0025.
- Mattout, Anna, Yair Aaronson, Badi Sri Sailaja, Edupuganti V. Raghuram, Arigela Harikumar, Jan Philipp Mallm, Kae Hwan Sim, Malka Nissim-Rafinia, Emmanuelle Supper, Prim B. Singh, Siu Kwan Sze, Susan M. Gasser, Karsten Rippe, and Eran Meshorer. 2015. “Heterochromatin Protein 1 $\beta$  (HP1 $\beta$ ) has distinct functions and distinct nuclear distribution in pluripotent versus differentiated cells”. *Genome Biology* 16 (1): 7–9. doi:10.1186/s13059-015-0760-8.
- McSwiggen, David T., Mustafa Mir, Xavier Darzacq, and Robert Tjian. 2019a. “Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences”. *Genes & Development* 33, numbers 23-24 (): 1619–1634. doi:10.1101/gad.331520.119.

- McSwiggen, David Trombley, Anders S Hansen, Sheila S Teves, Hervé Marie-Nelly, Yvonne Hao, Alec Basil Heckert, Kayla K Umemoto, Claire Dugast-Darzacq, Robert Tjian, and Xavier Darzacq. 2019b. “Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation”. 8 (). doi:10.7554/elife.47098.
- Mendez, Deanna L., Daesung Kim, Maksymilian Chruszcz, Gena E. Stephens, Wladek Minor, Sepideh Khorasanizadeh, and Sarah C. R. Elgin. 2011. “The HP1a Disordered C Terminus and Chromo Shadow Domain Cooperate to Select Target Peptide Partners”. *ChemBioChem* 12, no. 7 (): 1084–1096. doi:10.1002/cbic.201000598.
- Mendez, Deanna L., Rebecca E. Mandt, and Sarah C.R. Elgin. 2013. “Heterochromatin Protein 1a (HP1a) Partner Specificity Is Determined by Critical Amino Acids in the Chromo Shadow Domain and C-terminal Extension”. *Journal of Biological Chemistry* 288, no. 31 (): 22315–22323. doi:10.1074/jbc.m113.468413.
- Mennella, V., B. Keszthelyi, K. L. McDonald, B. Chhun, F. Kan, G. C. Rogers, B. Huang, and D. A. Agard. 2012. “Subdiffraction-resolution fluorescence microscopy reveals a domain of the centrosome critical for pericentriolar material organization”. *Nature Cell Biology* 14, no. 11 (): 1159–1168. doi:10.1038/ncb2597.
- Meshorer, Eran, Dhananjay Yellajoshula, Eric George, Peter J. Scambler, David T. Brown, and Tom Misteli. 2006. “Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells”. *Developmental Cell* 10, no. 1 (): 105–116. doi:10.1016/j.devcel.2005.10.017.
- Minc, Elsa, Yves Allory, Howard J. Worman, Jean Claude Courvalin, and Brigitte Buendia. 1999. “Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells”. *Chromosoma* 108 (4): 220–234. doi:10.1007/s004120050372.
- Miron, Ezequiel, Roel Oldenkamp, Jill M. Brown, David M. S. Pinto, C. Shan Xu, Ana R. Faria, Haitham A. Shaban, James D. P. Rhodes, Cassandravictoria Innocent, Sara de Ornellas, Harald F. Hess, Veronica Buckle, and Lothar Schermelleh. 2020. “Chromatin arranges in chains of mesoscale domains with nanoscale functional topography independent of cohesin”. *Science Advances* 6, no. 39 (). doi:10.1126/sciadv.aba8811.
- Mishra, Laxmi Narayan, and Jeffrey J. Hayes. 2018. “A nucleosome-free region locally abrogates histone H1-dependent restriction of linker DNA accessibility in chromatin”. *Journal of Biological Chemistry* 293, no. 50 (): 19191–19200. doi:10.1074/jbc.ra118.005721.
- Miura, Takumi, Mark P. Mattson, and Mahendra S. Rao. 2004. “Cellular lifespan and senescence signaling in embryonic stem cells”. *Aging Cell* 3, no. 6 (): 333–343. doi:10.1111/j.1474-9728.2004.00134.x.
- Muchardt, Christian, Marie Guillemé, Jacob-S Seeler, Didier Trouche, Anne Dejean, and Moshe Yaniv. 2002. “Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1 $\alpha$ ”. *EMBO reports* 3, no. 10 (): 975–981. doi:10.1093/embo-reports/kvf194.
- Müller-Ott, Katharina, Fabian Erdel, Anna Matveeva, Jan-Philipp Mallm, Anne Rademacher, Matthias Hahn, Caroline Bauer, Qin Zhang, Sabine Kaltofen, Gunnar Schotta, Thomas Höfer, and Karsten Rippe. 2014. “Specificity, propagation, and memory of pericentric heterochromatin”. *Molecular Systems Biology* 10, no. 8 (): 746. doi:10.15252/msb.20145377.
- Müller, Katharina P., Fabian Erdel, Maiwen Caudron-Herger, Caroline Marth, Barna D. Fodor, Mario Richter, Manuela Scaranaro, Joël Beaudouin, Malte Wachsmuth, and Karsten Rippe. 2009. “Multiscale Analysis of Dynamics and Interactions of Heterochromatin Protein 1 by Fluorescence Fluctuation Microscopy”. *Biophysical Journal* 97, no. 11 (): 2876–2885. doi:10.1016/j.bpj.2009.08.057.
- Munari, Francesca, Nasrollah Rezaei-Ghaleh, Shengqi Xiang, Wolfgang Fischle, and Markus Zweckstetter. 2013. “Structural Plasticity in Human Heterochromatin Protein 1 $\beta$ ”. Ed. by Roman Tuma. *PLoS ONE* 8, no. 4 (): e60887. doi:10.1371/journal.pone.0060887.

- Munari, Francesca, Szabolcs Soeroes, Hans Michael Zenn, Adrian Schomburg, Nils Kost, Sabrina Schröder, Rebecca Klingberg, Nasrollah Rezaei-Ghaleh, Alexandra Stützer, Kathy Ann Gelato, Peter Jomo Walla, Stefan Becker, Dirk Schwarzer, Bastian Zimmermann, Wolfgang Fischle, and Markus Zweckstetter. 2012. “Methylation of Lysine 9 in Histone H3 Directs Alternative Modes of Highly Dynamic Interaction of Heterochromatin Protein hHP1 $\text{extbeta}$  with the Nucleosome”. *Journal of Biological Chemistry* 287, no. 40 (): 33756–33765. doi:10.1074/jbc.m112.390849.
- Nagano, Takashi, Yaniv Lubling, Tim J. Stevens, Stefan Schoenfelder, Eitan Yaffe, Wendy Dean, Ernest D. Laue, Amos Tanay, and Peter Fraser. 2013. “Single-cell Hi-C reveals cell-to-cell variability in chromosome structure”. *Nature* 502, no. 7469 (): 59–64. doi:10.1038/nature12593.
- Nicetto, Dario, and Kenneth S. Zaret. 2019. “Role of H3K9me3 heterochromatin in cell identity establishment and maintenance”. *Current Opinion in Genetics & Development* 55 (): 1–10. doi:10.1016/j.gde.2019.04.013.
- Nielsen, Anders Lade, Mustapha Oulad-Abdelghani, José A Ortiz, Eumorphia Remboutsika, Pierre Chambon, and Régine Losson. 2001. “Heterochromatin Formation in Mammalian Cells”. *Molecular Cell* 7, no. 4 (): 729–739. doi:10.1016/s1097-2765(01)00218-0.
- Nielsen, Peter R., Daniel Nietlispach, Helen R. Mott, Juliana Callaghan, Andrew Bannister, Tony Kouzarides, Alexey G. Murzin, Natalia V. Murzina, and Ernest D. Laue. 2002. “Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9”. *Nature* 416, no. 6876 (): 103–107. doi:10.1038/nature722.
- Nieuwenhuizen, Robert P J, Keith A Lidke, Mark Bates, Daniela Leyton Puig, David Grünwald, Sjoerd Stallinga, and Bernd Rieger. 2013. “Measuring image resolution in optical nanoscopy”. *Nature Methods* 10, no. 6 (): 557–562. doi:10.1038/nmeth.2448.
- Nishibuchi, Gohei, Shinichi Machida, Akihisa Osakabe, Hiromu Murakoshi, Kyoko Hiragami-Hamada, Reiko Nakagawa, Wolfgang Fischle, Yoshifumi Nishimura, Hitoshi Kurumizaka, Hideaki Tagami, and Jun-ichi Nakayama. 2014. “N-terminal phosphorylation of HP1 $\text{extalpha}$  increases its nucleosome-binding specificity”. *Nucleic Acids Research* 42, no. 20 (): 12498–12511. doi:10.1093/nar/gku995.
- Noa, Amra, Hui-Shun Kuan, Vera Aschmann, Vasily Ziburdaev, and Lennart Hilbert. 2021. “The hierarchical packing of euchromatin domains can be described as multiplicative cascades”. Ed. by Ilya Ioshikhes. *PLOS Computational Biology* 17, no. 5 (): e1008974. doi:10.1371/journal.pcbi.1008974.
- Novo, Clara Lopes, Calvin Tang, Kashif Ahmed, Ugljesa Djuric, Eden Fussner, Nicholas P. Mullin, Natasha P. Morgan, Jasvinder Hayre, Arnold R. Sienerth, Sarah Elderkin, Ryuichi Nishinakamura, Ian Chambers, James Ellis, David P. Bazett-Jones, and Peter J. Rugg-Gunn. 2016. “The pluripotency factor Nanog regulates pericentromeric heterochromatin organization in mouse embryonic stem cells”. *Genes & Development* 30, no. 9 (): 1101–1115. doi:10.1101/gad.275685.115.
- Novo, Clara Lopes, Emily Wong, Colin Hockings, Chetan Poudel, Eleanor Sheekey, Simon Walker, Gabriele S. Kaminski Schierle, Geeta J. Narlikar, and Peter J. Rugg-Gunn. 2020. “Satellite repeat transcripts modulate heterochromatin condensates and safeguard chromosome stability in mouse embryonic stem cells” (). doi:10.1101/2020.06.08.139642.
- Ostapcuk, Veronika, Fabio Mohn, Sarah H. Carl, Anja Basters, Daniel Hess, Vytautas Iesmantavicius, Lisa Lampersberger, Matyas Flemr, Aparna Pandey, Nicolas H. Thomä, Joerg Betschinger, and Marc Bühler. 2018. “Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes”. 557, no. 7707 (): 739–743. doi:10.1038/s41586-018-0153-8.
- Ou, Horng D., Sébastien Phan, Thomas J. Deerinck, Andrea Thor, Mark H. Ellisman, and Clodagh C. O’Shea. 2017. “ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells”. *Science* 357, no. 6349 (). doi:10.1126/science.aag0025.

- Ozdil, Berrin, Günnur Güler, Eda Acikgoz, Duygu Calik Kocaturk, and Huseyin Aktug. 2019. “The effect of extracellular matrix on the differentiation of mouse embryonic stem cells”. *Journal of Cellular Biochemistry* 121, no. 1 (): 269–283. doi:10.1002/jcb.29159.
- Pagliari, Stefano, Kristian Franze, Crystal R. McClain, George W. Wylde, Cynthia L. Fisher, Robin J. M. Franklin, Alexandre J. Kabla, Ulrich F. Keyser, and Kevin J. Chalut. 2014. “Auxetic nuclei in embryonic stem cells exiting pluripotency”. *Nature Materials* 13, no. 6 (): 638–644. doi:10.1038/nmat3943.
- Pajerowski, J. D., K. N. Dahl, F. L. Zhong, P. J. Sammak, and D. E. Discher. 2007. “Physical plasticity of the nucleus in stem cell differentiation”. *Proceedings of the National Academy of Sciences* 104, no. 40 (): 15619–15624. doi:10.1073/pnas.0702576104.
- Pau, G., F. Fuchs, O. Sklyar, M. Boutros, and W. Huber. 2010. “EBImage—an R package for image processing with applications to cellular phenotypes”. *Bioinformatics* 26, no. 7 (): 979–981. doi:10.1093/bioinformatics/btq046.
- Pavani, S. R. P., M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun, and W. E. Moerner. 2009. “Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function”. *Proceedings of the National Academy of Sciences* 106, no. 9 (): 2995–2999. doi:10.1073/pnas.0900245106.
- Pavani, Sri Rama Prasanna, and Rafael Piestun. 2008. “Three dimensional tracking of fluorescent microparticles using a photon-limited double-helix response system”. *Optics Express* 16, no. 26 (): 22048. doi:10.1364/oe.16.022048.
- Peters, Antoine H.F.M., Dónal O’Carroll, Harry Scherthan, Karl Mechtler, Stephan Sauer, Christian Schöfer, Klara Weipoltshammer, Michaela Pagani, Monika Lachner, Alexander Kohlmaier, Susanne Opravil, Michael Doyle, Maria Sibilina, and Thomas Jenuwein. 2001. “Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability”. *Cell* 107, no. 3 (): 323–337. doi:10.1016/s0092-8674(01)00542-6.
- Phair, Robert D., and Tom Misteli. 2000. “High mobility of proteins in the mammalian cell nucleus”. *Nature* 404, no. 6778 (): 604–609. doi:10.1038/35007077.
- Piacentini, Lucia, Laura Fantì, Maria Berloco, Barbara Perrini, and Sergio Pimpinelli. 2003. “Heterochromatin protein 1 (HP1) is associated with induced gene expression in Drosophila euchromatin”. *Journal of Cell Biology* 161, no. 4 (): 707–714. doi:10.1083/jcb.200303012.
- Pillariseti, Anand, Jaydev P. Desai, Hamid Ladjal, Andrew Schifmacher, Antoine Ferreira, and Carol L. Keefer. 2011. “Mechanical Phenotyping of Mouse Embryonic Stem Cells: Increase in Stiffness with Differentiation”. *Cellular Reprogramming* 13, no. 4 (): 371–380. doi:10.1089/cell.2011.0028.
- Pokorná, Pavlína, Miroslav Krepl, and Jiří Šponer. 2021. “Residues flanking the ARKme3T/S motif allow binding of diverse targets to the HP1 chromodomain: Insights from molecular dynamics simulations”. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1865, no. 1 (): 129771. doi:10.1016/j.bbagen.2020.129771.
- Pombo, Ana, and Niall Dillon. 2015. “Three-dimensional genome architecture: players and mechanisms”. *Nature Reviews Molecular Cell Biology* 16, no. 4 (): 245–257. doi:10.1038/nrm3965.
- Quinodoz, Sofia A., Joanna W. Jachowicz, Prashant Bhat, Noah Ollikainen, Abhik K. Banerjee, Isabel N. Goronzy, Mario R. Blanco, Peter Chovanec, Amy Chow, Yolanda Markaki, Jasmine Thai, Kathrin Plath, and Mitchell Guttman. 2021. “RNA promotes the formation of spatial compartments in the nucleus”. *Cell* 184, no. 23 (): 5775–5790.e30. doi:10.1016/j.cell.2021.10.014.
- R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ricci, Maria Aurelia, Carlo Manzo, María Filomena García-Parajo, Melike Lakadamyali, and Maria Pia Cosma. 2015. “Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes In Vivo”. *Cell* 160, no. 6 (): 1145–1158. doi:10.1016/j.cell.2015.01.054.

- Richart, Alexandria N., Clair I.W. Brunner, Katherine Stott, Natalia V. Murzina, and Jean O. Thomas. 2012. “Characterization of Chromoshadow Domain-mediated Binding of Heterochromatin Protein 1 $\alpha$  (HP1 $\alpha$ ) to Histone H3”. *Journal of Biological Chemistry* 287, no. 22 (): 18730–18737. doi:10.1074/jbc.m111.337204.
- Ries, Jonas. 2020. “SMAP: a modular super-resolution microscopy analysis platform for SMLM data”. 17, no. 9 (): 870–872. doi:10.1038/s41592-020-0938-1.
- Ripley, B. D. 1977. “Modelling Spatial Patterns”. *Journal of the Royal Statistical Society: Series B (Methodological)* 39, no. 2 (): 172–192. doi:10.1111/j.2517-6161.1977.tb01615.x.
- Rivera-Mulia, Juan Carlos, Andrew Dimond, Daniel Vera, Claudia Trevilla-Garcia, Takayo Sasaki, Jared Zimmerman, Catherine Dupont, Joost Gribnau, Peter Fraser, and David M. Gilbert. 2018. “Allele-specific control of replication timing and genome organization during development”. *Genome Research* 28, no. 6 (): 800–811. doi:10.1101/gr.232561.117.
- Rose, Nathan R., and Robert J. Klose. 2014. “Understanding the relationship between DNA methylation and histone lysine methylation”. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1839, no. 12 (): 1362–1372. doi:10.1016/j.bbagr.2014.02.007.
- Ruan, Jianbin, Hui Ouyang, Maria F. Amaya, Mani Ravichandran, Peter Loppnau, Jinrong Min, and Jianye Zang. 2012. “Structural Basis of the Chromodomain of Cbx3 Bound to Methylated Peptides from Histone H1 and G9a”. Ed. by Pierre-Antoine Defossez. *PLoS ONE* 7, no. 4 (): e35376. doi:10.1371/journal.pone.0035376.
- Ryan, Daniel P, and David J Tremethick. 2018. “The interplay between H2A.Z and H3K9 methylation in regulating HP1 $\alpha$  binding to linker histone-containing chromatin”. *Nucleic Acids Research* 46, no. 18 (): 9353–9366. doi:10.1093/nar/ky632.
- Sage, Daniel, Thanh-An Pham, Hazen Babcock, Tomas Lukes, Thomas Pengo, Jerry Chao, Ramraj Velmurugan, Alex Herbert, Anurag Agrawal, Silvia Colabrese, Ann Wheeler, Anna Archetti, Bernd Rieger, Raimund Ober, Guy M. Hagen, Jean-Baptiste Sibarita, Jonas Ries, Ricardo Henriques, Michael Unser, and Seamus Holden. 2019. “Super-resolution fight club: assessment of 2D and 3D single-molecule localization microscopy software”. *Nature Methods* 16, no. 5 (): 387–395. doi:10.1038/s41592-019-0364-4.
- Sanulli, S., M. J. Trnka, V. Dharmarajan, R. W. Tibble, B. D. Pascal, A. L. Burlingame, P. R. Griffin, J. D. Gross, and G. J. Narlikar. 2019. “HP1 reshapes nucleosome core to promote phase separation of heterochromatin”. *Nature* 575, no. 7782 (): 390–394. doi:10.1038/s41586-019-1669-2.
- Schindelin, Johannes, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona. 2012. “Fiji: an open-source platform for biological-image analysis”. *Nature Methods* 9, no. 7 (): 676–682. doi:10.1038/nmeth.2019.
- Schmidtman, Elisabeth, Tobias Anton, Pascaline Rombaut, Franz Herzog, and Heinrich Leonhardt. 2016. “Determination of local chromatin composition by CasID”. *Nucleus* 7, no. 5 (): 476–484. doi:10.1080/19491034.2016.1239000.
- Schmiedeberg, Lars, Klaus Weisshart, Stephan Diekmann, Gabriele Meyer zu Hoerste, and Peter Hemmerich. 2004. “High- and Low-mobility Populations of HP1 in Heterochromatin of Mammalian Cells”. *Molecular Biology of the Cell* 15, no. 6 (): 2819–2833. doi:10.1091/mbc.e03-11-0827.
- Schotta, G. 2002. “Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing”. *The EMBO Journal* 21, no. 5 (): 1121–1131. doi:10.1093/emboj/21.5.1121.
- Schultz, David C., Kasirajan Ayyanathan, Dmitri Negorev, Gerd G. Maul, and Frank J. Rauscher. 2002. “SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins”. *Genes & Development* 16, no. 8 (): 919–932. doi:10.1101/gad.973302.

- Schwarz, Gideon. 1978. “Estimating the Dimension of a Model”. *The Annals of Statistics* 6, no. 2 (). doi:10.1214/aos/1176344136.
- Shaban, Haitham A., Roman Barth, Ludmila Recoules, and Kerstin Bystricky. 2020. “Hi-D: nanoscale mapping of nuclear dynamics in single living cells”. *Genome Biology* 21, no. 1 (). doi:10.1186/s13059-020-02002-6.
- Shanle, Erin K., Stephen A. Shinsky, Joseph B. Bridgers, Narkhyun Bae, Cari Sagum, Krzysztof Krajewski, Scott B. Rothbart, Mark T. Bedford, and Brian D. Strahl. 2017. “Histone peptide microarray screen of chromo and Tudor domains defines new histone lysine methylation interactions”. *Epigenetics & Chromatin* 10, no. 1 (). doi:10.1186/s13072-017-0117-5.
- Shechtman, Yoav, Lucien E. Weiss, Adam S. Backer, Steffen J. Sahl, and W. E. Moerner. 2015. “Precise Three-Dimensional Scan-Free Multiple-Particle Tracking over Large Axial Ranges with Tetrapod Point Spread Functions”. *Nano Letters* 15, no. 6 (): 4194–4199. doi:10.1021/acs.nanolett.5b01396.
- Shivanandan, A., H. Deschout, M. Scarselli, and A. Radenovic. 2014. “Challenges in quantitative single molecule localization microscopy”. *FEBS Letters* 588, no. 19 (): 3595–3602. doi:10.1016/j.febslet.2014.06.014.
- Shukron, O., A. Seeber, A. Amitai, and D. Holcman. 2019. “Advances Using Single-Particle Trajectories to Reconstruct Chromatin Organization and Dynamics”. 35, no. 9 (): 685–705. doi:10.1016/j.tig.2019.06.007.
- Smith, Austin. 2017. “Formative pluripotency: the executive phase in a developmental continuum”. *Development* 144, no. 3 (): 365–373. doi:10.1242/dev.142679.
- Smith, Austin G., John K. Heath, Deborah D. Donaldson, Gordon G. Wong, J. Moreau, Mark Stahl, and David Rogers. 1988. “Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides”. *Nature* 336, no. 6200 (): 688–690. doi:10.1038/336688a0.
- Spahn, Christoph, Jonathan B. Grimm, Luke D. Lavis, Marko Lampe, and Mike Heilemann. 2018. “Whole-Cell, 3D, and Multicolor STED Imaging with Exchangeable Fluorophores”. *Nano Letters* 19, no. 1 (): 500–505. doi:10.1021/acs.nanolett.8b04385.
- Stevens, Tim J., et al. 2017. “3D structures of individual mammalian genomes studied by single-cell Hi-C”. *Nature* 544, no. 7648 (): 59–64. doi:10.1038/nature21429.
- Strickfaden, Hilmar, Thomas O. Tolsma, Ajit Sharma, D. Alan Underhill, Jeffrey C. Hansen, and Michael J. Hendzel. 2020. “Condensed Chromatin Behaves like a Solid on the Mesoscale In Vitro and in Living Cells”. *Cell* 183, no. 7 (): 1772–1784.e13. doi:10.1016/j.cell.2020.11.027.
- Strogantsev, Ruslan, Felix Krueger, Kazuki Yamazawa, Hui Shi, Poppy Gould, Megan Goldman-Roberts, Kirsten McEwen, Bowen Sun, Roger Pedersen, and Anne C. Ferguson-Smith. 2015. “Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression”. *Genome Biology* 16, no. 1 (). doi:10.1186/s13059-015-0672-7.
- Strom, Amy R., Alexander V. Emelyanov, Mustafa Mir, Dmitry V. Fyodorov, Xavier Darzacq, and Gary H. Karpen. 2017. “Phase separation drives heterochromatin domain formation”. *Nature* 547, no. 7662 (): 241–245. doi:10.1038/nature22989.
- Strom, Amy R., Ronald J Biggs, Edward J Banigan, Xiaotao Wang, Katherine Chiu, Cameron Herman, Jimena Collado, Feng Yue, Joan C Ritland Politz, Leah J Tait, David Scalzo, Agnes Telling, Mark Groudine, Clifford P Brangwynne, John F Marko, and Andrew D Stephens. 2021. “HP1 $\alpha$  is a chromatin crosslinker that controls nuclear and mitotic chromosome mechanics”. *eLife* 10 (). doi:10.7554/eLife.63972.
- Sugimoto, K., T. Yamada, Y. Muro, and M. Himeno. 1996. “Human Homolog of Drosophila Heterochromatin-Associated Protein 1 (HP1) Is a DNA-Binding Protein Which Possesses a DNA-Binding Motif with Weak Similarity to That of Human Centromere Protein C (CENP-C)”. *Journal of Biochemistry* 120, no. 1 (): 153–159. doi:10.1093/oxfordjournals.jbchem.a021378.

- Tesar, Paul J., Josh G. Chenoweth, Frances A. Brook, Timothy J. Davies, Edward P. Evans, David L. Mack, Richard L. Gardner, and Ronald D. G. McKay. 2007. “New cell lines from mouse epiblast share defining features with human embryonic stem cells”. *Nature* 448, no. 7150 (): 196–199. doi:10.1038/nature05972.
- Thevathasan, Jervis Vermal, Maurice Kahnwald, Konstanty Ciesliński, Philipp Hoess, Sudheer Kumar Peneti, Manuel Reitberger, Daniel Heid, Krishna Chaitanya Kasuba, Sarah Janice Hoerner, Yiming Li, Yu-Le Wu, Markus Mund, Ulf Matti, Pedro Matos Pereira, Ricardo Henriques, Bianca Nijmeijer, Moritz Kueblbeck, Vilma Jimenez Sabinina, Jan Ellenberg, and Jonas Ries. 2019. “Nuclear pores as versatile reference standards for quantitative superresolution microscopy”. 16, no. 10 (): 1045–1053. doi:10.1038/s41592-019-0574-9.
- Thiru, Abarna, Daniel Nietlispach, Helen R Mott, Mitsuru Okuwaki, Debbie Lyon, Peter R Nielsen, Miriam Hirshberg, Alain Verreault, Natalia V Murzina, and Ernest D Laue. 2004. “Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin”. *EMBO Journal* 23 (3): 489–499. doi:10.1038/sj.emboj.7600088.
- Tosolini, Matteo, Vincent Brochard, Pierre Adenot, Martine Chebrou, Giacomo Grillo, Violette Navia, Nathalie Beaujean, Claire Francastel, Amélie Bonnet-Garnier, and Alice Jouneau. 2018. “Contrasting epigenetic states of heterochromatin in the different types of mouse pluripotent stem cells”. *Scientific Reports* 8, no. 1 (). doi:10.1038/s41598-018-23822-4.
- Tse, Christin, Takashi Sera, Alan P. Wolffe, and Jeffrey C. Hansen. 1998. “Disruption of Higher-Order Folding by Core Histone Acetylation Dramatically Enhances Transcription of Nucleosomal Arrays by RNA Polymerase III”. *Molecular and Cellular Biology* 18, no. 8 (): 4629–4638. doi:10.1128/mcb.18.8.4629.
- Turner, Abigail L., Matthew Watson, Oscar G. Wilkins, Laura Cato, Andrew Travers, Jean O. Thomas, and Katherine Stott. 2018. “Highly disordered histone H1-DNA model complexes and their condensates”. *Proceedings of the National Academy of Sciences* 115, no. 47 (): 11964–11969. doi:10.1073/pnas.1805943115.
- Vakoc, Christopher R., Sean A. Mandat, Benjamin A. Olenchock, and Gerd A. Blobel. 2005. “Histone H3 Lysine 9 Methylation and HP1 $\gamma$  Are Associated with Transcription Elongation through Mammalian Chromatin”. *Molecular Cell* 19, no. 3 (): 381–391. doi:10.1016/j.molcel.2005.06.011.
- Van, Phu, Wenxin Jiang, Raphael Gottardo, and Greg Finak. 2018. “ggCyto: next generation open-source visualization software for cytometry”. Ed. by Jonathan Wren. *Bioinformatics* 34, no. 22 (): 3951–3953. doi:10.1093/bioinformatics/bty441.
- Veneri, Paula, Camila Vazquez Echeagaray, Camila Osés, Martin Stortz, Alejandra Guberman, and Valeria Levi. 2020. “Dynamical reorganization of the pluripotency transcription factors Oct4 and Sox2 during early differentiation of embryonic stem cells”. *Scientific Reports* 10, no. 1 (). doi:10.1038/s41598-020-62235-0.
- Verschure, Pernelle J, Ineke van der Kraan, Erik M M Manders, Deborah Hoogstraten, Adriaan B Houtsmuller, and Roel van Driel. 2003. “Condensed chromatin domains in the mammalian nucleus are accessible to large macromolecules”. *EMBO reports* 4, no. 9 (): 861–866. doi:10.1038/sj.embor.embor922.
- Wang, Liang, Yifei Gao, Xiangdong Zheng, Cuifang Liu, Shuangshuang Dong, Ru Li, Guanwei Zhang, Yixuan Wei, Hongyuan Qu, Yuhan Li, C. David Allis, Guohong Li, Haitao Li, and Pulong Li. 2019. “Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism”. *Molecular Cell* 76, no. 4 (): 646–659.e6. doi:10.1016/j.molcel.2019.08.019.
- Weimann, Laura, Kristina A. Ganzinger, James McColl, Kate L. Irvine, Simon J. Davis, Nicholas J. Gay, Clare E. Bryant, and David Klenerman. 2013. “A Quantitative Comparison of Single-Dye Tracking Analysis Tools Using Monte Carlo Simulations”. Ed. by Attila Szolnoki. *PLoS ONE* 8, no. 5 (): e64287. doi:10.1371/journal.pone.0064287.

- Weinberger, Leehee, Muneef Ayyash, Noa Novershtern, and Jacob H. Hanna. 2016. “Dynamic stem cell states: naive to primed pluripotency in rodents and humans”. *Nature Reviews Molecular Cell Biology* 17, no. 3 (): 155–169. doi:10.1038/nrm.2015.28.
- Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wickham, Hadley, et al. 2019. “Welcome to the tidyverse”. *Journal of Open Source Software* 4 (43): 1686. doi:10.21105/joss.01686.
- Williams, R. Lindsay, Douglas J. Hilton, Shirley Pease, Tracy A. Willson, Colin L. Stewart, David P. Gearing, Erwin F. Wagner, Donald Metcalf, Nicos A. Nicola, and Nicholas M. Gough. 1988. “Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells”. *Nature* 336, no. 6200 (): 684–687. doi:10.1038/336684a0.
- Wysocka, Joanna, Tomek Swigut, Hua Xiao, Thomas A. Milne, So Yeon Kwon, Joe Landry, Monika Kauer, Alan J. Tackett, Brian T. Chait, Paul Badenhorst, Carl Wu, and C. David Allis. 2006. “A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling”. *Nature* 442, no. 7098 (): 86–90. doi:10.1038/nature04815.
- Xiang, Limin, Kun Chen, Rui Yan, Wan Li, and Ke Xu. 2020. “Single-molecule displacement mapping unveils nanoscale heterogeneities in intracellular diffusivity”. *Nature Methods* 17, no. 5 (): 524–530. doi:10.1038/s41592-020-0793-0.
- Xu, Jianquan, Hongqiang Ma, Jingyi Jin, Shikhar Uttam, Rao Fu, Yi Huang, and Yang Liu. 2018. “Super-Resolution Imaging of Higher-Order Chromatin Structures at Different Epigenomic States in Single Mammalian Cells”. *Cell Reports* 24, no. 4 (): 873–882. doi:10.1016/j.celrep.2018.06.085.
- Yamamoto, Ken, and Miki Sonoda. 2003. “Self-interaction of heterochromatin protein 1 is required for direct binding to histone methyltransferase, SUV39H1”. *Biochemical and Biophysical Research Communications* 301, no. 2 (): 287–292. doi:10.1016/s0006-291x(02)03021-8.
- Ye, Qian, Isabelle Callebaut, Arash Pezhman, Jean-Claude Courvalin, and Howard J. Worman. 1997. “Domain-specific Interactions of Human HP1-type Chromodomain Proteins and Inner Nuclear Membrane Protein LBR”. *Journal of Biological Chemistry* 272, no. 23 (): 14983–14989. doi:10.1074/jbc.272.23.14983.
- Yearim, Ahuvi, Sahar Gelfman, Ronna Shayevitch, Shai Melcer, Ohad Glaich, Jan-Philipp Mallm, Malka Nissim-Rafinia, Ayelet-Hashahar S. Cohen, Karsten Rippe, Eran Meshorer, and Gil Ast. 2015. “HP1 Is Involved in Regulating the Global Impact of DNA Methylation on Alternative Splicing”. *Cell Reports* 10, no. 7 (): 1122–1134. doi:10.1016/j.celrep.2015.01.038.
- Yin, Hang, Sarah Sweeney, Debasish Raha, Michael Snyder, and Haifan Lin. 2011. “A High-Resolution Whole-Genome Map of Key Chromatin Modifications in the Adult *Drosophila melanogaster*”. Ed. by Wolf Reik. *PLoS Genetics* 7, no. 12 (): e1002380. doi:10.1371/journal.pgen.1002380.
- Ying, Qi-Long, Marios Stavridis, Dean Griffiths, Meng Li, and Austin Smith. 2003. “Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture”. *Nature Biotechnology* 21, no. 2 (): 183–186. doi:10.1038/nbt780.
- Ying, Qi-Long, Jason Wray, Jennifer Nichols, Laura Battle-Morera, Bradley Doble, James Woodgett, Philip Cohen, and Austin Smith. 2008. “The ground state of embryonic stem cell self-renewal”. *Nature* 453, no. 7194 (): 519–523. doi:10.1038/nature06968.
- Zanacchi, Francesca Cella, Carlo Manzo, Angel S Alvarez, Nathan D Derr, Maria F Garcia-Parajo, and Melike Lakadamyali. 2017. “A DNA origami platform for quantifying protein copy number in super-resolution”. *Nature Methods* 14, no. 8 (): 789–792. doi:10.1038/nmeth.4342.
- Zhang, Xiaodong, Zhiwei Ye, Xinfu Zhang, Huizi Man, Zhenlong Huang, Ning Li, and Yi Xiao. 2019. “A targetable fluorescent probe for dSTORM super-resolution imaging of live cell nucleus DNA”. *Chemical Communications* 55 (13): 1951–1954. doi:10.1039/c8cc08575g.

- Zhao, Tao, Thomas Heyduk, C. David Allis, and Joel C. Eissenberg. 2000. "Heterochromatin Protein 1 Binds to Nucleosomes and DNA in Vitro". *Journal of Biological Chemistry* 275, no. 36 (): 28332–28338. doi:10.1074/jbc.m003493200.
- Zheng, Qinsi, Anthony X. Ayala, Inhee Chung, Aubrey V. Weigel, Anand Ranjan, Natalie Falco, Jonathan B. Grimm, Ariana N. Tkachuk, Carl Wu, Jennifer Lippincott-Schwartz, Robert H. Singer, and Luke D. Lavis. 2019. "Rational Design of Fluorogenic and Spontaneously Blinking Labels for Super-Resolution Imaging". *ACS Central Science* 5, no. 9 (): 1602–1613. doi:10.1021/acscentsci.9b00676.
- Zhou, Jiansheng, Jun Y Fan, Danny Rangasamy, and David J Tremethick. 2007. "The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression". *Nature Structural & Molecular Biology* 14, no. 11 (): 1070–1076. doi:10.1038/nsmb1323.