Dissertation

submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
M.Sc. Øyvind Ødegård Fougner (Oeyvind Oedegaard)
born in: Haugesund, Norway
Oral examination: 25.11.2019

A method to image the 3D structure of human genes in single cells with 10 kb resolution

Referees: Dr. Jan Korbel

Prof. Dr. Dirk-Peter Herten

This work was carried out at the European Molecular Biology Laboratory in Heidelberg from September 2015 to August 2019 under the supervision of Dr. Jan Ellenberg.

Part of the technical work described in this thesis has been published:

Carl Barton, Sandro Morganella, Øyvind Ødegård-Fougner, Stephanie Alexander, Jonas Ries, Tomas Fitzgerald, Jan Ellenberg, and Ewan Birney. ChromoTrace: Computational reconstruction of 3D chromosome configurations for super-resolution microscopy. PLoS computational biology, 14(3), p. e1006002. doi: 10.1371/journal.pcbi.1006002.

Nike Walther, M. Julius Hossain, Antonio Z. Politi, Birgit Koch, Moritz Kueblbeck, Øyvind Ødegård-Fougner, Marko Lampe, and Jan Ellenberg. A quantitative map of human Condensins provides new insights into mitotic chromosome architecture. The Journal of Cell Biology, 217(7), pp. 2309–2328. doi: 10.1083/jcb.201801048.

Table of contents

Abbreviations

3C	Chromatin conformation capture
3D	Three-dimensional (x, y, z)
4C	Circularized chromosome conformation capture
5C	Carbon copy chromosome conformation capture
B-DNA	B form of DNA as described by Watson, Crick, Franklin, Gosling, Stokes, and Wilson in 1953
bp	base pair
CT	Chromosome territory
CTCF	CCCTC-binding factor
DNA	Deoxyribonucleic acid
DNA-PAINT	DNA-based Point Accumulation for Imaging in Nanoscale Topography
EM	Electron microscopy
FCS	Fluorescence correlation spectroscopy
FISH	Fluorescence *in situ* hybridisation
G1-phase	Gap 1 phase in cell cycle
G2-phase	Gap 2 phase in cell cycle
Hi-C	Genome-wide chromatin conformation capture
M-phase	Mitotic phase in cell cycle
NA	Numerical aperture
PAINT	Point Accumulation for Imaging in Nanoscale Topography
PDB ID	Protein data bank accession number

RD	Replication domain
RNA	Ribonucleic acid
SMLM	Single molecule localisation microscopy
S-phase	Synthesis phase in cell cycle
STED	Stimulated emission depletion
STORM	Stochastic Optical Reconstruction Microscopy
TAD	Topologically associating domain

Summary

The spatial organisation of the genome is essential for its functions including gene expression, DNA replication and repair, as well as chromosome compaction and segregation. Below the level of the large linear chromosomal DNA molecules, more compact topologically associating domains (TADs) have been identified as fundamental units of chromosome structure. However, the actual three-dimensional (3D) folding of DNA within TADs still needs to be understood.

Based on theoretical simulations, we predicted that the nanoscale resolving power of super-resolution microscopy can in principle address this key open question. Here, we present the development of an experimental approach that combines super-resolution microscopy with Exchange-PAINT of barcoded *in situ* hybridisation probes and their computational analysis to extract the 3D path of the linear DNA sequence underlying TADs. We demonstrate that this method can resolve the physical structure of the DNA at a resolution of ~500 bp *in vitro* and ~10 kb in single human cells. Given the predicted genomic loop sizes and our ability to reconstruct the physical DNA path from the positions of combinatorial *in situ* hybridisation labels, the experimental and computational pipeline developed in this thesis is ready to be scaled-up to probe the 3D organisation of entire chromosomes at ~10 kb resolution in single human cells.

Zusammenfassung

Die räumliche Organisation des Genoms ist für seine Funktionen wie Genexpression, DNA-Replikation und -Reparatur sowie Chromosomenkompaktierung und -segregation von entscheidender Bedeutung. Den großen linearen chromosomalen DNA-Molekülen sind kompaktere, topologisch assoziierende Domänen (TADs) untergeordnet, die als grundlegende Einheiten der Chromosomenstruktur identifiziert wurden. Die tatsächliche dreidimensionale Faltung der DNA innerhalb von TADs muss jedoch noch aufgeklärt und verstanden werden.

Basierend auf theoretischen Simulationen sollte das nanoskalige Auflösungsvermögen höchstauflösender Mikroskope prinzipiell diese Verständnislücke füllen können. Wir präsentieren hier die Entwicklung eines experimentellen Verfahrens, das höchstauflösende Mikroskopie mit dem PAINT-Austausch von barcodierten In-situ-Hybridisierungssonden und computergestützter Analyse ihrer Lokalisation kombiniert, um den 3D-Pfad der den TADs zugrunde liegenden linearen DNA-Sequenz zu extrahieren. Wir zeigen, dass diese Methode die physikalische Struktur der DNA bei einer Auflösung von ~500 bp *in vitro* und ~10 kb in einzelnen menschlichen Zellen darstellen kann. Basierend auf den vorhergesagten Genomschleifengrößen und unserer Fähigkeit, den physikalischen Pfad der DNA aus den Positionen kombinatorischer In-situ-Hybridisierungsmarkierungen zu rekonstruieren, kann nun die in dieser Arbeit entwickelte experimentelle und rechnerische Pipeline für die Untersuchung der 3D-Organisation ganzer Chromosomen bei einer Auflösung von ~10 kb in individuellen menschlichen Zellen hochskaliert werden.

Chapter 1: Introduction

The structure of deoxyribonucleic acid (DNA)

The genome of a human cell encodes the information for all its constituents in about six billion base pairs (bp) of deoxyribonucleic acid (DNA) divided in two homologous sets of 23 chromosomes. 10 bp form one helix turn in the B-DNA form of the double helix and each bp is 0.43 nm long when measured *in vitro* (Franklin and Gosling, 1953; Watson and Crick, 1953; Wilkins, Stokes and Wilson, 1953) (Figure 1). However, it may have a slightly different conformation in solution (Wang, 1979). This means that the total length of DNA in each human cell nucleus measures about two metres [(0.34 nm x 10^9 bp) x (6 x 10^9)].

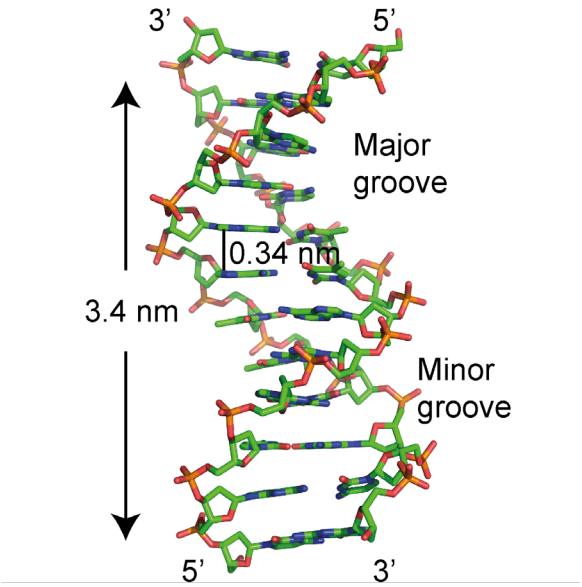

Figure 1: Crystal structure of B-DNA double helix. *Resolved at 1.9 Å resolution (PDB ID: 1BNA). 3', 5', 3- and 5-prime end of the DNA. Physical distances in nm are indicated.*

When considering that typical human cell nuclei are only a few picolitres in size or a few micrometres in diameter, it becomes apparent that a large amount of folding is required to fit 2 m of linear DNA polymer into a nucleus. At the same time the DNA fibre has to remain sufficiently open and spatially organised to generate messenger RNA (mRNA) for protein production in interphase. In addition, DNA has to be rearranged to progress from the G1-phase of interphase, with two copies of the genome, into the S-phase where the genome is duplicated. Following duplication, cells containing four copies of the genome in the G2-phase need to further compact the chromosomal DNA molecules in order to enter mitosis and accurately distribute an identical copy of the genome to the two newly forming daughter cells. It has been studied for decades how the genome is folded and how this is regulated so that the features mentioned above and other functionalities are facilitated. However, with the development of many new methods in molecular biology, many additional strides in our knowledge have only been made recently.

The current model of chromatin compaction

The DNA of 46 chromosomes exists in a complex with closely associated proteins, jointly referred to as chromatin, a term coined initially for the ability to stain the substance inside nuclei (Flemming, 1882). Although the presence of nucleic acids in cell nuclei has been known for more than a century (Miescher, 1871), their structural organisation and compaction mechanisms in the nucleus are still under debate and investigation.

The structure of the nucleosome

The first layer of compaction of the genome is organised by the most abundant structural proteins of chromatin called histones (Kossel, 1911). Two copies of H2A, H2B, H3 and H4 assemble in an octameric complex that forms a flat cylinder around which 147 bp (or 146) of DNA can coil to form the DNA-protein complex called the nucleosome core particle (commonly referred to as nucleosome) (Figure 2). The structure of the nucleosome core particle is known and has been resolved as highly as 1.9 Å (Richmond *et al.*, 1984; Davey *et al.*, 2002). An additional histone protein called linker histone (H1) binds to 20 nt of DNA in the linker region between the nucleosome core particles (Thoma and Koller, 1977; Simpson, 1978; Zhou *et al.*, 2015) (Figure 3).

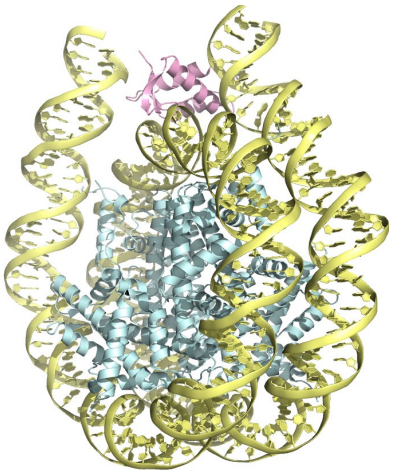

Figure 2: The structure of the nucleosome core particle *(PDB ID: 4qlc; adapted from: Zhou et al., 2015). DNA (yellow) Core histones (blue); linker histone H1 (pink); Image produced with PyMol v1.3.*

Nucleosome core particles and their respective linker regions are referred to as 11-nm (chromatin) fibre (appearing as beads on a string in an *in vitro* electron micrograph; Figure 3 A) (Grigoryev *et al.*, 2009) and are positioned approximately every 200 bp along the eukaryotic genome (Kornberg, 1974). They function as the first basic folding unit of DNA (Kornberg and Thomas, 1974; Olins and Olins, 1974). If one considers the genome to be a 2 m long one-dimensional polymer, the 11-nm fibre formed by beads-on-a-string nucleosomes spaced every 200 bp would be 0.53 m in length [2 m x (200 bp – 147 bp / 200 bp)], which corresponds to a 3.8-fold linear compaction when DNA

is coiled around nucleosomes [200 / (200 - 147)], and a 6.1-fold linear compaction when H1 further increases compaction [200 / (200 - 167)].

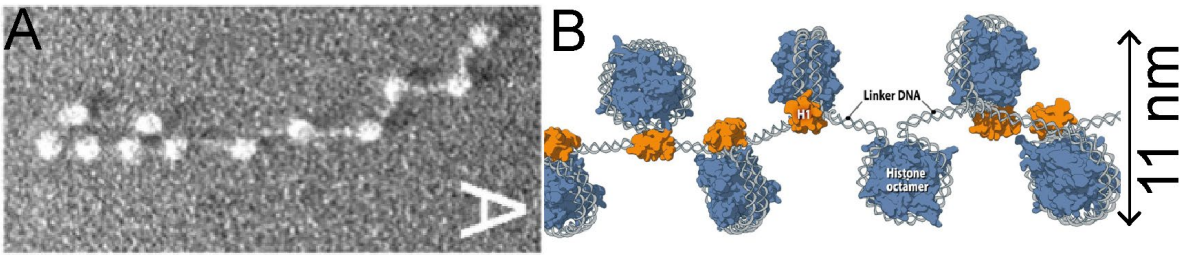

Figure 3: 11-nm chromatin fibre. *(A) Electron microscopy image* (Grigoryev *et al.*, 2009) *(B) Cartoon of 11-nm fibre containing nucleosome core particles (histone octamer plus DNA) and linker regions (histone H1 plus DNA). Modified from (MBINFO, 2017).*

No evidence for 30-nm fibre *in situ* or *in vivo*

In the hierarchical model of chromatin compaction, the next compaction level is referred to as the 30-nm fibre. This fibre supposedly consists of nucleosomes that are stabilised by the linker histone H1 and stacked-up next to each other in a fibre of about 30 nm in diameter. The most popular theoretical models of the 30-nm fibre are the one-start solenoid model (Robinson *et al.*, 2006) and the two-start zig-zag model of which existence has been shown *in vitro* (Schalch *et al.*, 2005; Song *et al.*, 2014). Nevertheless, the existence of such structures is highly debated and they have not been observed *in vivo* or *in situ* in higher eukaryotes (Reviewed in Tremethick, 2007; Joti *et al.*, 2012) although current electron microscopy methods can resolve individual nucleosomes (Eltsov *et al.*, 2008).

Replication domains (RDs) synchronise their replication origins

During replication of the mammalian genome in S-phase there are ~5,000 stable units of chromosomes called replication domains (RDs). In each RD there are about six replication origins that fire synchronously (Jackson and Pombo, 1998; Rivera-Mulia and Gilbert, 2016). To visualise these RDs, our group has labelled the co-replicating foci on single chromosomes in a sequence-unspecific manner and applied correlative confocal and super-resolution microscopy to investigate RD structure *in situ* (Xiang *et*

al., 2018) (Figure 4). We discovered RDs to have a median physical size of 150 nm and that they are spaced 300 nm apart from each other along the chromosome. RDs on the same chromosome no longer correlate in their movement if they are further than 550 nm apart, indicating that the boundaries between them are rather flexible (Xiang *et al.*, 2018).

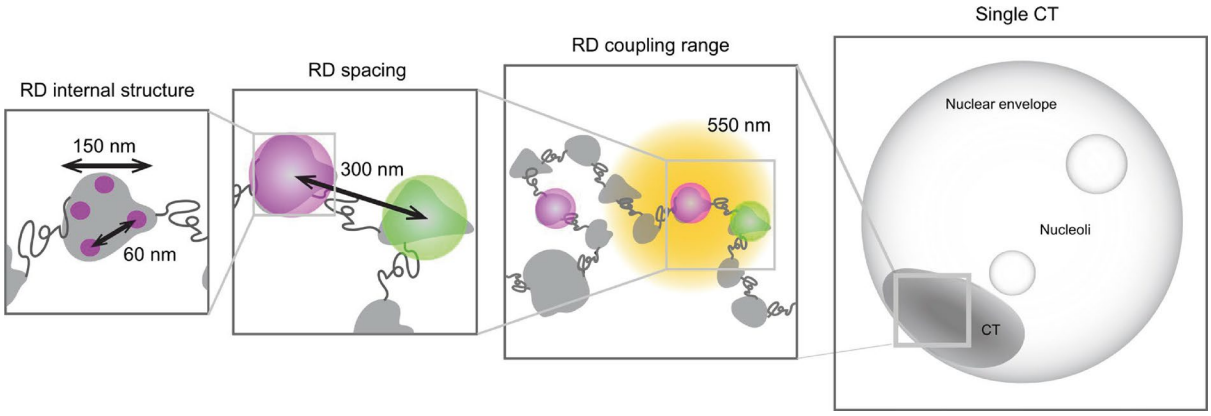

Figure 4: Quantitative model of replication domains organising chromosome territories *showing the 150 nm median size of replication domains (RDs), their 300 nm spacing and 550 nm coupling range within each chromosome territory (CT). Modified from (Xiang et al., 2018).*

Topologically associating domains (TADs) are stable units of the genome and are essentially the same structures as RDs

The development of chromatin crosslinking techniques, such as 3C (Dekker, 2002), 4C (Lomvardas *et al.*, 2006; Simonis *et al.*, 2006; Würtele and Chartrand, 2006; Zhao *et al.*, 2006), 5C (Dostie *et al.*, 2006) and especially Hi-C (Lieberman-Aiden *et al.*, 2009), have provided an indirect way to study higher order structures beyond the 11-nm fibre. With the Hi-C method the contact frequency between all points in the genome is mapped with a genomic resolution ranging between 1 kb and 6 kb depending on the restriction enzyme used. In the resulting contact frequency maps, regions of ~400-800 kb were observed to cluster spatially. These stretches were termed topologically associating domains (TADs) (Rao *et al*. 2014; Nora *et al*. 2012; Pope *et al*. 2014) and correspond to RDs (Moindrot *et al.*, 2012; Pope *et al.*, 2014; Dileep *et al.*, 2015). In the remainder of this work we will refer to both RDs and TADs as TADs.

TADs contain smaller loops

One of the basic principles of TAD structure is that a loop extrudes between two CCCTC-binding factor (CTCF) proteins binding sites spaced ~400-800 kb apart (see Figure 5). Theoretical modelling of Hi-C maps also suggests that these loops may contain sub-structures in the form of DNA loops of ~100 kb in size (Phillips-Cremins *et al.*, 2013; Rao *et al.*, 2014). The borders between TADs are demarked by the DNA sequence motif CCGCGNGGNGGCAG which serves as a binding site for CTCF (Dixon *et al.*, 2012; Dowen *et al.*, 2014). The CTCF protein, in turn, has been shown to associate with cohesin (Parelho *et al.*, 2008; Stedman *et al.*, 2008; Wendt *et al.*, 2008), and CTCF partakes in cohesin positioning (Busslinger *et al.*, 2017). Cohesin is a large protein complex with many protein subunits that was originally found to keep the sister chromatids crosslinked from S-phase until anaphase (Losada, Hirano and Hirano, 1998; Tóth *et al.*, 1999; Sumara *et al.*, 2000). However, cohesin has been shown later to crosslink chromatin throughout interphase (Splinter *et al.*, 2006) and disrupt TAD boundaries when removed (Nora *et al.*, 2017; Rao *et al.*, 2017; Schwarzer *et al.*, 2017; Wutz *et al.*, 2017).

TADs are evolutionarily conserved and divergent CTCF positioning across species correlates with dissimilar domain structure

Across evolution, the use of TADs or similar architectural structures to shape chromosomes into functional partitions has been conserved. Besides mammals, TADs have been described in *Drosophila* and observed on the X chromosome of *C. elegans* (Vietri Rudan *et al.*, 2015). Some plants also show clear genome partitioning and in both *S. pombe* and *S. cerevisiae*, self-interacting domains have been identified (Hsieh *et al.*, 2015). In several type of bacteria, studies have shown the presence of chromosomal interaction domains (CIDs) which resemble eukaryote TADs and play a role in transcriptional regulation (Marbouty *et al.*, 2015).

In mammals, TAD boundary positions are primarily conserved across species. The protein CTCF is enriched at the border of TADs and plays a role together with cohesin in chromatin loop formation. Conserved CTCF binding sites have been detected at

positions of strong contact insulation and conserved boundaries, whereas divergent binding sites are found within TADs, generating different sub-TAD structures which might contribute to the variation in gene expression observed between species (Vietri Rudan *et al.*, 2015).

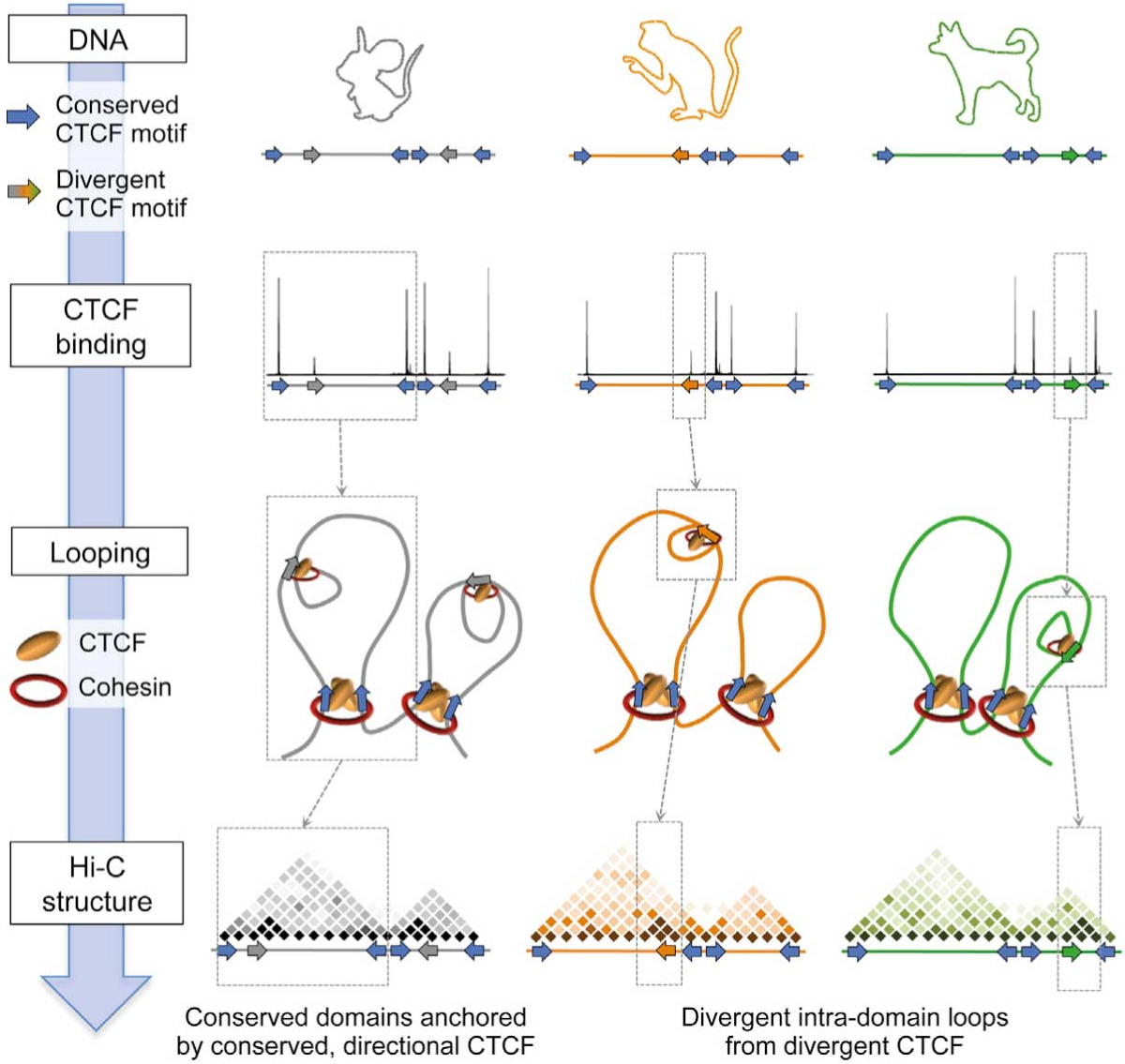

Figure 5: Relationship between CTCF binding sites, TAD/loop structures and Hi-C maps. *Adapted from (Vietri Rudan et al., 2015). Chromatin loops are formed between two CTCF binding sites stabilised by cohesin.*

TADs organise into active and inactive compartments

Electron microscopy (EM) is suitable to image DNA spread out on a large surface but does not have the power to resolve the DNA path inside the nucleus of a cell due to the very large degree of crowding of DNA strands. However, it can distinguish between euchromatin, which is not strongly stained except during cell division, and heterochromatin, which is strongly stained throughout the cell cycle (Heitz, 1928; Passarge, 1979). Euchromatin is associated with active regions of chromatin, whereas heterochromatin is compact and inactive (Cooper, 1959). Heterochromatin can either be constitutive (always off) or facultative with varying gene expression depending on differentiation (Schrader, 1921; Hughes-Schrader, 1948; Brown and Nur, 1964).

Upon correction of the average dependence of contact frequencies on genomic distance in Hi-C maps, a checkerboard pattern of around ~1 Mb appears at the level of TADs. When CTCF or cohesin is prohibited to bind to CTCF binding sequences in the genome, TAD boundaries are removed and groups of TADs combine and form so-called chromatin compartments (Nora *et al.*, 2017; Rao *et al.*, 2017; Schwarzer *et al.*, 2017; Wutz *et al.*, 2017). These compartments are either active (A), or inactive (B) correlate to a large extent with euchromatin and heterochromatin, respectively, and are defined by their epigenetic state rather than CTCF binding sites (Lieberman-Aiden *et al.*, 2009).

Mechanisms of compartmentalisation

There is currently no consensus on how chromatin compartments are formed and kept. Theoretical predictions have suggested that phase separation may play a crucial role both in keeping the compartments contained and in keeping them apart from each other while still allowing the DNA polymer to be flexible and adaptive (Jost *et al.*, 2014; Di Pierro *et al.*, 2016, 2017; Erdel and Rippe, 2018). Alternative models suggest that chromatin is both anchored to the nuclear lamina and to nuclear speckles, such as the nucleolus, which limits the conformations that each chromosome can have. Dynamic differences in the genome have also been proposed to serve as the driver of compartmentalisation and positioning (Ganai, Sengupta and Menon, 2014).

At the lowest level of compaction, confocal microscopy data has shown that each chromosome occupies a defined volume or “territory” (Cremer *et al.*, 1982; Bolzer *et al.*, 2005) (Figure 6, upper right image), although this has first been proposed from observations made in *Salmonella* more than a century ago (Rabl, 1885). In the regions where chromosome territories meet, there is, however, a significant amount of intermingling of fibres (Branco and Pombo, 2006). The lack of knowledge about chromatin organisation at the scale between 10 nm and 200 nm lies below the diffraction limit of light and consequently requires higher resolution than normal confocal microscopy can offer.

To summarise, each chromosome has its own distinct territory and is divided into an alternating pattern of active and inactive compartments with a size of approximately 1 Mb each. Each compartment consists of several TADs that are formed by cohesin-mediated loop extrusion at CTCF binding sites. The polymer that forms these loops is the 11-nm fibre which represents DNA coiled around histone proteins.

Until recently it has not been possible to connect information about the physical structure of DNA acquired by light or EM to contact frequency data resulting from Hi-C. This would require a technology that can spatially resolve individual loops of the DNA/nucleosome fibre and map it to specific DNA sequences in the genome. The major challenge is that none of the methods mentioned above adequately bridge the resolution gap between the nm and µm scale (Figure 6).

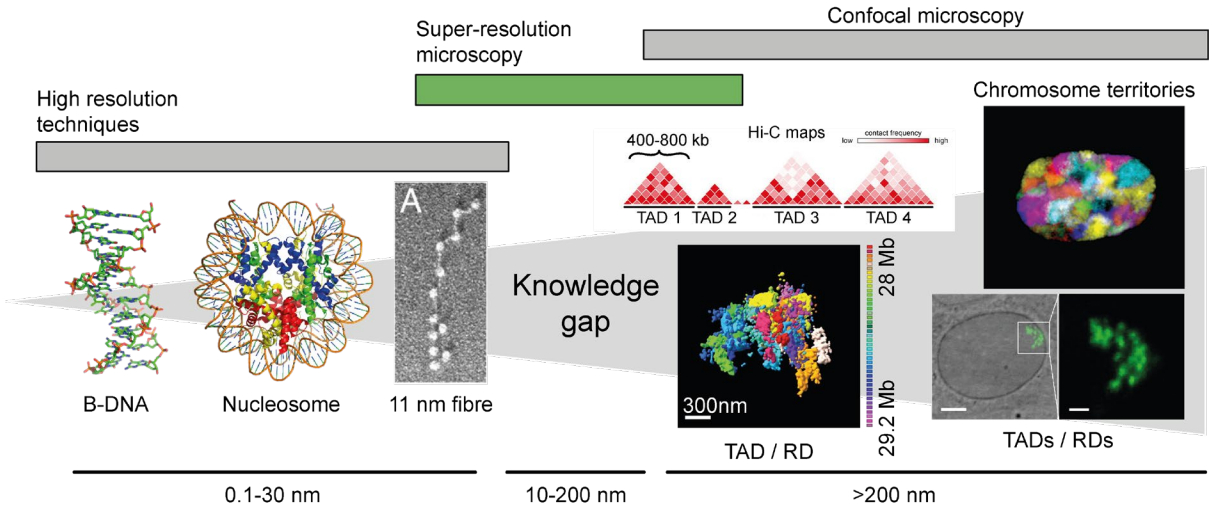

Figure 6: DNA and chromatin at different resolution scales. *High-resolution techniques (EM) enable structural analysis of B-DNA (PDB ID: 1BNA; Drew et al., 1981), nucleosomes (PDB ID: 2CV5; Tsunaka, 2005) and the 11-nm fibre in vitro (Grigoryev et al., 2009), but fail to resolve DNA at larger scales than this in situ because of the high density of DNA compaction in chromatin. Confocal microscopy can resolve individual chromosome territories (Bolzer et al., 2005; Speicher and Carter, 2005), and can depict TADs/RDs as diffraction-limited points (Xiang and Roberti et al., 2018). Super-resolution light microscopy methods could be utilised to fill the knowledge gap about chromatin organisation between 11 nm and 200 nm (Bintu et al., 2018) to investigate the internal structures of TADs beyond what could be deduced from contact frquency Hi-C maps (Schwartz and Cavalli, 2017).*

Circumventing the diffraction limit of light

The diffraction limit of light describes how close two objects can come together to still be resolvable by optical microscopes. All optical microscopes focus light of a given wavelength and all lenses and objectives have critical angles from which they collect or reject this light. The range of angles that is collected is described by the numerical aperture value (NA). Together with the wavelength, the smallest NA in the light path determines the best achievable resolution of a microscope (Abbe, 1874) (Equation 1).

$$d = \frac{\lambda}{2NA}$$

Equation 1: Abbe diffraction limit (d) *is determined by the wavelength (λ) and the numerical aperture (NA).*

Super-resolution microscopes circumvent the diffraction limit of light by only imaging fluorophores that are further apart from each other than the diffraction limit at a given time. This can be achieved in several ways. In this study stochastic super-resolution microscopy is used. Here, fluorescent molecules are activated stochastically at different times resulting in the emission of only a single fluorophore within a resolvable region. This is also called single molecule localisation microscopy (SMLM) and encompasses e.g. stochastic optical reconstruction microscopy (STORM) which uses sequential activation of photoactivatable fluorophores (Rust, Bates and Zhuang, 2006), points accumulation for imaging in nanoscale topography (PAINT) (Sharonov and Hochstrasser, 2006) and DNA-PAINT (Jungmann *et al.*, 2010). The latter uses transient binding of fluorescent molecules to the target structure to introduce fluorescence stochastically. Since only one fluorescent molecule is imaged at each location simultaneously, the subpixel localisation of this molecule can be calculated highly accurately if a sufficient number of photons is collected.

The scale of the problem

A confocally resolvable volume by a state-of-the-art microscope is ~140 nm in x and y and ~400 nm in z. In a human nucleus, this volume contains on average more than 2 Mb of DNA which is highly folded into compact chromatin. Thus, the internal loop structures of TADs, which are estimated to be 10s to 100s of kb, are clearly not resolvable by diffraction-limited microscopy methods. By contrast, state-of-the-art 3D SMLM (e.g. 4Pi-STORM or iPALM; resolution of ~20 x 20 x 20 nm ≈ 8 x10^-6 μm^3) can resolve approximately 1000-fold smaller volume elements and therefore potentially probe more than 80 million points in a single human nucleus, whereby each resolvable unit would contain less than 2,000 bp. This means that the looping substructures of the ~400-800 kb-sized TADs should be easily resolvable although one cannot expect to resolve individual nucleosomes.

Sequence-specific labelling of DNA

Sequence-specific labelling of DNA can be achieved with fluorescence *in situ* hybridisation (FISH). A good correlation between the contact frequencies recorded in Hi-C maps and the physical distance between exemplary genomic loci observed by FISH probes has been found (Lakadamyali and Cosma, 2015; Bintu *et al.*, 2018).

Modern FISH probes, such as oligopaint FISH (Beliveau *et al.*, 2015), do not directly attach fluorescent dyes to the primary probe hybridising to the genomic target sequence as it is done in traditional FISH techniques, such as BAC-based FISH approaches (BACFISH). In modern FISH probes, a non-genome-complementary docking handle is added to the end of the primary probe. This primary probe can then be targeted with a secondary DNA oligonucleotide, called imager strand, which carries the fluorophore (Figure 7). As the sequence and length of the docking handle/imager strand complex can be freely designed, it enables rapid binding, replacement or removal of the imager strand and significantly reduces the costs of the primary probe library. Moreover, by using different docking handle/imager strand sequence combinations, many regions can be labelled with unique "barcodes" without any crosstalk, as long as the DNA sequences of the docking handles are significantly different from each other. The length and GC content of the docking handle determines how strongly the imager strand binds to it, which can be modulated. A docking handle of >20 nt can be used for permanent binding, ~12 nt will result in an intermediate binding time of about 1000 s and 9-10 nt will give transient binding of ~1 s. (Pers. Comm., Ralf Jungmann (MPI, Martinsried, Germany)).

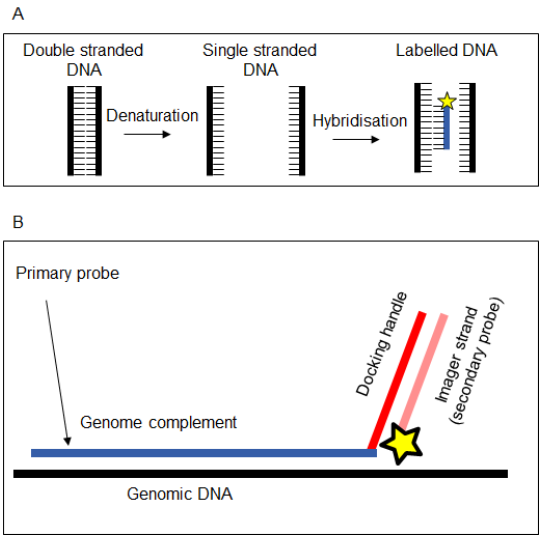

Figure 7: Traditional FISH approach (A) versus secondary imager strand approach (B) *Docking handle and imager strands. Black strand, genomic DNA; blue strand, genomecomplementary probe; red strand, docking handle; pink strand, imager strand; star, fluorophore.*

However, the best resolution reached by FISH to date has not been sufficient to determine the substructure of TADs or reconstruct the size and shape of hypothetical DNA loops in cells with genome-specific probes (Beliveau *et al.*, 2015; Ni *et al.*, 2017; Nir, Farabella, Pérez Estrada, *et al.*, 2018).

Aim and Approach

The recent advances in imaging and DNA labelling technology introduced above have provided a basis for me to develop a technology to resolve the 3D folding of the genome sequence at the kilobase scale *in situ*. If successful, this would allow for the first time to directly determine the internal looping structure of key chromosome structuring elements, such as TADs, in single human cells. To reach this ambitious goal, my strategy was to combine oligopaint FISH with high-resolution 3D light microscopy of genomic loci in single human cells. My overall aim was to develop a method that is not only able to resolve the 3D path of DNA through the dense structure of a TAD at ~10 kb resolution but is also scalable to an entire chromosomal DNA molecule.

I approached this ambitious challenge in three steps. My first goal was to establish the technology to resolve neighbouring genomic loci that are only 10 kb apart in 3D in the nucleus of a human cell. To this end, I first tested the labelling efficiency of oligopaint FISH and the resolving power of 3D locus imaging on pure DNA mini-chromosomes *in vitro*. Secondly, I set up comparable assays for labelling efficiency and resolution in human cells to determine the optimal combination of oligopaint FISH probe design and hybridisation conditions which preserve nuclear architecture as much as possible. After systematic optimization of experimental conditions and probe design, I then finally moved on to establish a multiplexing workflow that allows me to exchange ten differently barcoded FISH probes and record all labelled loci at high resolution in 3D from single nuclei.

The integrated experimental and computational pipeline I developed allows me to extract the path of the targeted linear genome sequence. Due to the power of oligopaint barcoding and multiplexing, it can be directly scaled with further automation to provide the technology to assemble the first directly observed 3D map of a whole human chromosome at 10 kb resolution. This technology will be invaluable for the field

to link the internal structure of TADs to the function of the DNA sequence they encode, for example during cell cycle transitions and cellular differentiation.

Chapter 2: Results

Measurement of maximal and minimal DNA-protrusion in space

As described in detail in the introduction, the actual compaction of DNA is surprisingly difficult to measure at sub-TAD resolution on large stretches of DNA. To experimentally investigate the maximal and minimal length that a region of DNA can occupy in space, DNA-origamis containing a 3-by-4 grid pattern with 20-nm-spaced docking handles (P1, 20-nm-grid) were assembled and attached to a glass surface and imaged with DNA-PAINT imager strands (Figure 8 A-C). The DNA scaffold of the 20-nm-grid is organised in 24 linear stretches with four docking handles being 64 bp apart on three of the stretches and six non-labelled linear scaffold stretches in between (Figure 8 C). 64 bp corresponds to a stretch of a 21.8-nm-long B-DNA(Franklin and Gosling, 1953; Watson and Crick, 1953; Wilkins, Stokes and Wilson, 1953). To estimate how DNA can maximally extend under these conditions, 69 of the 20-nm-grids containing at least three spots in one direction and four spots in the other direction were manually identified in overview images (Figure 8 D). Spot-to-spot distances were then extracted in the direction of the linear scaffold (direction with four spots) (Figure 8 E), yielding a mean nearest neighbour distance of 20.2 ± 0.2 nm (Figure 8 F).

Furthermore, the distance between the spots measured perpendicular to the scaffold (in the 3-spot-direction) was 19.9 ± 0.2 nm suggesting that two adjacent stretches of the DNA scaffold can be as close as 2.8 nm (Figure 8 G-I). If one considers the distances of DNA with each base pair occupying a cylinder of 0.32 nm in height [20.2 nm / 64 bp] and 1.4 nm in radius, this would mean that each base pair occupies a volume of 1.9 nm^3, suggesting that even in extremely dense regions of DNA there is some space between the 1 nm^3 -sized base pairs, especially between adjacent stretches of DNA.

This shows that short stretches of DNA have similar dimensions as the theoretical values derived from structural biology under *in vitro* conditions when measured with DNA-PAINT. It also suggests that extremely high resolution in the single-nanometre domain is required to separate two adjacent backfolded DNA double helixes even if they are far apart in the linear DNA molecule.

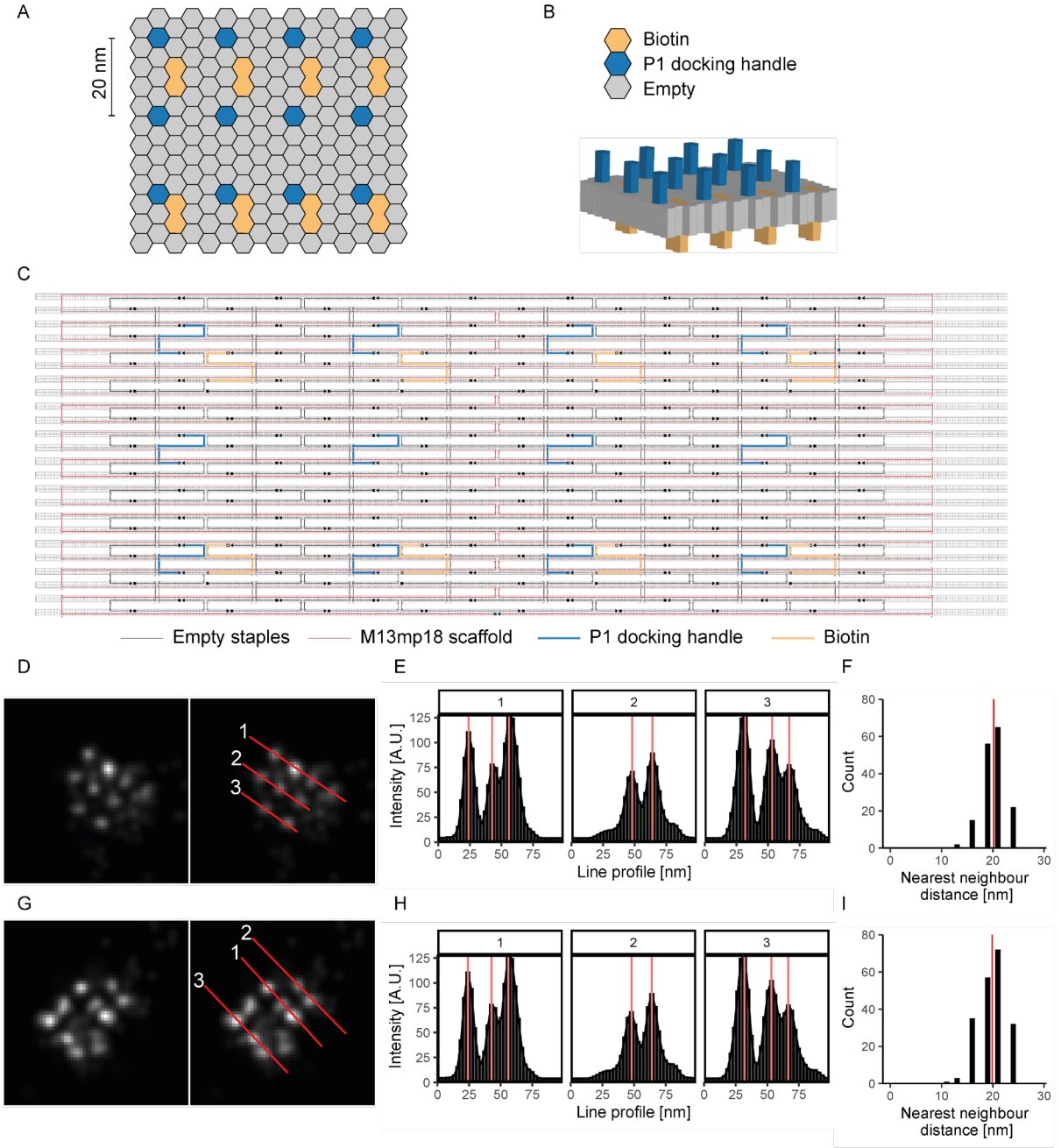

Figure 8: Schematic view of mini-chromosomes designed with Picasso Design (Schnitzbauer et al. 2017). *Layout (A) and 3D rendering (B) of the 20-nm-grid. (C) Detailed 20-nm-grid layout. (A-C) Orange, biotin-containing docking handle; blue, P1 docking handle; red, M13mp18 scaffold DNA; grey, empty position. (D) Representative image of a 20-nm-grid with and without line profiles in the direction of 4 docking handles. (E) Local maxima of line profile 1-3 as labelled in D. (F) Mean nearest neighbour distance between nearest local maxima in D. n = 69 traces and 237 connected maxima. (G) Example image of a 20-nm-grid with and without line profiles in the direction of 3 docking handles. (H) Local maxima of line profile 1-3 as labelled in G. (I) Mean nearest neighbour distance between nearest local maxima in G. n = 112 traces and 316 connected maxima.*

Resolving the path of mini-chromosomes *in vitro* at < 1kb resolution

To estimate the labelling efficiency and resolution of FISH with imager strands on less ordered DNA molecules than origamis, synthetic mini-chromosomes were generated by linearisation of 6.5 kb M13 phage DNA and deposited on glass slides. To label the mini-chromosome all along its length, ten sets of primary FISH probes were designed, each set containing three 32-nt-long probes with the same docking handle sequence. Probes were then hybridised to the mini-chromosome at a genomic distance of 632 nt between each set. In addition, to potentially visualise the path of the linker DNA between these ten loci, six probes with a generic imager strand were hybridised to all linker sequences with a spacing of 64 nt (Figure 9 A; Table 1). Each locus and linker DNA were imaged during 11 rounds (E_0-E_{10}) of DNA-Exchange-PAINT in collaboration with Ralf Jungmann (MPI, Martinsried, Germany) by sequentially adding imager strands complementary to each docking handle. The 20-nm origami grids and corresponding imager strands were added in all exchange rounds and used as reference to correct for sample drift during the 10-hours imaging. 98 grids detected in all rounds were distributed all over the 41-by-41 µm field of view and used for drift correction. After drift correction, a composite multicolour image was generated from all exchange rounds, whereby each exchange E_0-E_{10} was assigned a different pseudo-colour channel (Figure 9 B).

On individual mini-chromosomes, loci spaced apart by 632 bp could clearly be optically resolved and the path of several loci along the chromosome could be traced unambiguously (Figure 9 C). Unexpectedly, even the six individual probes with the generic barcode could often be resolved as individual discrete spots along the linker DNA despite being spaced 64 bp apart (highlighted E_0 in Figure 9 C-3). Line profiles were drawn through 44 stretches of mini-chromosomes if three or more spots were seen in the E_0 channel. A total of 219 peaks were identified and the mean distance between the nearest neighbours was 18 ± 8 nm which is about 90% of the distance of B-DNA (Figure 9 D-E).

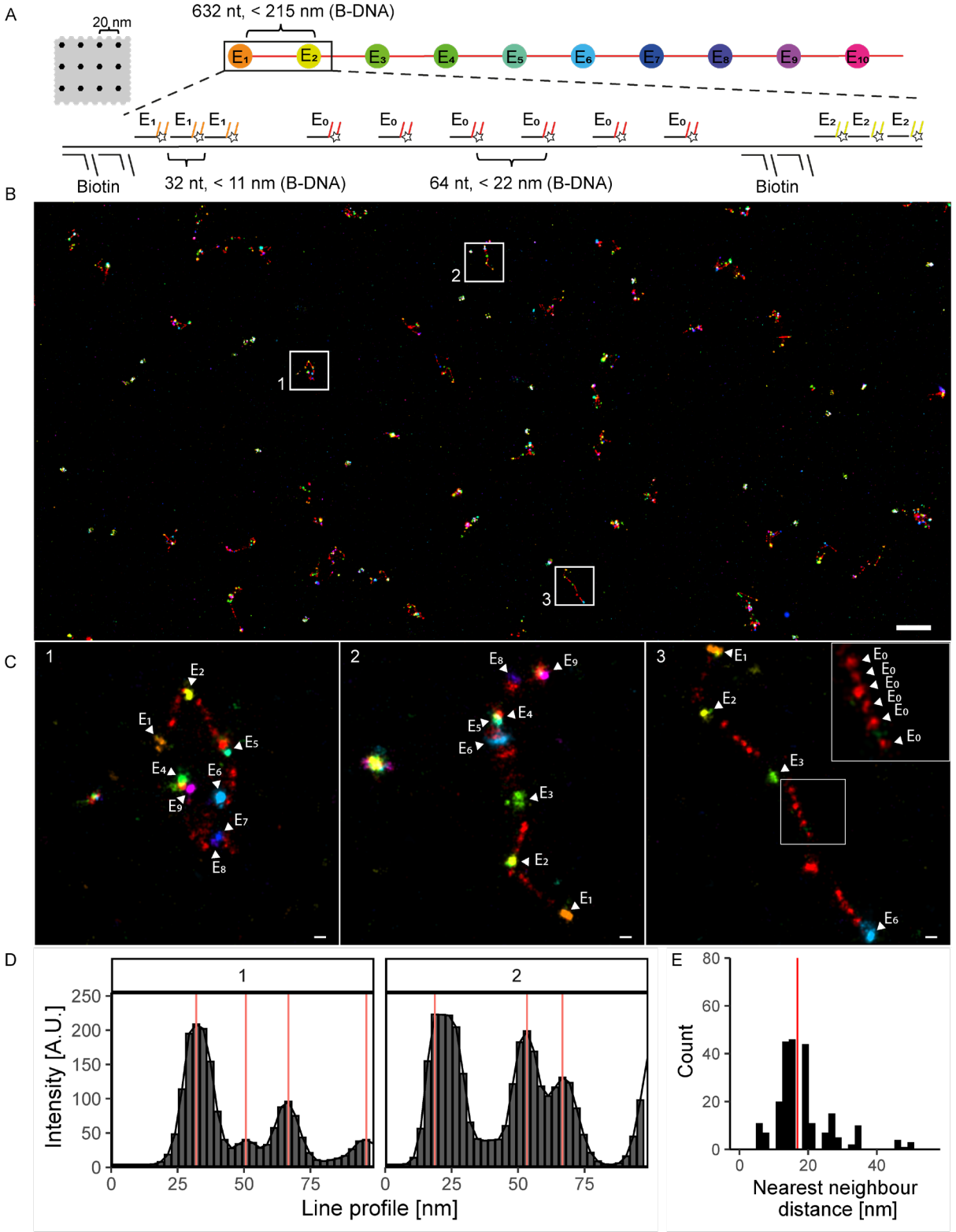

Figure 9: DNA-Exchange-PAINT of mini-chromosomes in vitro at < 1kb resolution. *(A) Mini-chromosome targeted at ten 632-nt-spaced loci with three probes with the same "colour" per locus. (B) Overview of merged DNA-Exchange-PAINT experiments. (C) Zoom-in of 3 regions in B where. Arrowheads denote the Exchange round. In region 3 insert a zoomed region of E_0 is shown. (D) Line profile of two representative regions across localisations in E_0-channel. Maximum value per peak*

highlighted by red line. (E) Nearest neighbour distances with median distance highlighted.

The majority of localisations had a precision of 1.3 ± 0.2 nm (Figure 10 A). The number of localisations per mini-chromosome gradually decayed over the extended imaging time (Figure 10 A-B). Using DBSCAN clustering for automatic identification of loci (E_1-E_{10}), manual annotation of mini-chromosomes and removal of signal from 20-nm-grids, 201 mini-chromosomes were found. Each mini-chromosome consisted on average of 10 ± 4 loci, i.e. about one per exchange. Theoretically, each mini-chromosome could have extended up to 2150 nm if it was in a straight B-DNA conformation. However, this was never observed and all chromosomes were contained within the 800 nm search window radius. In this search window around each selected mini-chromosome on average 0.9 ± 0.7 loci were detected per exchange with the exception of E_3, E_4, E_9, and E_{10} (Figure 10 B). These had a multiple testing-adjusted, statistically significant different number of clusters per mini-chromosome region compared to E_0. This suggests that E_3 and E_4 have some crosstalk between other exchanges. E_4 had more than twice the expected number of loci. E_{10} had a significantly lower number of loci per exchange in line with the fact that the z-focus was suboptimal at the end of the experiment. However, it does not exclude the possibility that E_{10} has a lower affinity to its docking handle (Figure 10 B). Overall, there is a significant decay of the imaging quality of the 20-nm-grids across the exchanges (Figure 10 C), probably due to photoactivated crosslinking between imager strand and docking handles that leads to docking handle decay over extended imaging (Pers. Comm., Ralf Jungmann (MPI, Martinsried, Germany)).

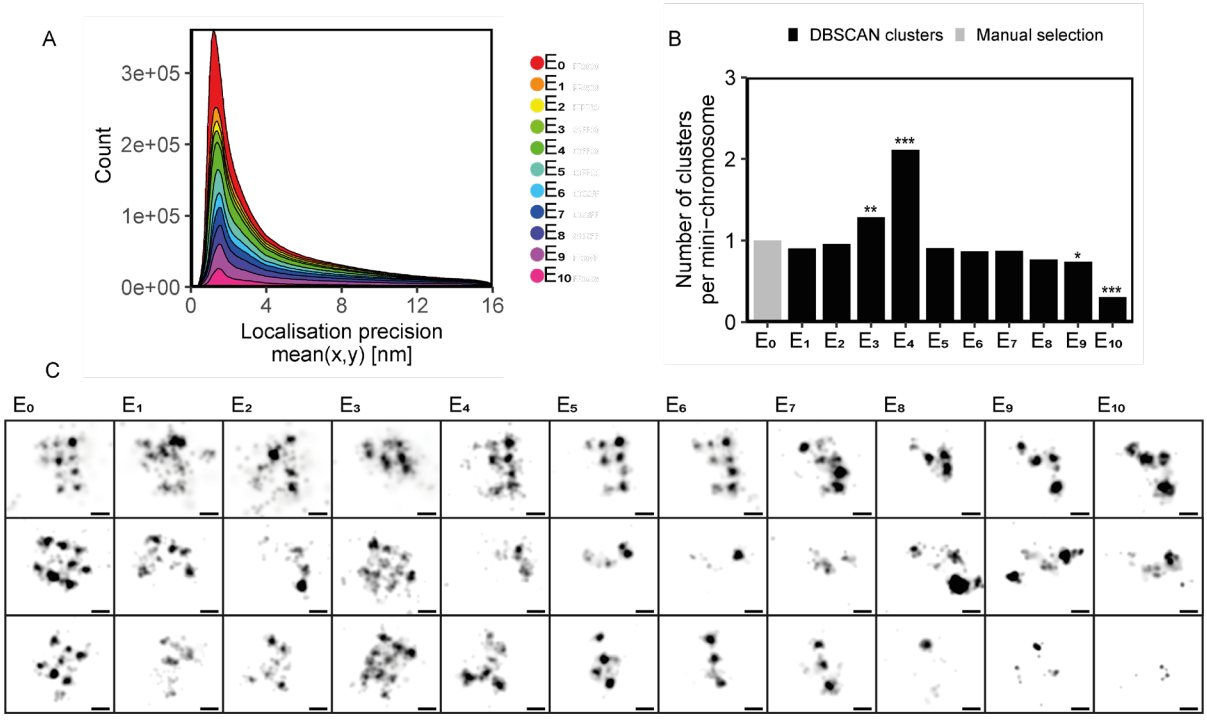

Figure 10: Quality assessment of DNA-Exchange-PAINT of mini-chromosomes. *(A) Localisation precision during exchange rounds. (B) Number of clusters per exchange round on each mini-chromosome. Statistics comparing all exchanges to E_0 with 1-way-anova and a Tukey's 'honestly significant difference' post hoc test (* $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$). (C) DNA origami reference structures across all exchanges. Scale bars, 20 nm.*

All possible traces between individual loci in a mini-chromosome cluster were determined with a simplified version of ChromoTrace (see section below) after removing the 20-nm-grids by removing regions with signal from more than 3 loci in a 74-nm-diameter threshold [d = root(40^2 nm^2 + 60^2 nm^2) + 1.5 nm] (Barton *et al.*, 2018). Connections with more than one "missing" locus were removed. If more than one trace was detected, the shortest was selected. The traces followed the general direction of the generic spacer probe E_0, but the path outlined by E_0 often showed extra loops which are not captured by the 632-nt-spaced probes (Figure 11 A). There are no examples of mini-chromosomes with more than 8 connected loci, also in accordance with the E_0 signal, which suggests that the mini-chromosomes are not extending any further (Figure 11 A-B). The mean distance between the loci is 130 ± 108 nm, corresponding to about half of the theoretical distance that a B-DNA helix could maximally extend (Figure 11 C). This confirms our expectation that the backbone of longer stretches of DNA is rather flexible and thus the overall distance between these loci is shorter due to folding.

When only considering the relative distances between genomic loci, the Euclidean distance shows a near-linear increase of about 14-15 nm per extra 632 nt in the range between 632 and 4424 bp. Longer connections are rare, making the reliability of the data weak above 3792 bp which corresponds to loci that are six exchanges apart (Figure 11 D).

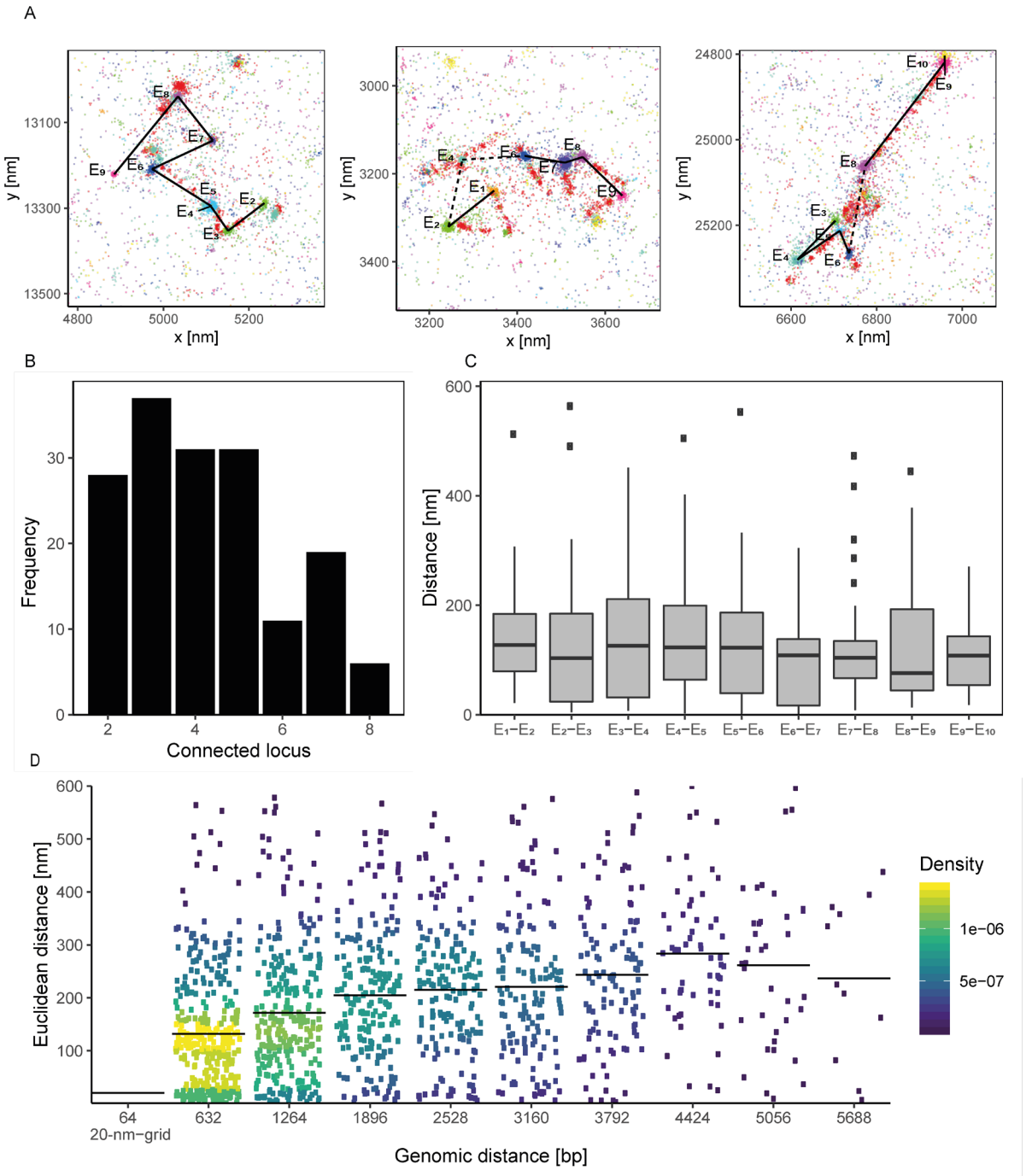

Figure 11: ChromoTrace connections along mini-chromosome paths. *(A) Examples of mini-chromosomes. Dotted lines indicate connections with a missing DNA-Exchange-PAINT signal. (B) Histogram of the number of connected loci with less than one missing probe. (C) Euclidean distance between neighbouring loci. (D) Relationship between Euclidean distance and genomic distance. Colour by density [A.U]. Mean Euclidean distance per genomic distance is indicated with a line.*

Based on these very promising *in vitro* results, we aimed to further develop this method of tracing the path of multiple-kb linear DNA for use in human cells, with the goal of tracing the 3D path of a chromosomal DNA molecule *in situ* in single nuclei.

ChromoTrace: Computational reconstruction of 3D chromosome configurations for super-resolution microscopy

In order to optimise the design of probe libraries that target the human genome and that can be resolved by 3D fluorescence microscopy, we developed a computer simulation to perform *in silico* modelling of DNA-PAINT FISH experiments in collaboration with Ewan Birney's group at EMBL-EBI. In theory, the limitations in the structural resolution and coverage of the human genome lie in the resolution power of the microscope and the number of unique imager strands that can be used, also referred to as 'colours'. Modern SMLM microscopes reach a 20-by-20-by-20 nm or even better resolution in the x-, y- and z-dimensions. This accuracy of localisation of the SMLM signal is in principle sufficient for 3D DNA structural analysis of B-DNA stretches of less than 100 bp and would represent the only limitation if all target loci could be perfectly labelled and given a different colour. Since only a limited number of exchange rounds/colours are experimentally feasible if one wants to label a large number of loci in the genome, the same colour has to be re-used for multiple loci. When probing an entire chromosome or even the full genome, the correspondence of a spot signal in one colour to the underlying unique genomic DNA sequence has to be derived from the combinatorial pattern of neighbouring spots in different colours. To determine the minimum number of colours required in order to correctly decode this assignment of spots to the DNA sequence of the entire genome, we simulated chromosomal DNA molecules as ensembles of polymer chains in a realistic nuclear geometry and developed an algorithm called ChromoTrace (Barton *et al.*, 2018) to trace DNA paths in 3D multi-colour spot data produced from different FISH probe libraries.

The ChromoTrace algorithm considers the x-y-z-coordinates, the colour of the locus and the expected spatial pattern of colours produced by the known linear genome sequence. The algorithm builds a distance graph of all loci in all colours that could be physically connected assuming maximally extended B-DNA. Then, using a suffix tree search it identifies paths in the distance graph with a single match to an expected

colour sequence from the underlying genome sequence. Once such unique “anchors” to the genome are identified, the algorithm searches for the next expected colour locus in the 3- and 5-prime directions of each anchor path until no more extensions can be made, which works even in highly compact regions (Barton *et al.*, 2018) (Figure 12 A-B). My contribution to this work has been to provide the necessary information about the compaction of the genome, the size of the nucleus and nucleolus as well as giving biological context all over the period necessary for the refinement procedure of the algorithm.

The simulations explored a large parameter space of probe library designs, varying number of colours and genomic resolution. Simulations with a 10 kb genomic resolution resulted in a consensus that beyond 10 colours the length of the paths does not increase much (Figure 12 C) and that the recall is maximal at 0.99 (data shown in Barton *et al*., 2018). The results also revealed that reconstructing the linear path in 3D is very sensitive to wrong colour assignments or off-target locus labelling but can deal relatively robustly with individual missing loci. This theoretical work highlighted the importance to aim for a maximum number of colours (10), very high labelling specificity and as good as possible labelling efficiency in the probe libraries to test experimentally.

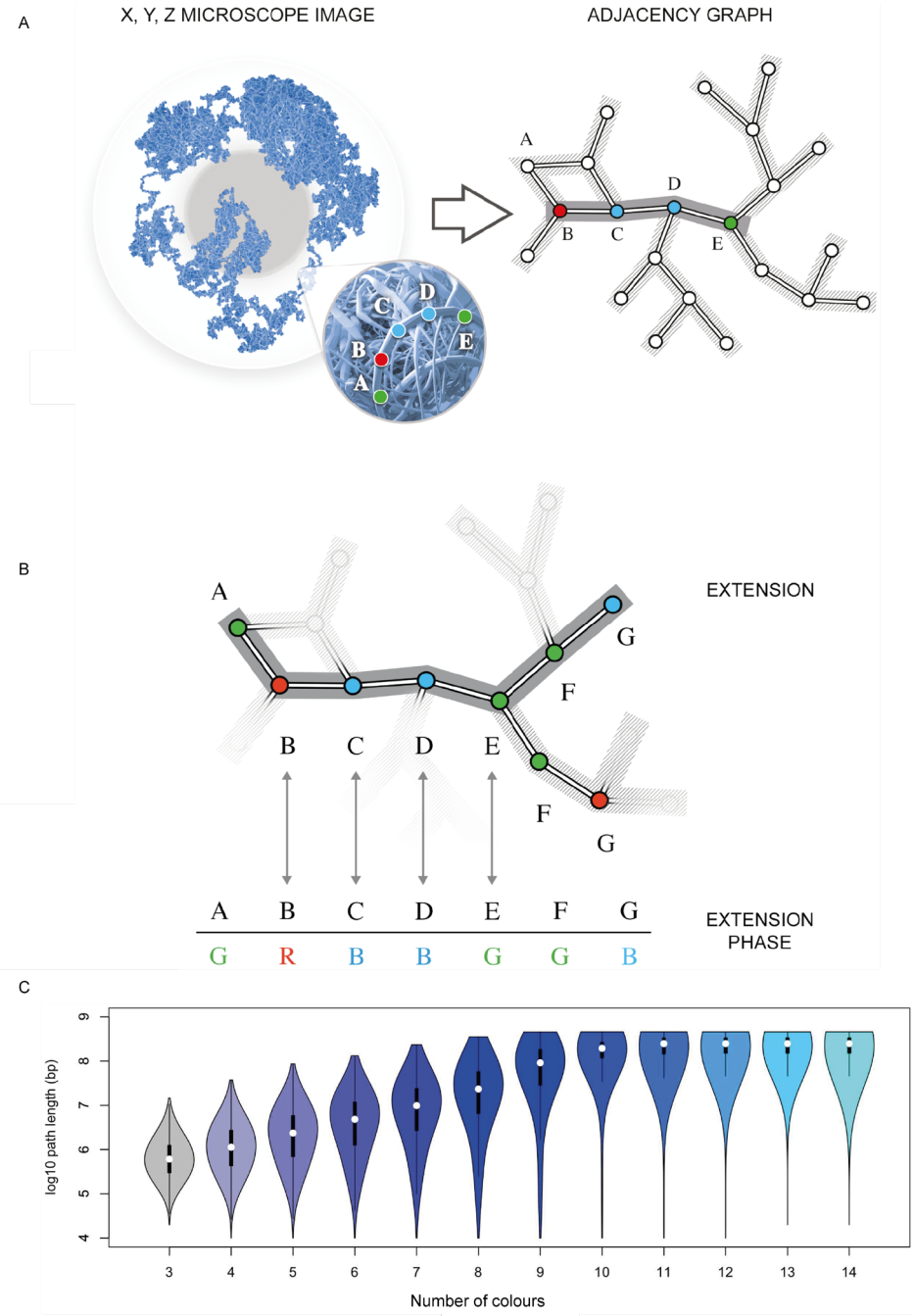

Figure 12: ChromoTrace algorithm. *Figures adopted from (Barton et al., 2018). (A) Illustration of x-y-z microscope image data and how the suffix tree is generated and used to find an unique "anchor". (B) From the "anchors" the genome path can be extended into more compact regions. (C) Violin plot of path length (bp) of simulated FISH data traced by ChromoTrace shows that the length of the reconstructed path does not increase more.*

Experimental optimisation of FISH protocol and probe design for nanoscopic DNA tracing in human cells

Visualisation of human chromosomal DNA molecules *in situ* requires significant development of the FISH probe library and the hybridisation protocol, which were originally designed for the simplified M13 phage DNA *in vitro* system. Specifically, for adherent cells in culture, it was necessary to overcome several issues such as autofluorescence, out-of-focus fluorescence, non-specific signal from off-target binding events and the much higher cell-to-cell variability of the obtained structural data presumably due to biological diversity in structural conformation by epigenetic mechanisms. At the same time my objective was to keep the FISH protocol as mild as possible to maintain genome and nuclear architecture as close as possible to its native state. I therefore systematically optimised the following parameters of the probe design: (i) Number of primary probes per genomic locus, (ii) optimal concentration of primary probes for hybridisation, (iii) optimal length of genome complementary region of primary probes, (iv) optimal temperature for hybridisation, (v) optimal number of dyes per imager strand, and (vi) optimal primary probe library density. Furthermore, I (vii) quantified the hybridisation efficiency by fluorescence correlation spectroscopy (FCS)-calibrated imaging and (viii) evaluated imager strand sequences for multiple exchange rounds.

Combining the resulting optimal test conditions, I used the best suitable combination and set out to perform exchange experiments on a confocal setup to achieve a high throughput in addition to high but diffraction-limited imaging. In doing this, I (ix) showed the feasibility of resolving the path of chromatin *in situ* at 10 kb resolution, (x) ensured that this labelling strategy was compatible with SMLM imaging and DNA-PAINT, and (xi) investigated if fluorogenic imager strands had an increase in fluorescence when bound to docking handles.

(i) Number of primary probes per genomic locus

As classical FISH probes (BAC-FISH) span much larger regions (100-200 kb) than targeted in this study (5 kb), an oligomer-based FISH approach was chosen, for which 10 kb are typically sufficient to visualise a target structure. In oligomer-based FISH probe libraries the size of the labelled region can be precisely controlled by the number of probes that target the region of interest, by the length of the individual probes and the spacing between probes. To minimise the length of DNA labelled within each locus, the minimal number of probes required for reliable detection was determined. As a starting point, a primary probe library was designed to target 10 kb on the MYC 335 enhancer on chromosome 8 using 96 probes made of 82 nt (60 nt complementary to the genome, two gap Ts and 20-nt docking strand). To determine the minimal number of probes needed to reliably detect one genomic locus, a decreasing ratio of imager strands labelled with Alexa Fluor 647 was added to bind the docking handles in HeLa Kyoto cells (HeLa-K) (Figure 13 A). Using an automated 3D spot picking algorithm, the labelled loci were identified and their intensity was determined and used as indicators for the number of primary probes bound (Figure 13 B). The number of spots detected per cell was used as an indicator of locus labelling efficiency. Cells with three or more spots were defined as fully labelled (Figure 13 C). In total, 53% of all cells had the expected three spots or more when all imager strands were labelled with Alexa Fluor 647. This decreased to 25% when only half of the imager strands were labelled. For ten exchange rounds this would mean that only 0.17% of cells would have all loci completely labelled with 96 probes per locus [0.53^10], and a negligible amount of cells would be completely labelled if only 48 probes were used. This suggests that under these hybridisation conditions the probability of observing fully labelled long stretches of chromosomal DNA in cells is very low. We therefore focused next on improving the hybridisation efficiency of our FISH protocol in cells to come closer to the *in vitro* conditions, where only 3 probes were sufficient to reliably detect one locus in a large fraction of M13 mini-chromosomes.

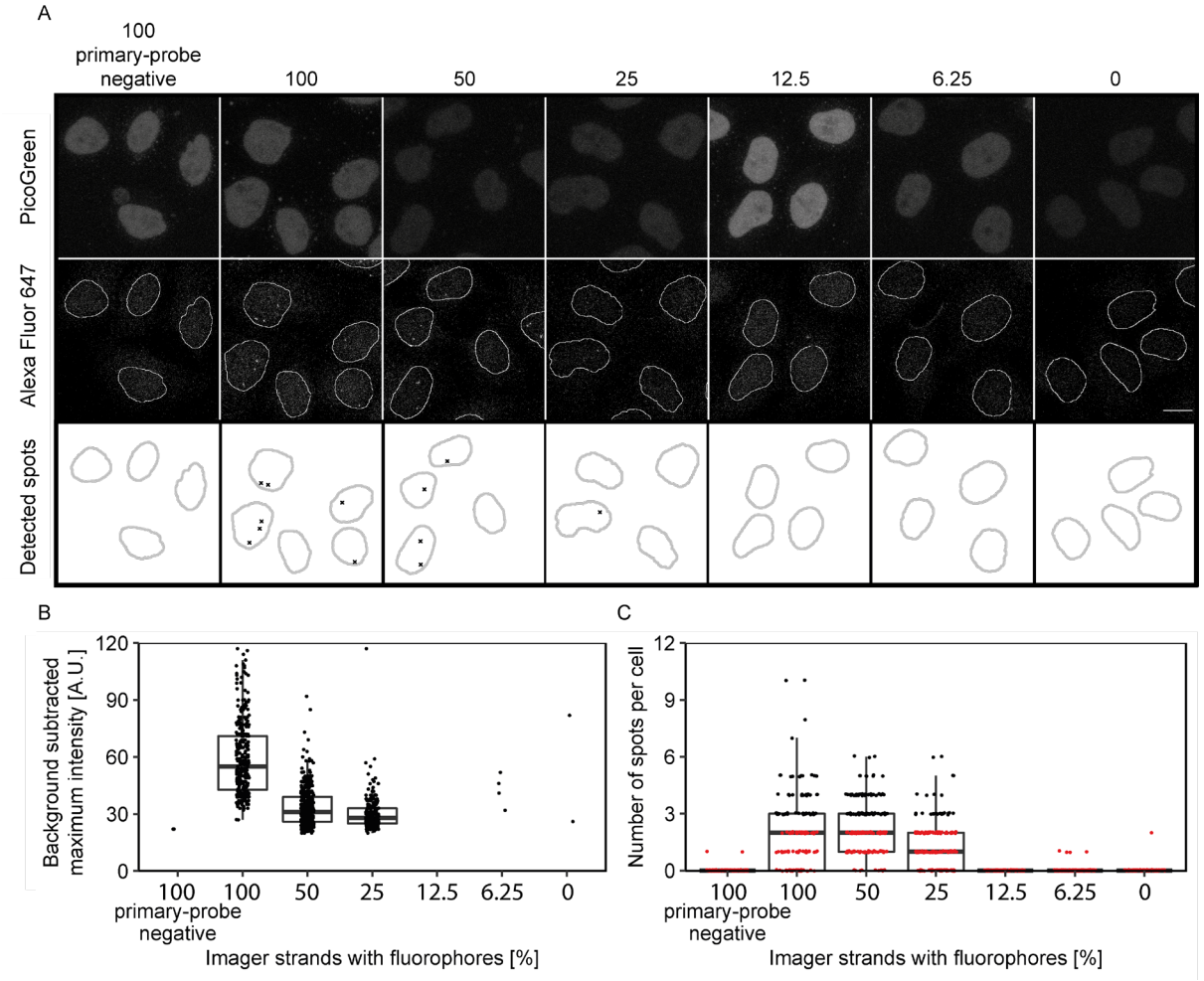

Figure 13: Titration of probe number required to label each locus. *(A) Example images of HeLa-Kyoto cells labelled to target the MYC 335 enhancer with 96 probes. Decreasing percentage of imager strand with fluorophore is indicated. Nuclei are encircled with a white (2nd row)/black (3rd row) nuclear mask. Scale bar, 10 µm. (B) Maximum intensity of detected spots. (C) Number of spots per cell. Black dots, cells with three or more spots; red, cells with less than three spots. X-axis, concentration of primary FISH probes for hybridisation (B, C).*

(ii) Concentration of primary probes for hybridisation

To increase the hybridisation efficiency a series of optimisation steps were performed. We started with increasing the concentration of the primary probes for binding, using 24 probes out of the 96 probes targeting the MYC 335 enhancer and compared the hybridisation efficiency to the standard conditions for 96 probes as a reference (Figure 14 A). Again, we used the mean background subtracted intensity per locus as a measure of hybridisation efficiency. Using 24 probes resulted in ~44% of the intensity of the full 96 probe library at the same concentration (1.9 ng/µl). The intensity

increased by about 5% when probe concentration was doubled. At very high concentrations (228 ng/µl) the unspecific background increased to a level which prevented the detection of any spots (Figure 14 B). However, using the number of spots per cell as an indicator of locus labelling efficiency (Figure 14 C) showed that regardless of the concentration used, fewer spots were observed with 24 than with 96 probes. This suggests that the concentration used for hybridisation cannot compensate for the number of probes per locus.

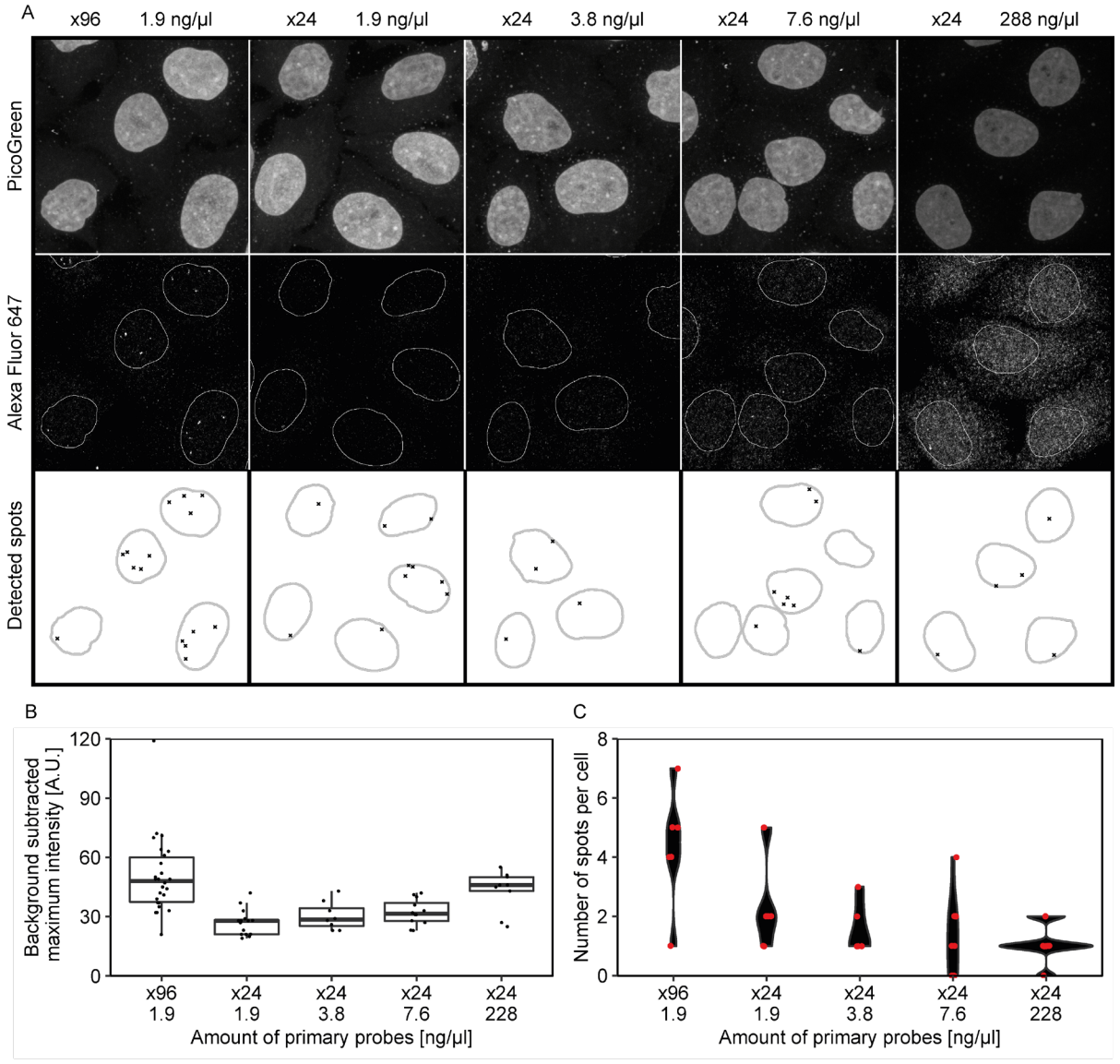

Figure 14: Titration of primary probe library concentration. *(A) Representative images of HeLa-K cells labelled with the MYC 335 enhancer with 96 probes (x96) or 24 probes (x24). Scale bar, 10 µm. Automated nuclei segmentation represented by white masks. (B) Background-subtracted maximum intensity of detected spots. (C) Number of spots per cell.*

(iii) Length of genome complementary region of primary probes

The quality of a FISH probe library is determined by several factors. All the probes in the library need to have complementarity with their targets in the genome. The longer the complementary region is, the stronger the probe will bind to its target and stay bound during stringent washes that remove potential off-target binding to other, not perfectly matching sites in the genome. However, long probes consume a significant genomic length per probe limiting the achievable genomic resolution. Thus, the probes need to strike a compromise to be as short as possible for high genomic resolution while still binding strong enough over off-target binding which needs to be removed with reasonably stringent washes. To investigate if the complementary region could be made shorter than the 60 nt used in standard libraries, three primary probe libraries were designed with either a 60-, 50- or 40-nt-long complementarity to the genome, targeting a 5-kb-long region downstream of the MYC 335 enhancer. All probes had a 20-nt-long docking handle sequence that permanently binds to its imager strand. As a positive control, the MYC 335 promoter library with 96 probes was used. After spot detection and quantification, the probes with reduced length did not show any statistically significant differences, although a significant difference could be observed between the 96 probes in the positive control and the 48 probes in the probe libraries with 60, 50 and 40 bp (Figure 15 A-B 75°C). In summary, the shortest 40-nt-long probe performed similarly to the 60-nt-long probe and can thus be used when it is critical to enable us targeting smaller genomic regions.

(iv) Optimal temperature for hybridisation and length of genome complementarity of primary probe library

Another critical step of *in situ* hybridisation is the optimal melting temperature so that on the one hand the double helix of genomic DNA is effectively opened for hybridisation with the primary probe and on the other hand nuclear architecture is preserved in a reasonably native state. The best melting temperature was optimised together with the length of the genome complementarity. The higher the temperature, the more accessible the genome but the more perturbed the chromatin structure is. Three identical experiments were performed with the only difference being the

hybridisation temperature. Based on the quantification of the hybridisation efficiency, we could show that 65°C was not sufficient to efficiently bind the primary probes since the locus intensities were significantly lower than in both the 70°C and the 75°C dataset. There was a small but less significant increase in the intensities between the 70°C and the 75°C dataset (Figure 15 A-B). Thus, the optimal hybridisation temperature is 75°C when relatively few FISH probes per locus are used, as in this study where high genomic resolution is a key aim. However, when more primary probes per locus can be used, 70°C should also provide a sufficient signal.

In addition, the 50-nt-long probe library had a significantly higher locus intensity in both the 70°C and the 75°C when both data sets are combined to increase statistical power. For this reason, the 50-nt-long primary probe and the 75°C hybridisation temperature were selected for future experiments.

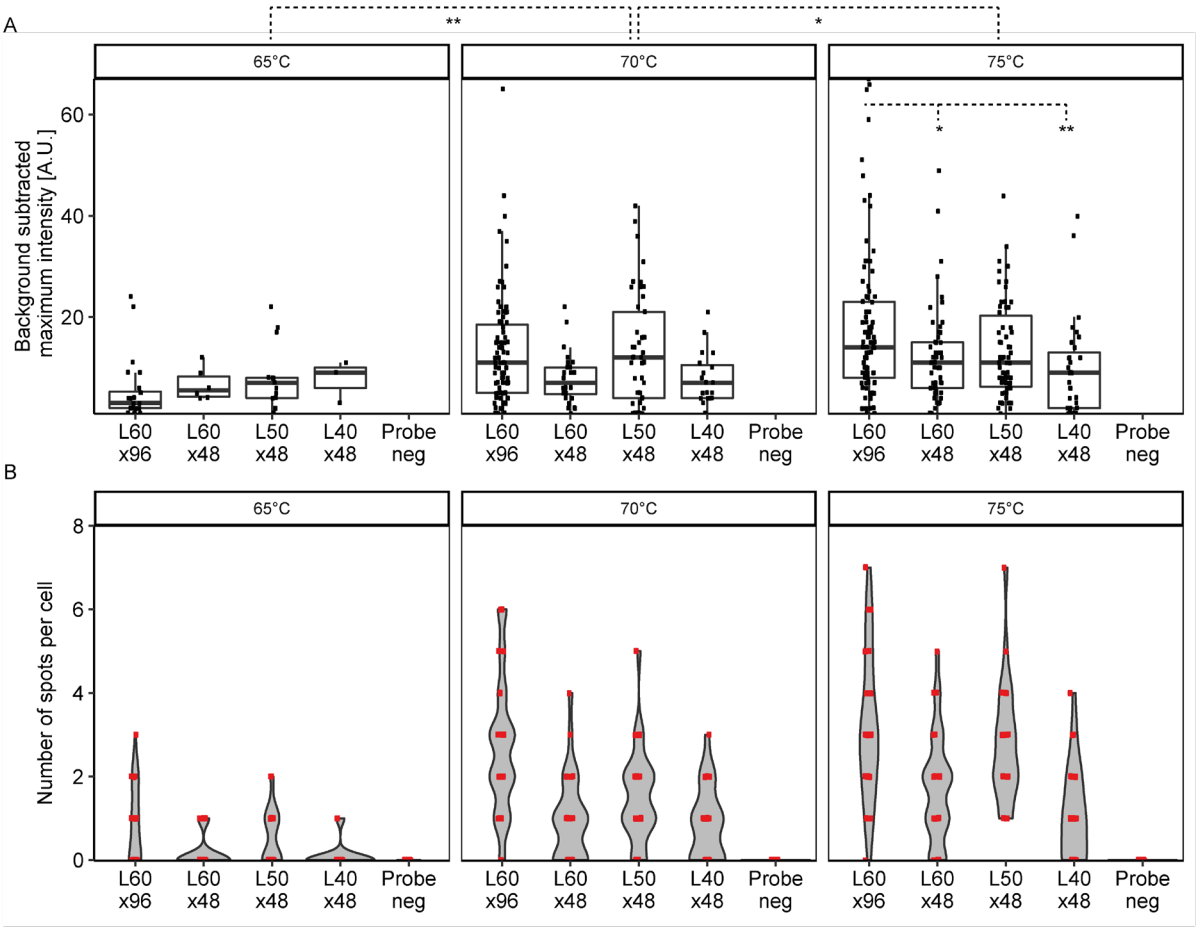

Figure 15: Temperature and primary probe length optimisation on HeLa-K cells. *Length of genome complementarity of primary probe library (L) was 60, 50 or 40 nt (L60, L50, L40, respectively). 96 or 48 probes per locus were used (x96 or x48, respectively). Samples were hybridised at 65°C, 75°C or 60°C. (A) Background-subtracted maximum intensity of detected spots. Statistics indicated with dotted lines were performed on all-against-all for 75°C; between 70°C and 75°C and comparing all exchanges to E_0 with 1-way-anova and a Tukey's 'honestly significant difference' post hoc test (* P ≤ 0.05; ** P ≤ 0.01; *** P ≤ 0.001). (B) Number of spots per cell.*

(v) Number of dyes per imager strand

Another option to increase signal over background per locus is to add more than one fluorophore to the imager strands. Adding a second fluorescent dye per imager strand could in principle increase the signal two-fold but would also increase the intensity of potential off-target binding events and might affect the binding affinity of the imager strands. In addition to the triploid and karyotypically variable HeLa-K cancer cells, diploid RPE-1 telomerase-immortalised cells were used to have a clear expectation for the number of loci present per cell. The MYC locus was then hybridised with 48 primary probes at the standard concentration and imager strands containing either a

single 5'-linked Atto565 dye (1x) or an imager strand with two 3'- and 5'-linked Atto565 dyes (2x) were bound to the primary library and compared to primary probe-negative controls. A significant intensity increase of 2.3 and 2.4 was detected comparing the 1x and 2x samples both for RPE-1 and HeLa-K, respectively (Figure 16 A-B). Although being slightly higher than the expected two-fold increase, this difference could be explained by the variability in intensities. The number of spots did however not show a significant change between the conditions, although no cells without spots were observed in the 2x conditions. In summary, adding an extra fluorophore to the imager strands is very advantageous and doubles the fluorescence intensity of the FISH signal. In a hypothetical case where only a few primary probes are bound to their targets and the 1x imager strand does not give a signal beyond the detection threshold, a 2x imager strand might introduce the extra photons required for detection. Thus, in theory, the same signal intensity could be reached with 48 primary probes as if a primary probe library of 96 probes was used. Importantly, this approach is only advantageous when the excess imager strand can be washed stringently away as background fluorescence also increases two-fold when using the 2x imager strand. A similar experiment was performed for DNA-PAINT with a reported intensity gain of 1.6, but the signal gain in intensity was exactly cancelled out by an equal increase in background signal (Pers. Comm., Ralf Jungmann (MPI, Martinsried, Germany)).

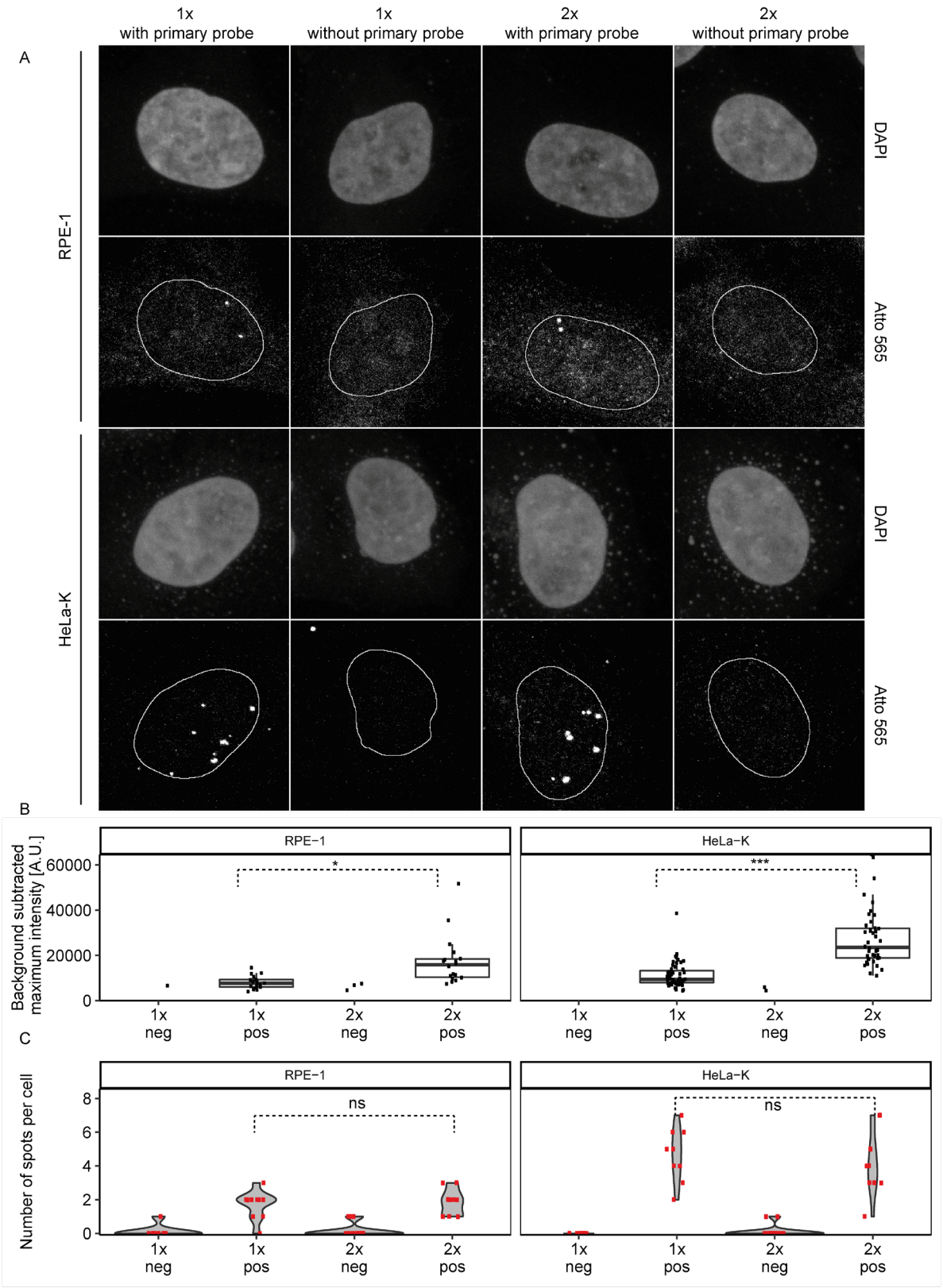

Figure 16: Comparison between imager strands with one or two conjugated fluorophores. *(A) Representative images of MYC locus targeted with 48 primary probes in RPE-1 and HeLa-K cells. Automated nuclei segmentation represented by white masks. Scale bar 10 µm. (B) Background-subtracted maximum intensity of detected spots. (C) Number of spots per cell in RPE-1 and HeLa-K cells. Statistics*

indicated with dotted lines between 1x and 2x positive controls using a 1-way-anova and a Tukey's 'honestly significant difference' post hoc test (P ≤ 0.05; ** P ≤ 0.01; *** P ≤ 0.001).*

(vi) Highest primary probe library density across large genomic distances

Having optimised the number and length of probes as well as the number of fluorophores per probe, we next assessed the probe density which could be achieved on the genome. Here, we aimed at identifying the densest genomic probe spacing possible while keeping the highest labelling specificity in line with the optimal conditions identified above (e.g. 40- or 50-nt-long genome complementarity, at least 48 probes per loci). To this end, a probe design algorithm was developed that optimises melting temperature (Tm) differences between all possible probe sequences in the unique target and potential off-target regions in the human reference genome (GRCh38) and confirms the presence of the selected probes in the genome of the human cell lines we employed experimentally (e.g. HeLa-K or RPE-1). The pipeline developed in collaboration with Carl Barton (Ewan Birney's group, EBI, Hinxton, UK) takes all possible probes with a sliding window of one base and searches for putative off-targets in the genome. In case off-targets are found, the Tmis computed and used to filter probes unsuitable for FISH. A threshold of maximally 60°C off-target melting temperature corresponding to the washing step following the primary hybridisation was used.

The algorithm runs on the premise that 48 primary probes with 50- or 40-nt-genome complementary regions provide sufficient signal over the background to reliably detect all alleles of a locus in about 50% of cells. To calculate the genomic resolution we could achieve with a probe set of 48 targeting one locus with the same docking handle for labelling, we surveyed the genomic region of the TAD containing the MYC gene, for which well-established FISH probes against the MYC 335 enhancer were already available (Figure 17 A). From the pre-computed table of potential probe sequences that satisfy a maximum off-target Tm threshold of 60°C, have less than 30% C- or G-content and lack repetitive sequences, we selected a subset of 48 probes resulting in a maximum probe density (minimal neighbour distance 21 bp) and lowest possible off-target Tm. This allowed us to reduce the genomic length of each locus detected by a

probe set with the same docking handle sequence for DNA-PAINT labelling to less than 5 kb. There was a clear difference between the Tm of the off-targets and the intended targets (Figure 17 B). The probe libraries computed can thus contain 48 or more probes in one locus matching these criteria per 5-kb-stretch of genomic DNA (in 95% of the loci in the region that was surveyed). This decreases for regions of 4 kb and 3.6 kb. We therefore concluded that 5 kb is a good compromise between high genomic resolution and a sufficient number of probes in the set targeting one locus for reliable labelling.

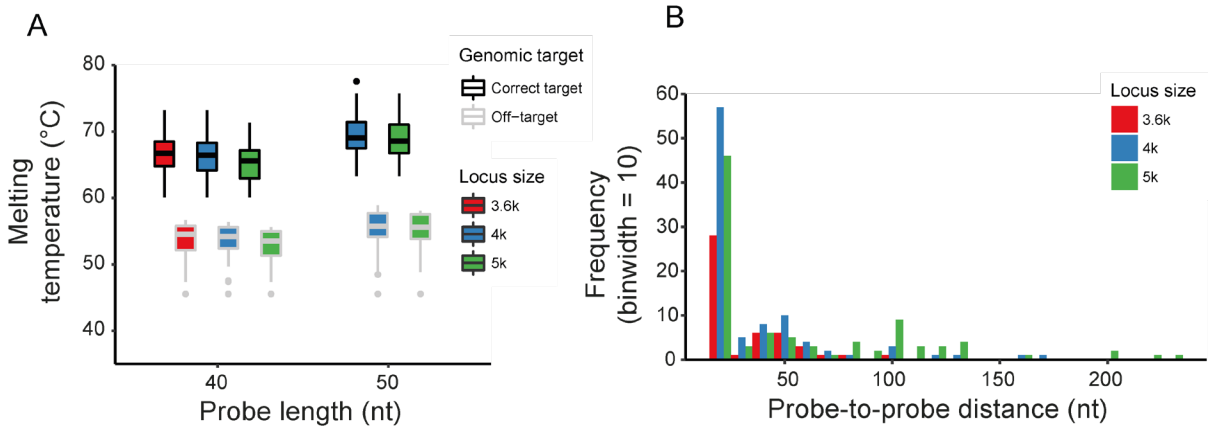

Figure 17: In silico evaluation of probe library design algorithm. *(A) Melting temperature plot of potential E_1-E_{10} probe library designs where (Red) is a locus size of 3.6 kb, (Blue) is a locus size of 4 kb, and (Green) is a locus size of 5 kb. (B) Frequency plot of probe-to-probe distance in the same library as in B.*

(vii) Quantitating the hybridisation efficiency by FCS-calibrated imaging

The overall much lower locus labelling efficiency in cells than *in vitro* suggested that the primary probe hybridisation efficiency is rather low since the intensity of imager strands targeting a tertiary imager strand (imager strand that binds an unlabelled docking handle that again is hybridised to the primary probe) essentially gives the same signal intensity as secondary imager strands (an imager strand is hybridised to the docking handle of the primary probe) (Pers. Comm., Franziska Kundel, EMBL, Heidelberg).

I set out to determine the number of fluorophores bound to one locus in order to be able to calculate the absolute hybridisation efficiency. To this end, we made use of

FCS-calibrated imaging that measures the concentration and number of molecules per imaged voxel in a confocal image stack can be calculated based on calibration with the same dye in solution (Politi *et al.*, 2018). Upon 100% hybridisation efficiency, a locus labelled with 96 or 48 primary probes that are all bound by single dye-linked imager strands should contain the same number of fluorophores. After ensuring linear correspondence between intensity and fluorophore concentration of our FCS system (Figure 18 A), we however detected a very variable range between 1-50% of the expected dyes per locus with a mean of only ~5% (Figure 18 B), suggesting a rather low hybridisation efficiency in human cells *in situ*. This is presumably due to the low accessibility of the target DNA sequences in chromatin and is challenging to improve without destroying nuclear architecture by harsh denaturing treatments. Nevertheless, hybridisation efficiency will be a key parameter to improve further in future developments of FISH-based technologies (see discussion).

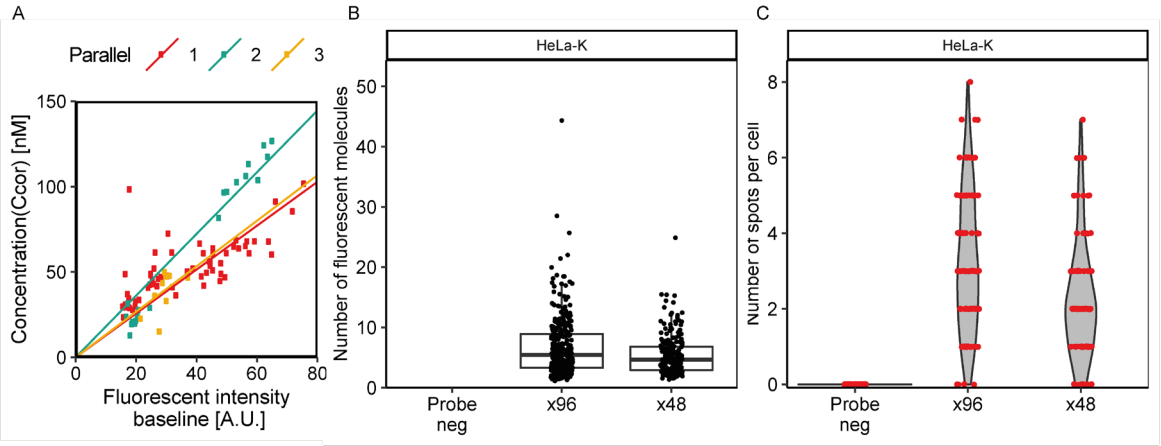

Figure 18: FCS-calibrated imaging of HeLa-K cells. *Labelled with 96 (96x) or 48 (48x) primary probes with 20-nt-long permanently bound imager strands conjugated to AlexaFluor 647. (A) Correspondence between intensity and fluorophore concentration. (B) Number of fluorophores per loci. (C) Number of spots per cell.*

(viii) Evaluation of imager strand sequences for multiple exchange rounds

To extend our method to 10 colours, we designed 10 distinct bifunctional 12-nt-long docking handles (E_1-E_{10} docking handles) that can be used either for transient binding of 9-nt-long imager strands (E_1-E_{10} imager strands) for canonical DNA-PAINT imaging where blinking is induced by short-lived binding events or for long-lived binding of 12-nt-long imager strands for confocal-, STORM- or STED-based detection exploiting the photophysical properties of dyes bound stably.

To perform sequential imaging of multiple loci, imager strands have to be efficiently washed away before each new imaging round without removing the primary probes hybridised to the genome. We achieved this by using 12-bp-long docking-handle-imager-strand complexes with an estimated free energy of -17.8 ± 0.5 kcal/mol under imaging conditions. For comparison, the 20-nt-long imager strand used in the earlier optimisations has an estimated free energy of -29.15 kcal/mol when in complex with its docking handle. In a BLAST search for the different 12-nt-long sequences of the imager strands no hits were found in the human reference genome GRCh38. These imager strands are estimated to have a docking time of ~1000 s, sufficiently long for automated acquisition of several labelled cells on a confocal microscope. To ensure that every imager strand could be washed away after imaging and did not produce unspecific background signal in the absence of primary probes, we performed a first experiment to test ten exchanges.

To this end, we targeted ten loci within the large MYC gene in RPE-1 cells, namely between the MYC promoter and the MYC 335 enhancer, and performed the optimised FISH protocol. When no primary probe was present, 8/10 imager strands (E_1-E_5 and E_7-E_9) gave little or no signal. 9/10 imager strands (E_1-E_9) produced spot-like signal in the presence of the primary probe library (Figure 19 A). As desired, all detected spots disappeared after a 25% formamide (FA) wash (in 1x PBS) (Figure 19). This showed that most of these imager strands are good candidates for sequential imaging, with only two exceptions. E_{10} shows very high nuclear background both in the absence and the presence of the primary probe library. Spot-like signals more intense than the background are also present without the primary probe and thus unspecific. E_6 shows low background but spot-like signals appear in the negative control similar to E_{10}. In

conclusion, all 10 FISH probes were efficiently washed away by the formamide buffer. 8/10 FISH probes gave specific, spot-like signals only in the presence of the primary probe library, while two probes showed unspecific spot-like signals. This is potentially due to binding of these probes to highly repetitive regions which are absent from the reference genome assembly GRCh38 used for BLAST searches to exclude off-target binding of the imager strand sequences.

The nuclei (Hoechst channel) of the cells shown in Figure 19 show clear signs of having undergone a FISH protocol,. Nevertheless, they have defined borders around the nucleus and visible nuclear compartments showing that these permeabilised cells have been successfully crosslinked before the harsh treatment during the FISH protocol. There is not a lot of DNA-signal outside the nucleus showing that this protocol does not disrupt the nuclear integrity as some harsh treatments may do (data not shown). In some cells one or two regions have bright spots indicating highly dense/accessible DNA which might correspond to collapsed regions of DNA.

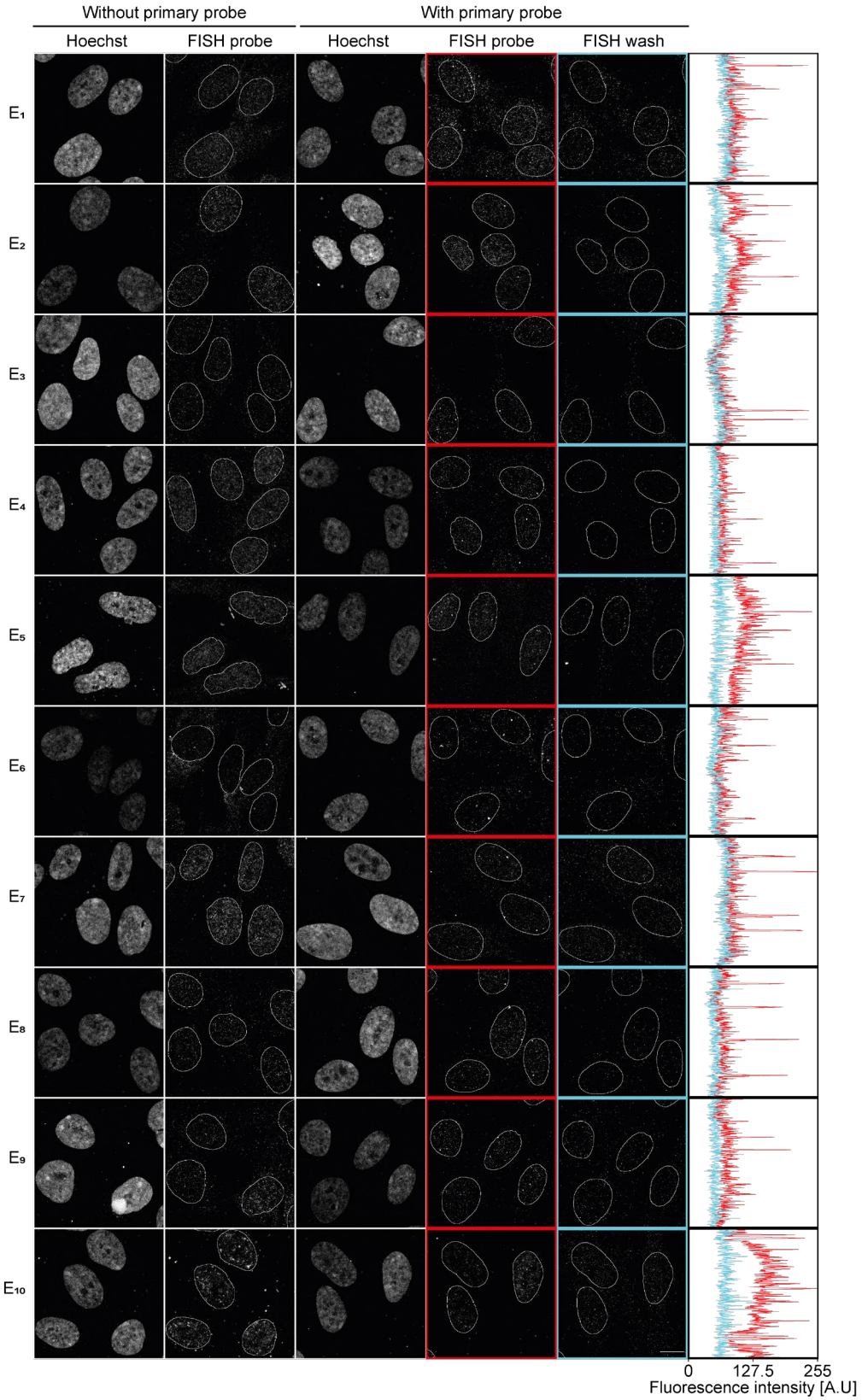

Figure 19: Control of imager strands binding in the presence or absence of the primary probe library. *(Red) Primary probe labelled with Atto565 imager strand. (Cyan) Washout of imager strand with 25% FA. (Plots) Column sums of red and cyan*

images. Brightness and contrast were linearly adjusted for display purposes. Scale bar, 10 µm.

To quantify our labelling efficiency for the targeted genomic locus, spots were automatically segmented in the absence or the presence of the primary probe as well as after the FA wash. The number of detected loci per cell was determined and plotted (Figure 20 A). Quantification confirms what was observed qualitatively in Figure 19 and shows that 8 imager strands detect spots and furthermore that E_6 and E_{10} label non-specific sequences even in the absence of the primary probe library (Figure 20 A). In the diploid RPE-1 cells two copies of each locus are expected, however, only a single spot is detected in most cells (Figure 20 B). This observation is similar to the results obtained with the 48 probes and the 20-nt-long permanently bound imager strand and is therefore likely caused by the poor hybridisation efficiency of the primary probe. Interestingly, this may indicate that the two genomic alleles of the same locus differ in their accessibility.

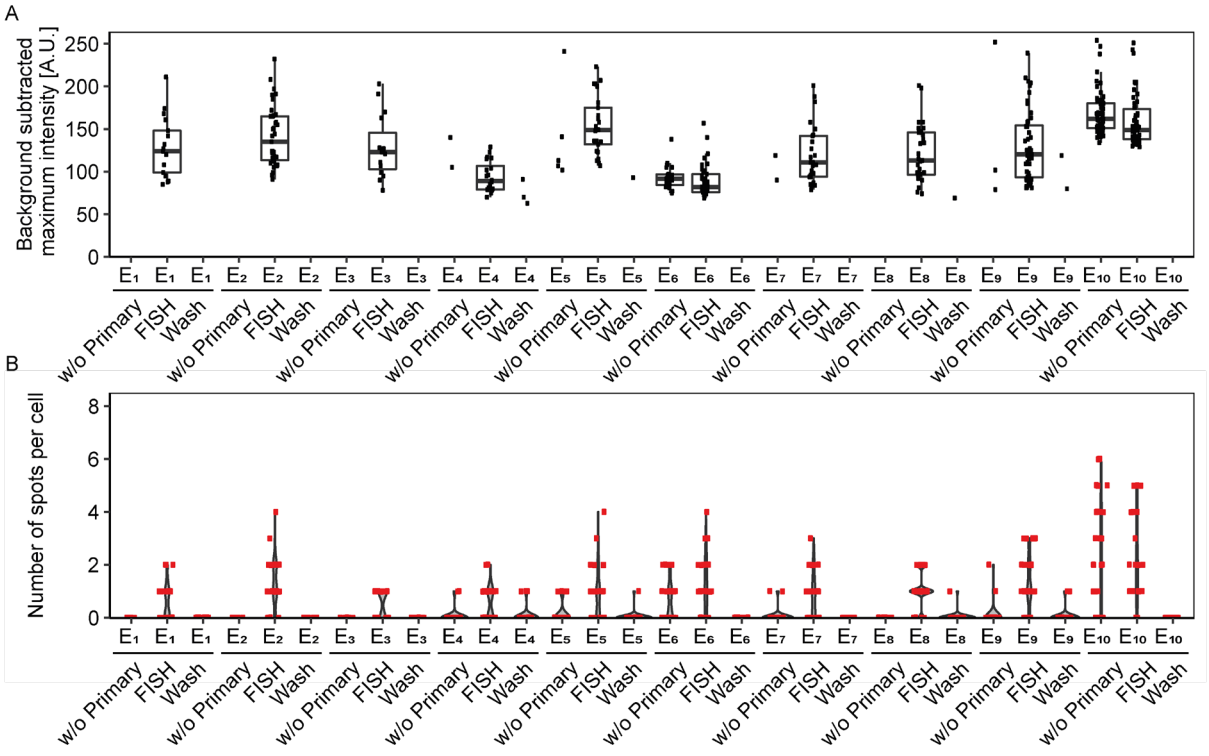

Figure 20: Quantification of imager stands E_1-E_{10} binding in the presence or absence of the primary probe library. *(A) Background-subtracted maximum intensity of detected spots. (B) Number of spots per cell.*

(ix) Resolving the path of chromatin *in situ* at 10 kb resolution

Given that 8/10 of our first set of imager strand designs proved suitable for systematic exchange labelling, we acquired a larger data set to test in how many cells we would be able to completely label one allele of the target MYC locus with 10 probes and elucidate its 3D path. To this end, we hybridised our probe library that targets 10 loci between the MYC promoter and the MYC 335 enhancer to the genome of RPE-1 cells and imaged 20 different fields of view with approximately 5 cells each in 10 consecutive imager strand exchange rounds resulting in a total dataset of 158 RPE-1 cells (Figure 21). For each exchange round and field of view 3D drift was corrected and labelled loci were detected (Figure 22). As expected for targeting a single gene locus with a total genomic length of 90 kb, labelled loci re-appeared in close spatial proximity across the exchanges (Figure 22). Considering the quantification of hybridisation efficiency (Figure 23 A-B), one can appreciate some variability in the background-subtracted maximum intensity between different imager strands but no clear trend of decay or increase over time during the 10 rounds of exchanges and imaging. This suggests that the repeated washes and prolonged imaging do not affect the pre-bound primary probe library or the ability of the imager strands to bind the docking handles even after several hours of working on the same sample. It would rather point to some degree of variability in the hybridisation efficiency of each of the primary probe sets, which may be caused by the degree of compaction of the underlying targeted locus. On average, 24% of the cells have exactly the expected number of spots (two) in the different exchanges, and 39% of them have two or more spots. This is in the same range as for the 20-nt-long docking handle.

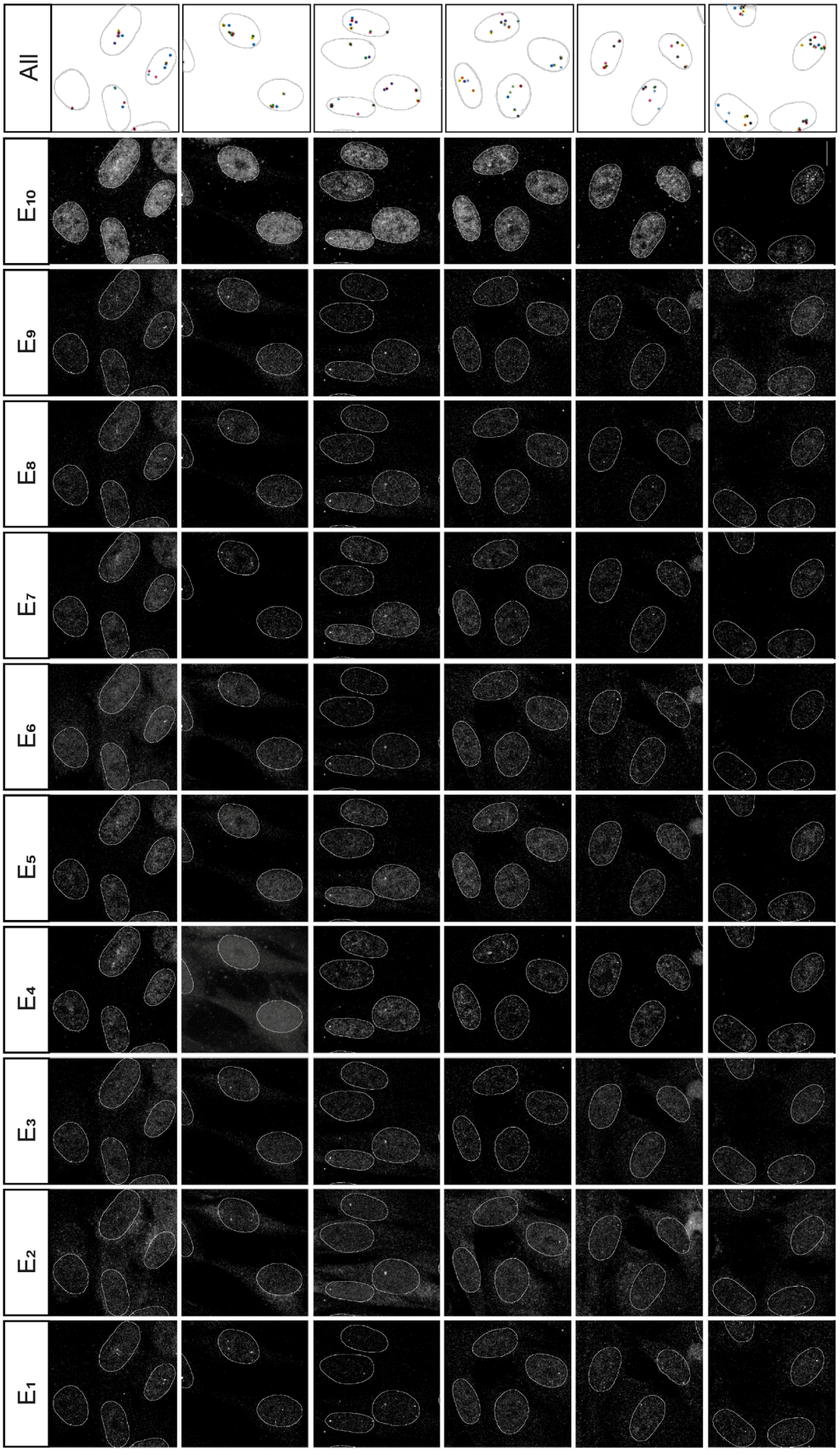

Figure 21: Representative images of individually contrasted RPE-1 cells. *Cells labelled at 10 loci near MYC 335 enhancer and imaged with 12-nt-long Atto565 bearing E_1-E_{10} imager strands. Imaged on a confocal microscope. Scale bar, 10 µm.*

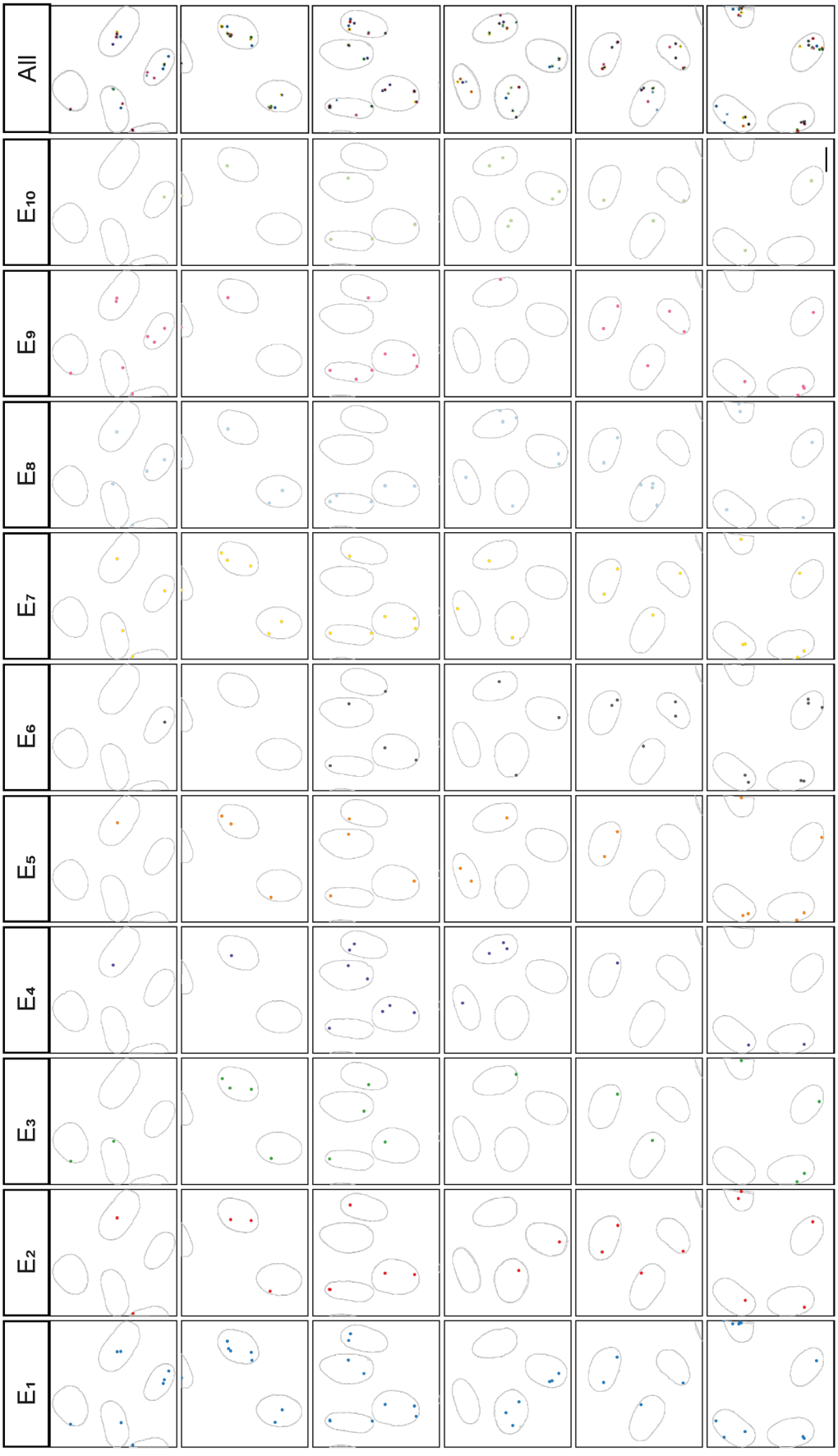

Figure 22: Detected spots from images in Figure 21.

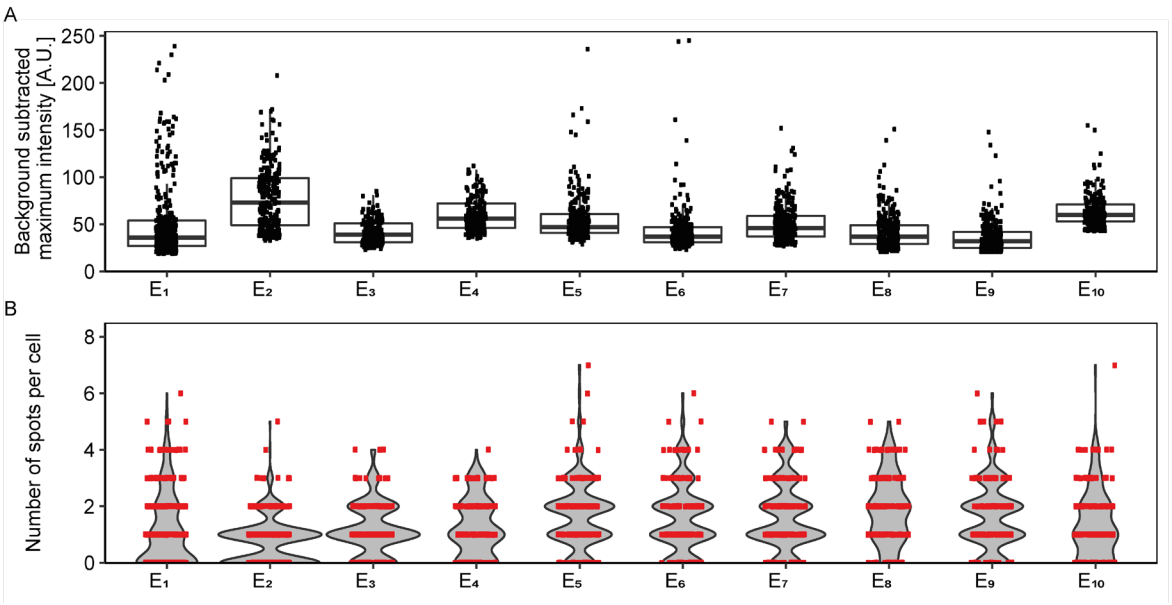

Figure 23: Quantification of spot detection from images in Figure 21. *(A) Background-subtracted maximum intensity. (B) Number of spots per cell.*

To extract the connections between individually labelled loci belonging to one allele of the MYC gene, we used a simplified version of ChromoTrace. This first identified genomically neighbouring spots that were located close to each other within less than the 3D distance expected for 10 kb of DNA as a linear 11-nm nucleosome fibre, which is consistent with them residing on the same chromosomal DNA molecule. In the rare cases where several close neighbours were found, the closer one was chosen. The algorithm then connected all labelled loci that satisfied this single allele criterion into traces (Figure 24). For about 10% of the cells (9/96), one allele of the targeted MYC region could be detected in all exchanges. Given that in diploid RPE-1 cells we would have expected two alleles per cell, this corresponds to an approximate overall complete labelling efficiency of 5% of the imaged cells. In addition, a total of 15, 4, and 22 traces could be found with a length of 9, 8 and 7 spots per trace, respectively, from which structural information about a significant part of the locus can be extracted. Even the very short traces in Figure 24 (two or three connected loci) provide some information, as they indicate that the beginning of the genomic region (E_1-E_3) is often not connected with the end (E_4-E_{10}), suggesting poor labelling efficiency of the in-between loci.

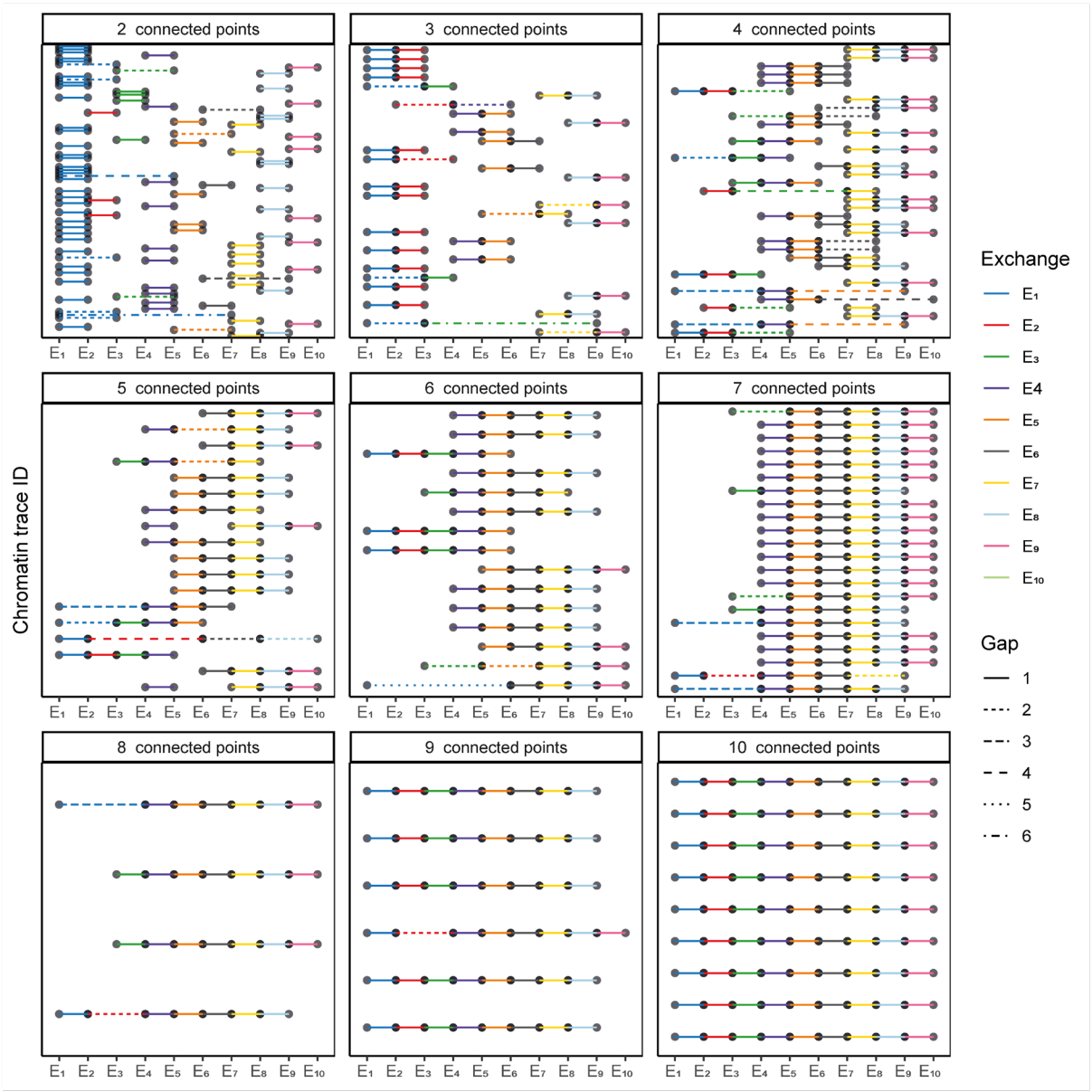

Figure 24: Traces of connected points from 96 RPE-1 cells with exchange round E_1-E_{10}. *Points were connected with a distance threshold of 850 nm per 10 kb stretch (10 000 bp * 0.34 nm/bp / 3.8 compaction = 895).*

The 90-kb-long genomic target region is situated in the middle between the MYC promoter and the MYC 335 enhancer and covers about 1/3 of the distance between them with a probe set targeting a stretch of 5 kb, spaced equidistantly every 10 kb. From Hi-C data in RPE-1 cells it is predicted to cover about one third of the distance between the MYC promoter and the MYC 335 enhancer (GRCh37: chr8q24.21: 128414228-128414237)). Plotting the 2D path of the nine completely labelled alleles in Figure 25 A suggests that this region is rather flexible, as it does not appear to have a reproducible distance signature present in all traces. Comparing the 3D path of this genomic region between individual cells confirms this flexibility (Figure 20 B). To

estimate the diameter of the volume that this target region typically occupies, we ran a principal component analysis and extracted the first eigenvalue (how extended the point cloud is in the longest direction) which was 200 ± 100 nm for the 9 completely labelled traces. Notably, the second and third eigenvalues were about half of the first eigenvalue. By inspection of the volumes this aligns with larger variability in the z-direction. The average of all eigenvalues was 140 nm which is similar to the 150 nm measured feret diameter of single replication domains (Xiang *et al.*, 2018).

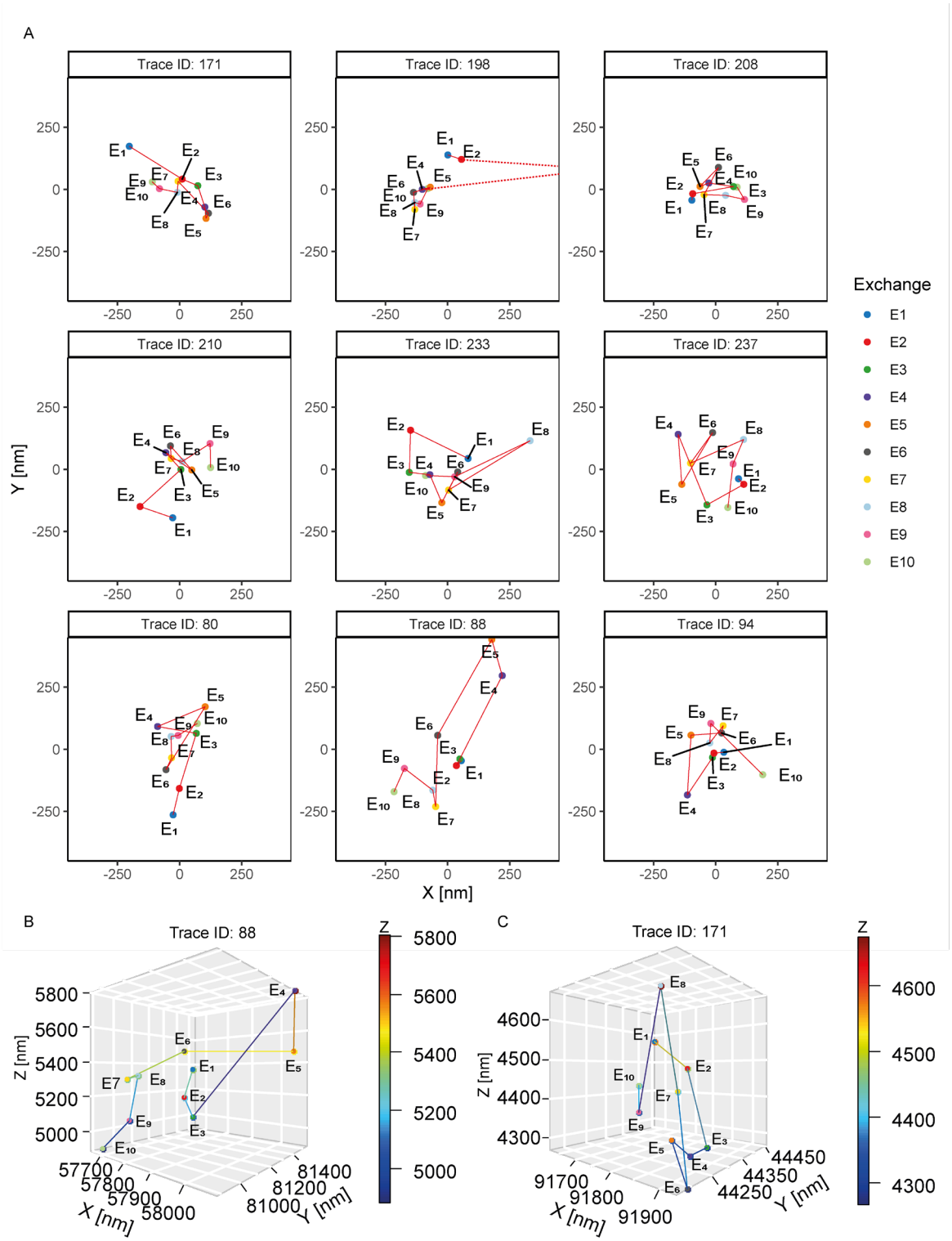

Figure 25: Traces of the loop region between MYC promoter and MYC 335 enhancer in RPE-1 cells. *(A) All 9 traces with signal from all 10 Exchanges. (B-C) 3D rendering of Trace 88 and 171 from A.*

To investigate whether there is any systematic difference in compaction of the 10 kb domains between the individual labelled loci, all pairwise distances between neighbouring loci were computed. Globally, all 10-kb-spaced loci extended on average

250 ± 170 nm which is only 28% of a completely stretched out 11-nm fibre [250/(10000*0.34/3.8)]. E_2-E_3 and E_3-E_4 appear to be slightly further apart but are not significantly so in a 1-way-annova with a Tukey's 'honestly significant difference' post hoc test. They do have a smaller sample size which may explain this divergence from the mean (Figure 26).

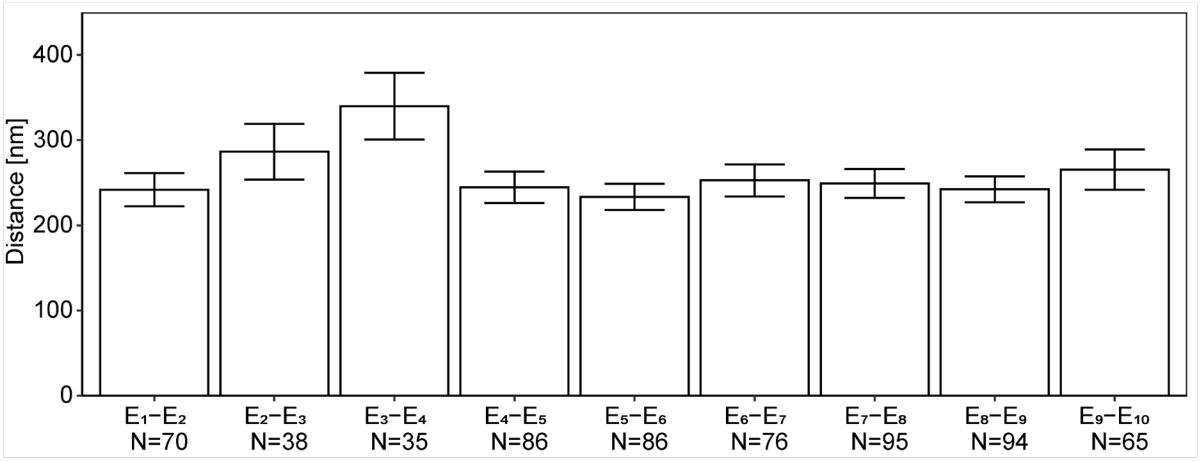

Figure 26: Mean distance between detected loci. *Error bar, standard error of mean. N, sample size per group. Calculated from all traces found in 76 cells. No significant difference according to a 1-way-anova and a Tukey's 'honestly significant difference' post hoc test (* P ≤ 0.05; ** P ≤ 0.01; *** P ≤ 0.001).*

Overall, we can conclude several new aspects of the structure of a 90 kb genomic region spanning part of the MYC gene locus from this data. The region is relatively compact filling a volume of about 0.002 μm^3, which is similar to the current best estimates of a single replication domain equivalent to a TAD (Xiang *et al*., 2018). Its individual nine 10 kb long segments all have a very similar average physical length, arguing against a specific prominent region that is much more or less compact. Finally, its 2D and 3D looping structure appears highly variable between individual cells, indicating a very flexible state of this region in RPE-1 cells.

No clear structures observed in unstructured region with 10 kb genomic resolution

To explore if any structures that could not be seen visually in the images above are present, the distances between all points of each connected path was computed. In the resulting distance map it can clearly be seen that the major contribution to the distance between two loci is the genomic distance between them (Figure 27 A). No

interactions between points more than 3 exchanges apart (corresponds to 20 kb genomic distance) were common enough to reduce the distance between the points below ~200 nm.

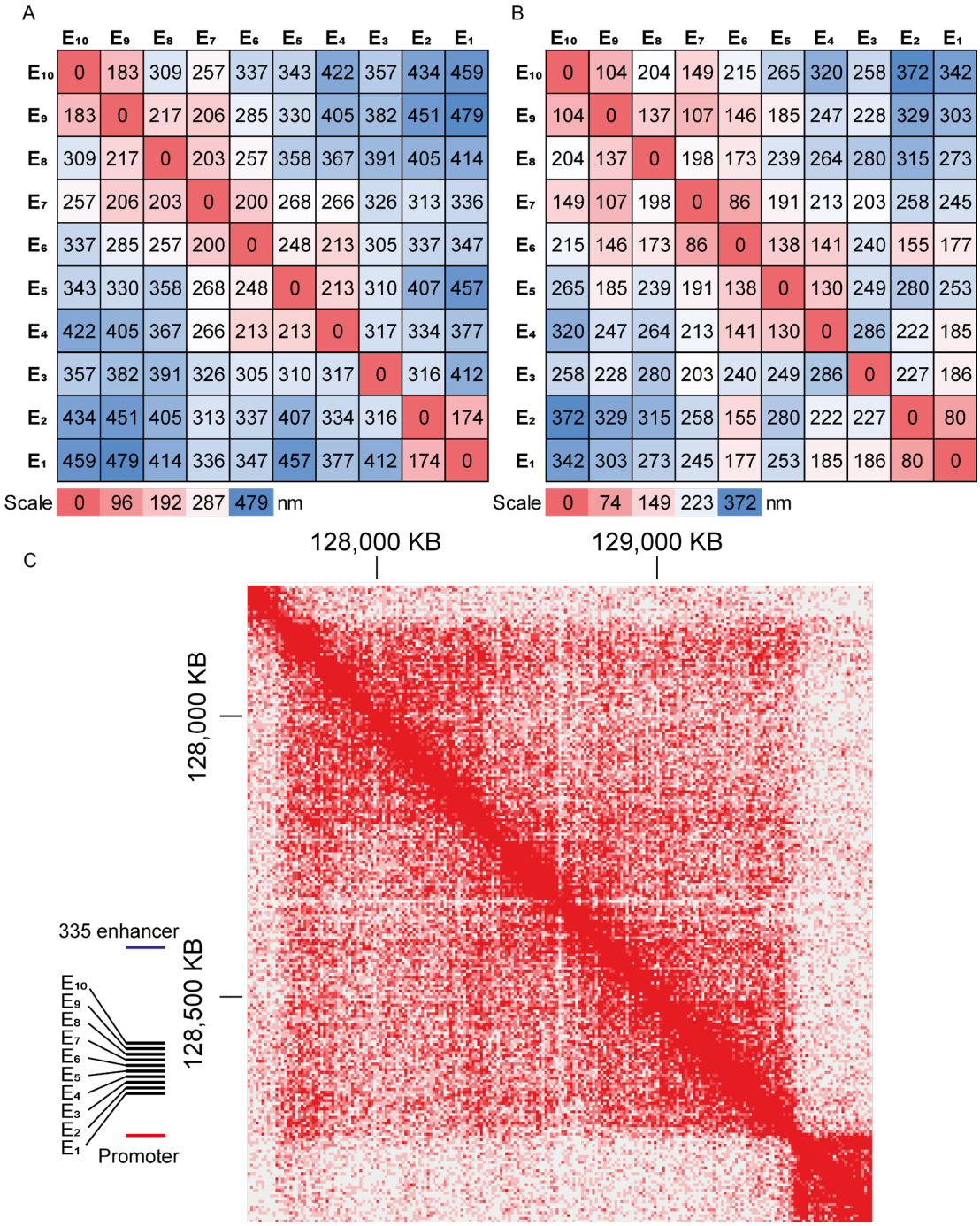

Figure 27: Heatmap of ten loci-to-loci distances between the MYC promoter and enhancer. *(A) Mean Euclidean distances between all detected loci (B) Standard deviation of the distances between all detected loci. (C) Hi-C map from (Darrow et al.,*

2016) *labelled with the 10 E_1-E_{10} probe loci (Black), Promoter (Red), and 335 enhancer (Blue). Hi-C rendering with Juicebox 1.11.08 (Durand et al., 2016).*

(x) Secondary imager strand is compatible with SMLM techniques such as STORM and DNA-PAINT

To ensure that the general labelling strategy is compatible with super-resolution microscopy, we used the MYC 335 enhancer reference probe library that spans 10 kb with 96 genomic probes and hybridised it in HeLa cells. Permanently bound imager strands (20 nt) bearing an Alexa Fluor 647 dye that has good blinking properties for STORM imaging was bound to the docking handles and imaged on a SMLM microscope (in collaboration with Jonas Ries, EMBL, Heidelberg). Qualitatively, these proof-of-concept experiments show that specific loci are labelled and that sub-diffraction structures can be resolved within each diffraction-limited locus (Figure 28 A, B).

We also validated that correlative diffraction-limited and super-resolution experiments can be made by using the bi-functional 12-nt-long docking handle, to which both intermediate binding 12- and short binding 9-nt-long PAINT imager strands can be bound. We hybridised 48 probes with the E_1 docking handle targeting the MYC 335 enhancer and first added the 12-nt-long Atto565-conjugated imager strand to the sample to acquire diffraction-limited images using a spinning disk confocal microscope. Subsequently, the 12-nt-long imager strand was washed away and a 9-nt-long imager strand was added. PAINT imaging was performed by keeping an excess of imager strands in the imaging media during acquisition of a high-speed video that captures the binding and dissociation of many imager strands. A composite of these two images show that the diffraction-limited signal decomposes into several super-resolved DNA-PAINT signals. However, there is a very significant amount of background in the DNA-PAINT channel (Figure 28 C-E) which will have to be overcome to rigorously interpret the super-resolved localisations. Nevertheless, we can conclude that correlative confocal and PAINT imaging can be performed with a single bi-functional docking handle, which will be very useful for applications where large volumes have to be scanned quickly in low resolution mode before zooming in on high resolution reconstruction of particular loci of interest.

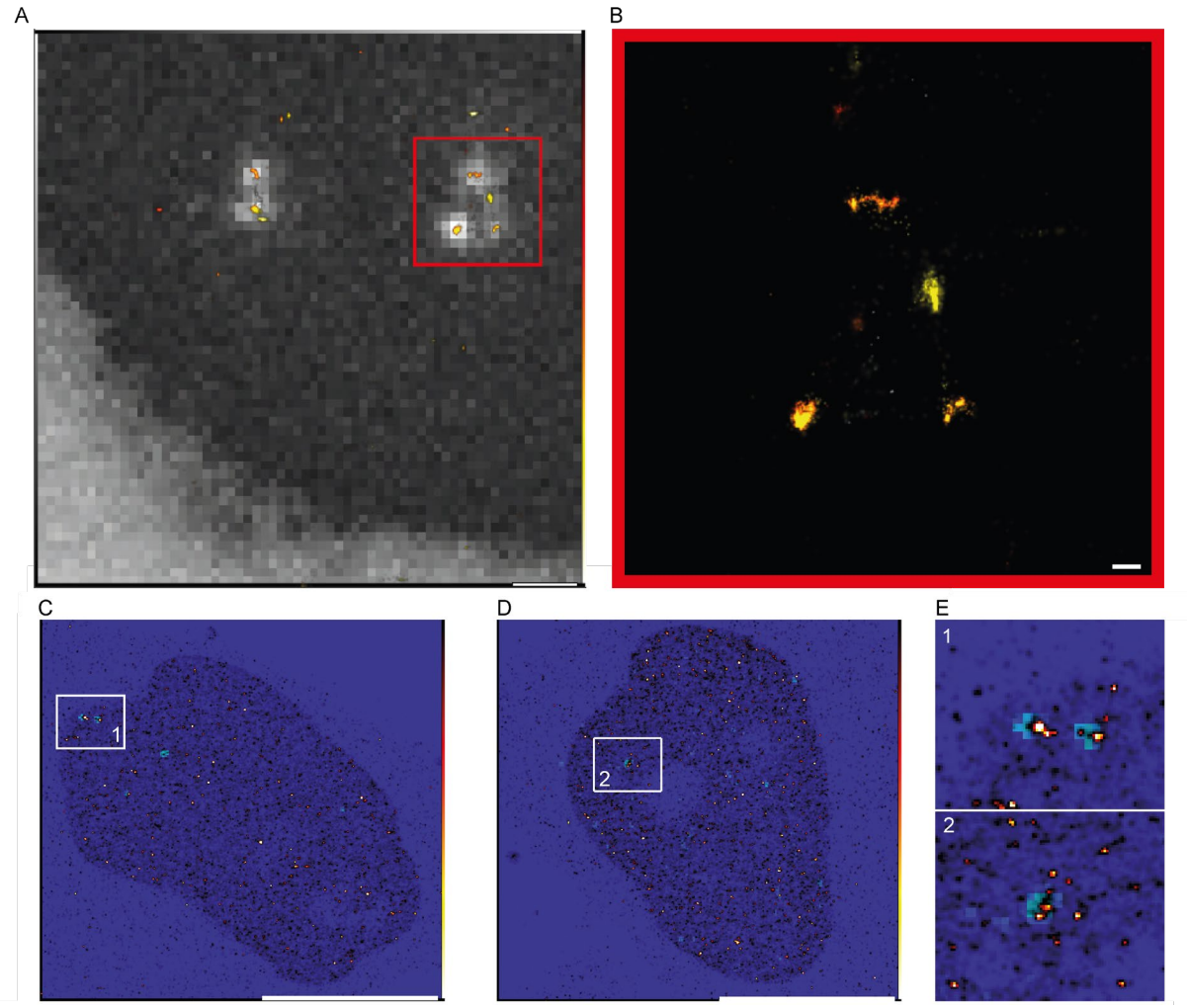

Figure 28: Representative examples of SMLM images of MYC gene in HeLa-K cells. *(A) STORM image of 96 primary probes with 20-nt-long imager strand conjugated to Alexa Fluor 647. DAPI signal (Black), diffraction-limited signal from Alexa Fluor 647 (Grey) and SMLM signal ("Red hot"). Scale bar, 1000 nm. (B) Zoomed region from . Scale bar, 100 nm. (C-D) Spinning disk confocal DNA-PAINT image of 48 primary probes with 12-nt-long docking handles. Diffraction-limited Atto565-bearing 12-nt-long imager strand (Blue) and super-resolved DNA-PAINT with 9-nt-long Cy3B bearing imager strand. (E) Zoomed regions from C and D.*

(xi) Fluorogenic DNA-PAINT imager strands

One potential way of reducing the background observed in DNA-PAINT imaging with excess imager strands in the incubation medium during imaging is to use fluorogenic DNA-PAINT imager strands. These fluorogenic imager strands are dark in solution, taking advantage of Atto655's known quenching in close proximity of guanine bases (Heinlein *et al.*, 2003; Jungmann *et al.*, 2010), and only become fluorescent once bound to the docking handle.

Since my preliminary data obtained in cells suggested that the fluorescence of freely diffusing imager strands can be a challenge when imaging deeply in the nucleus, two new imager strands were designed with a 3- or 4-bp-long hairpin structure (3QP1 and 4QP1, respectively) that bring the 3' guanine close to the 5' Atto655 dye to quench it when in solution (Figure 29 A). Using fluorescence spectrophotometry, I showed that these imager strands were effectively quenched in solution and that the shorter hairpin 3QP1 indeed regained most of its fluorescence upon addition of the complementary docking handle to open up the hairpin. (Figure 29 B). This novel fluorogenic imager strand design, in addition to using crowding agents such as 500 kDa dextran that allow to lower the imager strand concentration in solution, could allow to perform DNA-PAINT imaging with short transiently binding imager strands with low background in the future.

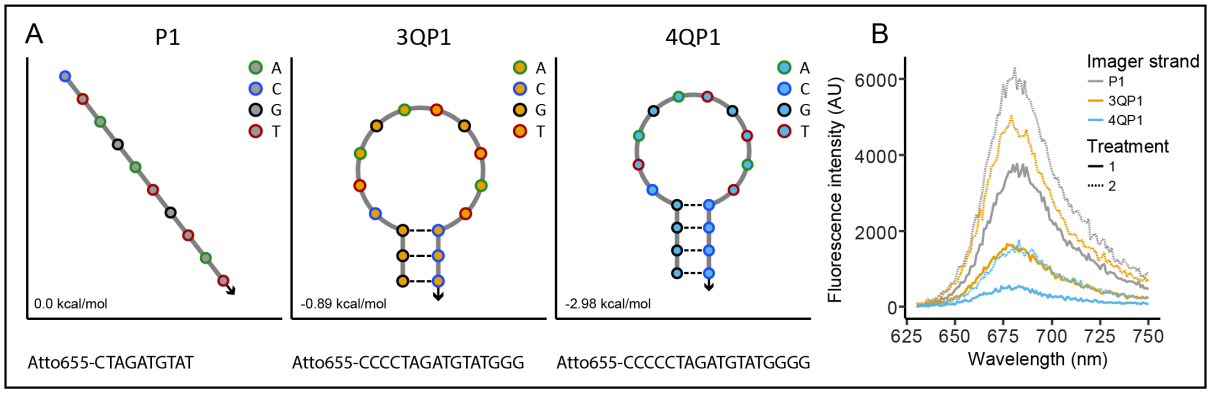

Figure 29: Imager strands containing fluorogenic dyes. *(A) Fluorogenic DNA-PAINT imager strand design. P1 consists of a 9-nt-long sequence with a guanine 4 nt and 7 nt downstream of the Atto655, 3QP1 has 3 cytosines between the fluorophore and the P1 sequence and 3 guanine nt downstream of P1. Predicted secondary structure with -0.89 kcal/mol is shown. 4QP1 contains one additional C and G compared to 3QP1. Predicted secondary structure with -2.89 kcal/mol. Secondary structure prediction performed at 24°C with 500 mM NaCl and can be compared to P1 imager strand (http://www.nupack.org/). (B) Fluorescence spectrophotometry showing that 3QP1 and 4QP1 are fluorogenic.*

EdU-PAINT

Although FISH is a very powerful method for sequence specific labelling, this method includes harsh steps such as treatment with 0.1 N HCl and 0.5% Triton X-100 as well as sample heating at 75°C. In principle, docking handles could be bound directly to the genome after mild *in vivo* incorporation of chemically reactive, non-canonical nucleotides, such as EdU. To test the feasibility of such an approach, the P1 docking handle used in the *in vitro* experiments was covalently bound to EdU bases and incorporated co-replicatively in HeLa cells using click chemistry. As a positive control we coupled Alexa Fluor 647 directly to EdU by click chemistry and as a negative control we used cells that had not incorporated clickable EdU nucleotides. This method is potentially very useful as the imager strand labelling closely resembles the direct coupling of Alexa Fluor 647 to EdU but has the advantage of being applicable for PAINT super-resolution imaging. In the future, strategies should be developed to incorporate more colours with orthogonal click chemistry on several different nucleotides. Moreover, since single DNA base labelling would allow increased resolution, the four DNA bases could potentially be labelled in different colours if enough orthogonal clickable groups become available that are compatible with incorporation during replication. At present, this method could be readily combined with FISH to try to elucidate the path of DNA between loci labelled in a sequence specific manner by FISH, similar to the E_0 imager strand we used in our *in vitro* experiments.

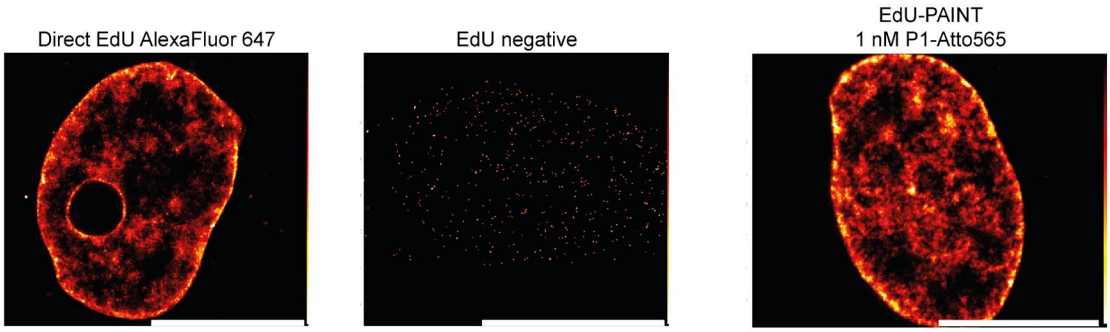

Figure 30: EdU labelled Hela-K cells imaged with SMLM. *Either directly conjugated to Alexa Fluor 647, or without EdU imaged with DNA-PAINT or conjugated to DNA-PAINT docking handles. Imaged on Leica GSDIM. Scale bar 10 µm.*

Taken together, I have established the methodology for creating and using probe libraries that allow flexible utilisation in different imaging modalities for DNA imaging in nuclei of mammalian cells. I have also presented a method that enables reconstruction of 3D chromatin paths in stretches of DNA both *in vitro* and *in situ*. In combination with the ChromoTrace algorithm, they are valuable methods for deciphering the 3D structure of larger stretches of chromatin and can be scaled up for visualisation of an entire chromosomal molecule.

Contributions to the work of collaborators

A quantitative map of human condensins provides new insights into mitotic chromosome architecture (Walther *et al.*, 2018)

Proteins such as the condensin complexes play a crucial role in the organisation of the genome as cells go through mitosis. One of the primary attributes of structural proteins is their abundance in addition to their functional properties. FCS-calibrated imaging was therefore performed on subunits of the condensin I and condensin II complexes in order to measure their concentration and protein abundance. Moreover, their distribution was probed by STED imaging and revealed a relatively sparse protein distribution appearing as resolvable spots. My contribution was to determine protein localisation at subpixel precision utilising the highly anisotropic STED data. For this, I have developed a script which automates the detection and extraction of subpixel protein positions. The final script is a combination of the FIJI (Schindelin *et al.*, 2012) plugin ThunderSTORM v1.3 (https://github.com/zitmen/thunderstorm; Ovesný *et al*. 2014) and DBSCAN (Ester *et al.*, 1996).

From the precise position of subunits of the condensin protein complexes, a three-step hierarchical looping model of mitotic chromosome compaction could be proposed (Walther *et al.*, 2018). in which condensin II initially anchors loops of a maximum size of ~450 kb at the chromatid axis. Upon condensing I binding, the loop size is then further reduced to ~90 kb in prometaphase and ~70 kb in anaphase, when maximum chromosome compaction occurs during sister chromatid segregation.

Chapter 3: Conclusions and discussion

In this work I have established a reliable methodological framework for imaging the folding of genomic DNA in human cells *in situ* using sets of short *in situ* hybridization probes that target unique genomic loci and have bi-functional docking handles to bind fluorophore bearing imager strands. My overall goal was to optimise hybridization conditions, probe design, imaging conditions and computational data analysis to achieve the highest possible genomic resolution in order to unravel the internal structure of chromosomes and their underlying TADs.

Distance measurements from DNA-PAINT on purified M13 phage genomes conform with theoretical B-DNA distances predicted from crystal structures

The first aim was to test the accuracy of the imager strand approach *in vitro*. The first measurements we made with single imager strands were done using docking handles incorporated into DNA origamis. The results aligned extremely well with the expected theoretical structure of the B-DNA double helix. I then moved to the next more complex system and used purified M13 phage genomes as mini-chromosomes and targeted ten loci about 600 nt apart with three primary probes each. Again, I could achieve high hybridisation and labelling efficiencies and could observe that significant looping occurred within the 6.5 kb phage genome *in vitro*. Moreover, we observed that the distance between individual loci was smaller than maximally extended B-DNA, suggesting that looping and folding occurs on DNA at a level below half a kilobase *in vitro*. I also designed a generic imager strand E_0 to decorate the DNA between loci at every 64 nt. Unexpectedly, many of the individual probes with E_0 could often be resolved, showing that a genomic locus can be efficiently labelled with a single primary probe/imager strand complex and that a genomic resolution of less than 100 bp is possible on pure DNA *in vitro*. These 64 nt stretches are typically spaced by only 18 nm, which is about 90% of B-DNA and shows that the DNA-Exchange-PAINT method can provide structural information across very different scales ranging from nm to µm. To summarise, two points along the region predicted to be in a DNA double helix conformation were 93.0 ± 0.1% apart when compared to the theoretical B-DNA distances, and the same length on a flexible mini-chromosome was 85 ± 40% apart, suggesting more flexible molecules.

FISH on adherent cultured cells such as Hela-K and RPE-1 require 48 probes per locus under these conditions for sufficient detection efficiency

The application of this promising protocol for *in situ* labelling of human cell nuclei turned out to require massive optimisation to obtain the best compromise between acceptable labelling efficiency of single loci and maintenance of a close to native nuclear architecture. I first deployed systematic efforts to set up quantitative assays for hybridization and locus labelling efficiency based on computational image analysis and successfully optimized the FISH protocol for the more variable and less accessible *in situ* samples. In order to successfully perform ten rounds of imager strand exchange experiments in cells, 48 rather than 1-3 primary hybridisation probes per locus were required due to the lower hybridisation efficiency of primary probes to chromatin compared to pure DNA. The hybridisation efficiency increased with the amount of probes and reached a quantity where the background also started to increase. In our hands a concentration of 8-10 ng/µl is optimal when 96 probes of 82 nt (60 bp genome complementarity, 2 nt gap sequence and a 20 nt docking handle) is used. This correlates to a 3.1 nmol/µl concentration of each individual probe. The relatively large number of primary probes per set necessary to effectively bind one locus results in a maximal possible genomic resolution of 5 kb when 50 nt long genome complementary regions are used in each probe. To combat the high background in cells which was mostly absent *in vitro*, I also made imager strands bearing two fluorophores and measured that they are indeed close to two times brighter than imager strands with one fluorophore. However, these imager strands generate a higher background which render them useful only when the excess imager strands can be washed away stringently.

Conservation of nuclear architecture

The most critical challenge in multiplex FISH labelling remains to efficiently hybridize the primary probes to the genomic DNA while keeping chromatin and nuclear structure as intact as possible for data interpretation. One of the main tasks has been to optimize every step of the classical FISH procedure to reach the best preservation of nuclear architecture and reach workable labelling efficiencies. Typically, samples dedicated to FISH labelling go through PFA fixation, Triton X-100 permeabilization, protein precipitation with HCL and, finally, genome denaturation using high temperatures. The most careful procedure developed here successfully preserves nuclear architecture as a qualitative observation of stained nuclei do not differ from live cells. Indeed, nuclear shape is fully conserved, no disruption of membrane and DNA leakage to the cytoplasm has ever been observed, and nuclear bodies such as nucleoli are intact. On average, this protocol gave a hybridization efficiency of 5-10%, which we suspect to be due to incomplete decoration of the genomic locus by the primary probe (Pers. Comm., Franziska Kundel, EMBL, Heidelberg). Despite the low FISH probe labelling, about 40% of cells showed 2 or more detectable loci and samples have been of sufficiently good quality to proceed with imaging of 10 probes on the MYC locus in single mammalian cells.

10 colour confocal microscopy with super-resolution capabilities

We have developed the ChromoTrace algorithm to perform *in silico* modelling of DNA-PAINT FISH experiments in order to optimize the design of probe libraries that can be resolved by 3D fluorescence microscopy. Simulations predict that most of a chromosome labelled every 10 kb can be reconstructed if labelled with 10 distinguishable colours (Barton *et al.*, 2018). Ten spectrally distinguishable fluorophores with super-resolution capabilities are not yet available. Additionally, the use of spectrally different fluorophores may introduce localization imprecision due to chromatic aberrations. To overcome this challenge, primary FISH probes were designed with a docking handle sequence. Each docking handle is used as a unique barcode for a single genomic locus. The docking handle is recognized and bound by a fluorescent imager strand.

The docking handle has been designed to form a bi-functional module with a short or a long imager strand. A 12-nt-long imager strand is used for microscopy techniques such as confocal microscopy, STED, or STORM. In this case the imager strand will bind the docking handle for approximately 1000 seconds and allow the acquisition of several confocal stacks or up to 50,000 frames of a STORM video with twenty-millisecond exposure time. The 9-nt-long or 10-nt-long imager strands are used for DNA-PAINT imaging and have a residence time of about 1 second. We have shown that the 12- and the 9-nt-long imager strands co-localize in FISH-labelled nuclei with DNA-PAINT acquired with a spinning disk confocal microscope and that the bi-functionality of our system can be used.

We have tested 10 docking handles together with the 12-nt-long imager strands and observed that eight of them show high specificity and low background with similar spot intensities. The 2 probes with high background are potentially binding to repetitive regions which were absent from the reference genome assembly GRCh38 used for our BLAST searches to exclude off-target binding of the imager strand. This should be addressed prior to upscaling the library to cover an entire chromosome or genome.

RPE1 cells have been used in this study as they are expected to be near-diploid. We have observed that the number of detected loci per cell varies from the expected two copies of the MYC gene in G1-phase and four copies in G2-phase. In our FISH exchange experiments, typically, 40% of cells have the number of expected spots or more.

As the labelling efficiency measured by FCS is around 5-10%, we expect that improvements will still be needed to scale up to full chromosomes or genomes. For instance, if 100 genomic loci would be probed with a 10% hybridization efficiency, this would lead to a negligible amount of full traces [0.10^100 = 0]. This prediction assumes that each successful detection of a locus is independent which might not be the case. Nevertheless, this step should be addressed to insure successful upscaling. Using harsher conditions for sample preparation would increase labelling efficiency but is not desirable as it would be done at the cost of nuclear architecture preservation. The risk would be to lose any biological relevance. Other denaturing methods using less stringent sample treatments exist and would be worth testing with the aim of increasing labelling efficiency.

Alternative labelling approaches

Peptide nucleic acids as an alternative approach

In the future it will be worth to explore alternative hybridisation approaches to improve efficiency and thereby use fewer probes per locus and achieve better genomic resolution. Among the approaches to consider is the use of primary probes based on peptide nucleic acids (PNA). PNA is an artificial synthetic peptide that can form a triple helix with DNA at specific genomic sites (Nielsen *et al.*, 1991). It can bind to DNA without the need to denature the genome by heating and without the need of FA that is normally used to "melt" DNA in FISH experiments (Genet, Cartwright and Kato, 2013). Advances in this technology has allowed the targeting of single loci by utilisation of signal amplification (Yaroslavsky and Smolina, 2013). Modifications of this technique might prove immensely powerful. For example, by choosing 10 PNA sequences that bind every 5000 bp on average, one could unwind the genome every 500 bp (Pers. Comm. Franziska Kundel, EMBL, Heidelberg). By either directly labelling the PNA probes with docking handles or by designing FISH probes for the PNA-unwound regions, one could target the genome at unprecedented resolution with enough colours to resolve the path of the whole genome in one go at single cell level.

Using endo- and exo-nucleases as an alternative approach

A second strategy that can improve the structural integrity of the nucleus is based on the CO-FISH method that originally was developed to detect the tandem repeat orientation within centromeric regions of chromosomes. Recently, this technique has been further developed and has been shown to successfully label non-repetitive regions in a method called RASER-FISH (Brown *et al.*, 2018). Cells labelled overnight with a mix of BrdU and BrdC have these nucleotide analogues incorporated onto one of the DNA strands in their genome before being fixed and permeabilized. Upon treatment with UV light, the sites where BrdU or BrdC have been incorporated will be nicked and can thereafter be treated with Exonuclease III to extend the nicks and produce a single-stranded genomic DNA region. This region can be targeted by FISH probes and is much more accessible than the double-stranded genome which has to

be denatured at high temperatures. This method has already been used to observe TAD-sized structures in cells (Miron *et al.*, 2019).

Image automation with microfluidics and feedback microscopy

Automation of image acquisition would increase the throughput and the reproducibility of experiments with multiple exchange rounds with the 12-nt-long imager strands. Typically, each Exchange round consists of adding the imager strand and a general DNA stain to the cells, followed by a gentle wash. For confocal microscopy z-stacks are acquired and cover the whole nucleus with ~3 µm buffer at the top and bottom of the cell at ~20 positions with averaging and with sub-Nyquist sampling. In total, each exchange round takes about 1 h and results in a minimum of 10 h imaging protocol and even ~20 h imaging when images are also acquired after the FA wash to ensure successful washout of imager strand. In the end such an experiment records about 70-80 cells imaged with 10 colours but the buffer exchanges increase the risk of losing x-y-z focus when done manually, especially since pipetting dexterity naturally decreases during a continuous 20-hours workflow performed by a single person. To fully automate confocal imaging of several exchanges I performed preliminary tests of a microfluidics devise (Vutara 365, Bruker). With the help of the staff members of the Advanced Light Microscopy Facility at EMBL, I programmed instruments to allow communication between the microscope and the fluidics device to ensure that imaging starts after the fluidics program is finished and that the next Exchange cycle is started after the last image (Sebastian Schnorrenberg, ALMF, Heidelberg; Aliaksandr Halavatyi, ALMF, Heidelberg).

Current state-of-the-art FISH methodologies

Recently, great advancements on FISH methodologies have shown that large probe libraries can be used to investigate chromosomes at the level of TAD-like structures in single cells. When averaging across many cells, the TAD borders obtained by imaging align well with Hi-C data but their localisation varies greatly between single cells (Bintu *et al.*, 2018). Methods of FISH signal amplification by concatenation of imager strands have also been shown to be a useful advance and can amplify a signal 10-450 fold depending on the target (Kishi *et al.*, 2018). Driven by these successes, FISH-based imaging approaches are now starting to complement Hi-C to derive genome structure (Bintu *et al.*, 2018; Nir, Farabella, Estrada, *et al.*, 2018; Mateo *et al.*, 2019).

Many details about global genome and chromosome structure at TAD resolution are starting to be revealed by imaging. However, studying the internal structure of single TADs and the critical functional element of individual loops has proven to be very difficult inside cells *in situ*.

With the methods development presented in this thesis, we can now detect specific genomic sites at 10 kb resolution with the potential of reducing it further to 5 kb resolution by tiling the probe library end-to-end and applying the technology developed here. Consequently, we should be able to sample chromatin sub-structures at 3-6 times higher genomic resolution compared to previous studies (Bintu *et al.*, 2018). For illustration, our methods allow us to position 20 probes instead of three along a 100-kb-large loop and therefore to visualise the human genome *in situ* in detail and especially resolve the internal looping architecture of single TAD-sized domains.

At last, we can say that a reliable framework has been established and several new avenues have opened recently to tackle the steps requiring improvement before upscaling to full chromosomes or genomes in the future.

Chapter 4: Materials and methods

Methods

DNA origami and mini-chromosome self-assembly

DNA origami were assembled as previously described and in collaboration with (Schnitzbauer *et al.*, 2017) using the M13 phage scaffold (Cat.# N4040S; New England Biolabs). To make small *in vitro* "mini-chromosomes", M13 phage DNA (Cat.# N4040S; New England Biolabs) was cut with the restriction enzymes BamHI (Cat.# R0136L; New England Biolabs) and BglII (Cat.# R0144S; New England Biolabs) leading to a 6566 nt long linear single stranded DNA molecule. Ten loci along the mini-chromosome were targeted with 28 or 32 nt long primary FISH probes with a spacing of 632 nt between each locus. Each probe contained a 3' extension called docking handle for binding of a secondary probe. Each locus started with two biotin-docking-handle sequences (Supplementary Table 1) followed by three copies of a locus-specific 9-nt long docking-handle (E_1-E_{10}). In addition, six copies of the 9-nt long E_0 docking-handle sequence were distributed evenly between the loci, giving a traceable line between the locus-specific probes. Mini-chromosomes were assembled in a 20 µl reaction mixture [10 nM linearized M13 phage scaffold (Cat.# N4040S; New England Biolabs); 300 nM (each) primary probe library

Supplementary Table 2); 50 nM biotin adapter sequence, 1x TAE buffer; 12.5 mM $MgCl_2$] by denaturation for 5 min at 80°C, followed by a 60°C-to-4°C temperature gradient over 3 h in a mastercycler nexus gradient instrument (Cat.# 6331000017; Eppendorf). After assembly, the labelled mini-chromosomes were separated from non-bound oligos on an $MgCl_2$ enriched agarose gel [1.5% agarose (Cat.# A9539; Sigma-Aldrich); 1x TAE buffer; 10 mM $MgCl_2$; 1.2x SYBR Safe DNA Gel Stain (Cat.# S33102; Thermo Fisher Scientific)] at 4°C using an $MgCl_2$ enriched agarose loading buffer [5% glycerol (Cat.# 104091; Merck Millipore); 0.0042% xylene cyanol FF (Cat.# X4126; Sigma-Aldrich); 0.0042% bromophenol blue (Cat.# 114391; Sigma-Aldrich)]. Assembled mini-chromosomes were cut out from the gel under a UV transilluminator and separated from the gel in a freeze-n-squeeze DNA gel extraction spin column (Cat.# 7326165; Bio-Rad) by centrifugation for 3 min at 1000 rcf.

Attaching DNA origami and mini-chromosomes to IBIDI chambers

Origami and mini-chromosomes were attached to glass bottom slides (Cat.# 80607; Ibidi GmbH) that were pre-cleaned with isopropanol in the following way: each well was incubated with 40 µl BSA-biotin buffer [1 mg/ml BSA-biotin (Cat.# A8549; Sigma-Aldrich); 10mM Tris-HCl (pH 8); 100 mM NaCl] and washed 3 times with 180 µl of buffer A+ [10 mM Tris-HCl (pH 8); 100 mM NaCl; 0.05% (v/v) Tween 20 (Tween 20; Cat.# P2287; Sigma-Aldrich) at pH 8.0]. Wells were incubated for 5 min with 40 µl streptavidin buffer [0.5 mg/ml streptavidin (Cat.# S888; Thermo Fisher Scientific); 10 mM Tris-HCl (pH 8); 100 mM NaCl; 0.05% (v/v) Tween 20], then washed 2 times with 180 µl Buffer A+ and 2 times Buffer B+ [5 mM Tris-HCl (pH 8); 10 mM MgCl2; 1 mM EDTA, 0.05% (v/v) Tween 20]. Then, 40 µl of origami and mini-chromosomes mix [~0.1 nM origami; ~0.025 nM mini-chromosomes] was added and incubated for 8 min. Wells were washed 3 times with 180 µl Buffer B+ before starting the imaging session.

Imaging mini-chromosomes with DNA Exchange-PAINT

Mini-chromosomes were imaged in eleven exchange rounds (E_0-E_{10}) by sequentially adding imager strands targeting different docking handles on the mini-chromosome. Each exchange round started by addition of 60 µl imager strand mixture [5 nM exchange round-specific imager strand; 5 nM reference origami-specific imager strand (P3-Cy3B); 5 mM Tris-HCl (pH 8); 10 mM MgCl2; 1 mM EDTA, 0.05% (v/v) Tween 20; 1x Trolox solution; 1x PCA solution; 1x PCD solution]. While the imager strand solution was transiently binding and unbinding to its docking handle, 8000 frames were acquired with a 300 ms long exposure time. Imaging was performed on a Nikon Eclipse Ti microscope (Nikon Instruments) with a 100x oil-immersion objective (CFI Apo TIRF 100x, NA 1.49; Nikon Instruments) with a 160 nm pixel size. A 561 nm laser (200 mW nominal; Coherent Sapphire) was filtered (ZET561/10; Chroma Technology) and directed to the objective with a multi-band beam splitter (ZT561rdc; Chroma Technology). Light from the fluorescent molecules were filtered (ET600/50m; Chroma Technology) and collected with an EMCCD camera (iXon X3 DU-897; Andor Technologies). The optimal laser intensity was found by increasing the laser power

until the duration of the blinking event decreased. The intensity was set below this value to obtain maximum number of photons per binding event while still not having excessive bleaching during the docking events.

Table 1: DNA-PAINT imager strands and docking sites

Target	Internal ID	Docking handle (5'-3')	Imager strand (5'-3')
Origami (E_0-E_{10})	P3-Cy3B	ttTCAATGTAT	ATACATTGA-Cy3B
E_0	P1-Cy3B	ttATACATCTA	TAGATGTAT-Cy3B
E_1	X61-Cy3B	ttTCCTCAATTA	TAATTGAGGA-Cy3B
E_2	X62-Cy3B	ttACAATTTTCC	GGAAAATTGT-Cy3B
E_3	X63-Cy3B	ttATTTTACACC	GGTGTAAAAT-Cy3B
E_4	X64-Cy3B	ttTCTTATACAC	GTGTATAAGA-Cy3B
E_5	X65-Cy3B	ttACTACTTATC	GATAAGTAGT-Cy3B
E_6	X66-Cy3B	ttTAAATTTCCC	GGGAAATTTA-Cy3B
E_7	X67-Cy3B	ttACTCTATTCA	TGAATAGAGT-Cy3B
E_8	X68-Cy3B	ttATCAATCTTC	GAAGATTGAT-Cy3B
E_9	X69-Cy3B	ttTTTCTAAACC	GGTTTAGAAA-Cy3B
E_{10}	X70-Cy3B	ttTCAATATCTC	GAGATATTGA-Cy3B

Lowercase sequences are gap-sequences

Table 2: 100x Trolox solution (4 ml)

Name	Amount	Reference
Trolox	100 mg	Trolox (Cat.# 238813; Sigma-Aldrich)
Methanol	430 µl	Methanol (Cat.# 32213; Sigma-Aldrich)
1 M NaOH	345 µl	NaOH (Cat.# 31627.29; VWR)
H_2O	3.2 ml	

20 µl aliquots; -20°C for up to 6 months

Table 3: PCA solution

Name	Amount	Reference
PCA	154 mg	PCA (Cat.# 37580-25G-F; Sigma-Aldrich)
NaOH	adjusted to pH 9.0	NaOH (Cat.# 31627.29; VWR)
H2O	To 10 ml total volume	

20 µl aliquots; -20°C for up to 6 months

Table 4: PCD solution

Name	Amount	Reference
PCD	9.3 mg	PCD (Cat.# P8279-25UN; Sigma-Aldrich)
Glycerol-KCl-EDTA-Tris buffer	13.3 ml	50% glycerol; 50 mM KCl; 1 mM EDTA; 100 mM Tris–HCl (pH 8.0)

20 µl aliquots; -20°C for up to 6 months

Subpixel localisation and drift correction of DNA-PAINT images of 20-nm origami grids and mini-chromosomes

Each exchange round video of DNA-PAINT binding and dissociation was processed independently with Picasso Localize version 989 (Schnitzbauer *et al.*, 2017). Subsequently redundant cross-correlation drift correction and drift correction using the 20-nm-grids as reference was performed in Picasso Render version 989 (Schnitzbauer *et al.*, 2017). For line profiles, a raster image with a pixel size of 2.67 nm was generated.

Line profile analysis of 20-nm origami grids and mini-chromosomes

To determine the distance between points in the DNA-PAINT images, raster images were loaded into ImageJ and line profiles were manually drawn. All local maxima with an intensity higher than 10% of maximum intensity were extracted from the 5 pixel broad line profiles and the nearest neighbour distances were calculated.

Primary FISH probe library design

The human reference genome GRCh38.p12 was used to search for unique FISH probes. First, highly repetitive sequences were removed, then all possible probe binding sites (with a sliding window of 1 nt) were blasted against the reference genome to identify possible off-targets and the melting temperature of the most similar off-target for each probe was computed. Probes for which the off-target melting temperature was below and close to 54°C for 40-nt-long probes or 58°C for 50-nt-long probes were selected. This pre-computed probe library could be queried in a fast and flexible way to design FISH probes targeting any locus. In a given experiment, the genomic size of a locus is only limited by the number of probes per kb and locus selection can either prioritise genome spacing or off-target temperature.

Cell culture

HAP-1 cells were provided by Bas van Steensel (Netherlands Cancer Institute, Amsterdam, Netherlands) and grown in IMDM media (Cat.# 12440053; Thermo Fisher Scientific) with 10% (v/v) FBS (Cat.# 10270106; Thermo Fisher Scientific). RPE-1 cells were provided by Jan Korbel (EMBL, Heidelberg, Germany) and grown in a 1:1 mixture of DMEM and Ham's F-12 medium (Cat.# 11320074; Thermo Fisher Scientific) supplemented with 10% (v/v) FBS (Cat.# 10270106; Thermo Fisher Scientific). Hela Kyoto cells were obtained from Shuh Narumiya (Kyoto University, Kyoto, Japan) and grown in DMEM high glucose medium (Cat.# 41965039; Thermo Fisher Scientific) supplemented with 10% (v/v) FBS (Cat.# 10270106; Thermo Fisher Scientific), 100 U/ml penicillin-streptomycin (Cat.# 15140122; Thermo Fisher Scientific), 1 mM sodium pyruvate (Cat.# 11360070; Thermo Fisher Scientific) and 2 mM L-glutamine (Cat.# 25030081; Thermo Fisher Scientific). All cell lines were grown in 5% CO_2 at 37°C in a cell culture incubator and passaged at 70-90% confluency every 2-3 days by trypsinisation with 0.05% Trypsin-EDTA (Cat.# 25300054; Thermo Fisher Scientific).

Cell fixation

50 µl aliquots containing 25,000-50,000 cells were seeded per channel in glass bottom channel slides (Cat.# µ-Slide VI 0.5 Glass Bottom; Ibidi GmbH) and grown overnight before fixation with 4% w/v PFA (Cat.# 15710; Electron Microscopy Sciences) in 1x PBS for 15 min. Cells were permeabilised with 0.5% (v/v) Triton X-100 (Cat.# T8787; Sigma-Aldrich) in 1x PBS for 20 min and with 0.1 N hydrochloric acid (Cat.# 109057; Merck Millipore) for 15 min. Fixed cells were stored in 50% v/v FA (Cat.# AM9342; Thermo Fisher Scientific) in 2x SSC buffer (Cat.# AM9763; Thermo Fisher Scientific) for up to 8 weeks.

Fluorescence *in situ* hybridization (FISH)

Genome labelling was achieved by fluorescence *in situ* hybridization. Primary hybridisation buffer (H1FA50) [50% (v/v) FA (Cat.# AM9342; Thermo Fisher Scientific); 2x SSC (SSC (20x), RNase-free; Cat.# AM9763; Thermo Fisher Scientific); 10% w/v Dextran sulfate sodium salt from *Leuconostoc spp.* (Cat.# D8906-10G; Sigma-Aldrich); 0.4 µg/µl RNase A (Ribonuclease A from bovine pancreas; Cat.# R-4642; Sigma-Aldrich)] and primary probe library dissolved in 1x TE buffer, pH 7.5, were pre-warmed to room temperature for 0.5-2 h with shaking and protected from light. PFA-fixed cells were re-permeabilised in 0.5% (v/v) Triton X-100 (Sigma-Aldrich; Cat.# T8787; Sigma-Aldrich) in 1x PBS for 10 min before rinsing with 50% (v/v) FA in 2x SSC buffer. Cells were incubated in H1FA50 buffer for 1 h at 37°C in a humidified chamber (ThermoBrite; Leica Biosystems) prior to adding ~0.00019 pmol/µl of each individual probe in H1FA50 buffer [amounts to 0.5 ng/µL for 96 probes of 82 nt length] followed by 1 h incubation at 37°C. Denaturation was performed for 3 min at 75°C in the humidified chamber, followed by overnight hybridisation at 37°C. Washes were performed by first washing 3 times in 2x SSC + 0.2% (v/v) Tween 20 for 5 min at room temperature, followed by 2 times washing in 0.2x SSC + 0.2% (v/v) Tween20 at 60°C for 7 min followed by a 5 minute wash in 4x SSC + 0.2% (v/v) Tween20.

For permanently bound (20-nt long) imager strands, hybridisation was performed with 10 nM imager strand in H2FA25 buffer [25% (v/v) FA; 2x SSC; 10% (w/v) Dextran sulfate; 0.1% (v/v) Tween20] for 2 h or overnight at 30°C in the dark. The samples

were subsequently washed 3 times for 5 min in 25% FA in 2x SSC, and 3 times for 5 min in 2x SSC and stained with 5 ng/µl DAPI in 2x SSC for 10 min before use.

Sequential imaging of FISH probes with confocal microscope

In the region between the MYC promoter and the MYC 335 enhancer (GRCh37: chr8q24.21: 128414228-128414237) 10 5-kb-long loci spaced 10 kb apart from each other were targeted with 48x 50-nt-long FISH probes. Each locus had a unique 5' 12-nt-long docking handle extension that was targeted with a complementary imager strand (Supplementary Table 4 - Supplementary Table 12). For each exchange round 100 nM of the imager strand was incubated for 5 min and washed with 500 mM NaCl in 1x PBS, followed by imaging. Subsequently, 25% FA in 1x PBS was used to detach the bound imager strand and the sample was rinsed with 500 mM NaCl in 1x PBS before re-staining with DAPI in 500 mM NaCl in 1x PBS. The washed sample was imaged to confirm the complete wash of E_1 prior to E_2 addition. The protocol was repeated for all the imager strands for 10 colour multiplexing.

Table 5: Bifunctional 12-nt-long docking handles

Name	Docking handle (9nt)	Docking handle (12nt)	Imager (9nt)	Imager (12nt)
E_1	…ttATACATCTA	…ttATACATCTACGG	TAGATGTAT-dye	CCGTAGATGTAT-dye
E_2	…ttTCTTCATTA	…ttTCTTCATTAGCG	TAATGAAGA-dye	CGCTAATGAAGA-dye
E_3	…ttTCAATGTAT	…ttTCAATGTATGGC	ATACATTGA-dye	GCCATACATTGA-dye
E_4	…ttAAAAAGTTC	…ttAAAAAGTTCGAG	GAACTTTTT-dye	CTCGAACTTTTT-dye
E_5	…ttTAGTTAGAG	…ttTAGTTAGAGCCC	CTCTAACTA-dye	GGGCTCTAACTA-dye
E_6	…ttTTGATGATA	…ttTTGATGATAGCC	TATCATCAA-dye	GGCTATCATCAA-dye
E_7	…ttATAAAGTGT	…ttATAAAGTGTCCA	ACACTTTAT-dye	TGGACACTTTAT-dye
E_8	…ttATATGATCT	…ttATATGATCTCCG	AGATCATAT-dye	CGGAGATCATAT-dye
E_9	…ttTATTAAGCT	…ttTATTAAGCTCGC	AGCTTAATA-dye	GCGAGCTTAATA-dye
E_{10}	…ttTTAAAACAG	…ttTTAAAACAGCCT	CTGTTTTAA-dye	AGGCTGTTTTAA-dye

Lowercase sequences are gap-sequences; All sequences are in 5' to 3' direction

Laser-scanning confocal microscopy

Protocol validation and optimisation was performed by imaging on an inverted laser-scanning microscope (LSM780; Carl Zeiss AG) using a 63x oil objective at 23-26°C. Multi-position z-stacks covering whole nuclei in x-y-z were acquired with a pixel size

that sampled diffraction-limited spots ~3-5 times in x-y and ~3 times in z (typically: 0.0900 x 0.0900 x 0.47 µm^3). For experiments with fluid exchange, a 5 ml syringe was connected with tubing to the luer adapter of the IBIDI slide. Liquid was removed from one side of the channel using the syringe while fluid was added to the other end of the channel.

A plexiglass IBIDI stabiliser was developed in collaboration with the mechanical workshop at EMBL Heidelberg to avoid x, y and z drift during long imaging times (Figure 31). The stabilizing device was fixed to the IBIDI slides with picodent twinsil (Cat.# 13001000; Picodent).

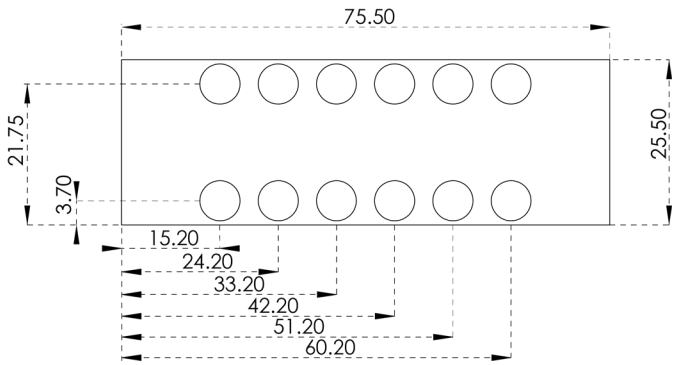

Figure 31: IBIDI stabiliser. *9 mm thick plexiglass with twelve ⌀ = 7 mm holes drilled for IBIDI luer adaptor access. Modified from (ibidi.com). Numbers in mm.*

Spot detection

A spot detection algorithm based on ImageJ (Schindelin *et al.*, 2012) was written to detect and quantify spots from the confocal exchange experiments. The algorithm first aligned the nuclei that had been imaged several times over many exchanges based on the nuclear channel using a phase correlation based algorithm (Parslow, Cardona and Bryson-Richardson, 2014), then interpolated the images to achieve an isotropic pixel size, followed by a median filter with a kernel with a size close to the diffraction limit. Thresholds were computed by removing the 0.01% brightest pixels and setting the threshold to the highest remaining pixel value (this efficiently removed the spots from the histograms). Based on this threshold the images were converted to binary images to define the region of each spot. 3D coordinates were extracted from each spot region by computing the weighted mean from the pixels in the interpolated image.

Materials

Table 6: List of reagents and materials

Name	Cat.#	Company
FBS	10270106	Thermo Fisher Scientific
DMEM and Ham's F-12	11320074	Thermo Fisher Scientific
IMDM	12440053	Thermo Fisher Scientific
DMEM	41965039	Thermo Fisher Scientific
Penicillin-streptomycin	15140122	Thermo Fisher Scientific
Sodium pyruvate	11360070	Thermo Fisher Scientific
L-glutamine	25030081	Thermo Fisher Scientific
Trypsin-EDTA	25300054	Thermo Fisher Scientific
Glass bottom channel slides	µ-Slide VI 0.5 Glass Bottom	Ibidi GmbH
PFA	15710	Electron Microscopy Sciences
Triton X-100	T8787	Sigma-Aldrich
Tween 20	P2287	Sigma-Aldrich
Hydrochloric acid	109057	Merck Millipore
Formamide	AM9342	Thermo Fisher Scientific
SSC buffer	AM9763	Thermo Fisher Scientific
175 cm^2 tissue culture flasks	10078780	Fisher Scientific
M13 phage scaffold	N4040S	New England Biolabs
BamHI	R0136L	New England Biolabs
BglII	R0144S	New England Biolabs
Mastercycler nexus gradient	6331000017	Eppendorf
Agarose	A9539	Sigma-Aldrich
SYBR Safe DNA Gel Stain	S33102	Thermo Fisher Scientific
Glycerol	104091	Merck Millipore
Xylene cyanol FF	X4126	Sigma-Aldrich
Bromophenol blue	114391	Sigma-Aldrich
Freeze-n-squeeze DNA gel extraction spin column	7326165	Bio-Rad
µ-slide VI 0.5 glass bottom	80607	Ibidi GmbH
BSA-biotin	A8549	Sigma-Aldrich
Streptavidin	S888	Thermo Fisher Scientific

Picodent twinsil	13001000	Picodent
Trolox	238813	Sigma-Aldrich
NaOH	31627.29	VWR
Methanol	32213	Sigma-Aldrich
PCA	37580-25G-F	Sigma-Aldrich
PCD	P8279-25UN	Sigma-Aldrich

Bibliography

Abbe, E. and Translated by Fripp, H. E. (1874) *A Contribution to the Theory of the Microscope and the nature of Microscopic Vision*. London: Proceedings of the Bristol Naturalists' Society - Google Books. Available at: https://books.google.de/books?id=TC8UAAAAYAAJ&dq=Bristol+Naturalists+Society+1875&pg=PA200&redir_esc=y#v=onepage&q&f=false (Accessed: 23 August 2019).

Barton, C. *et al.* (2018) 'ChromoTrace: Computational reconstruction of 3D chromosome configurations for super-resolution microscopy.', *PLoS computational biology*, 14(3), p. e1006002. doi: 10.1371/journal.pcbi.1006002.

Beliveau, B. J. *et al.* (2015) 'Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes.', *Nature communications*, 6(May), p. 7147. doi: 10.1038/ncomms8147.

Bintu, B. *et al.* (2018) 'Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells.', *Science (New York, N.Y.)*, 362(6413), p. eaau1783. doi: 10.1126/science.aau1783.

Bolzer, A. *et al.* (2005) 'Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes.', *PLoS biology*, 3(5), p. e157. doi: 10.1371/journal.pbio.0030157.

Branco, M. R. and Pombo, A. (2006) 'Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations', *PLoS Biology*. Edited by P. Becker, 4(5), pp. 780–788. doi: 10.1371/journal.pbio.0040138.

Brown, J. M. *et al.* (2018) 'A tissue-specific self-interacting chromatin domain forms independently of enhancer-promoter interactions', *Nature Communications*, 9(1), pp. 1–15. doi: 10.1038/s41467-018-06248-4.

BROWN, S. W. and NUR, U. (1964) 'HETEROCHROMATIC CHROMOSOMES IN THE COCCIDS.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 145(3628), pp. 130–6. doi: 10.1126/science.145.3628.130.

Busslinger, G. A. *et al.* (2017) ‘Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl.’, *Nature*. Nature Publishing Group, 544(7651), pp. 503–507. doi: 10.1038/nature22063.

Cooper, K. W. (1959) ‘Cytogenetic analysis of major heterochromatic elements (especially Xh and Y) in Drosophila melanogaster, and the theory of “heterochromatin”’, *Chromosoma*, 10(1–6), pp. 535–588. doi: 10.1007/BF00396588.

Cremer, T. *et al.* (1982) ‘Rabl’s model of the interphase chromosome arrangement tested in Chinise hamster cells by premature chromosome condensation and laser-UV-microbeam experiments’, *Human Genetics*, 60(1), pp. 46–56. doi: 10.1007/BF00281263.

Darrow, E. M. *et al.* (2016) ‘Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 113(31), pp. E4504-12. doi: 10.1073/pnas.1609643113.

Davey, C. A. *et al.* (2002) ‘Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution.’, *Journal of molecular biology*, 319(5), pp. 1097–113. doi: 10.1016/S0022-2836(02)00386-8.

Dekker, J. (2002) ‘Capturing Chromosome Conformation’, *Science*, 295(5558), pp. 1306–1311. doi: 10.1126/science.1067799.

Dileep, V. *et al.* (2015) ‘Topologically-associating domains and their long-range contacts are established during early G1 coincident with the establishment of the replication timing program.’, *Genome research*. doi: 10.1101/gr.183699.114.

Dixon, J. R. *et al.* (2012) ‘Topological domains in mammalian genomes identified by analysis of chromatin interactions’, *Nature*, 485(7398), pp. 376–380. doi: 10.1038/nature11082.

Dostie, J. *et al.* (2006) ‘Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements’, *Genome Research*, 16(10), pp. 1299–1309. doi: 10.1101/gr.5571506.

Dowen, J. M. *et al.* (2014) ‘Control of Cell Identity Genes Occurs in Insulated Neighborhoods in Mammalian Chromosomes’, *Cell*, 159(2), pp. 374–387. doi: 10.1016/j.cell.2014.09.030.

Drew, H. R. *et al.* (1981) ‘Structure of a B-DNA dodecamer: conformation and dynamics.’, *Proceedings of the National Academy of Sciences of the United States of America*, 78(4), pp. 2179–83. doi: 10.1073/pnas.78.4.2179.

Durand, N. C. *et al.* (2016) ‘Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom.’, *Cell systems*. NIH Public Access, 3(1), pp. 99–101. doi: 10.1016/j.cels.2015.07.012.

Eltsov, M. *et al.* (2008) ‘Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ.’, *Proceedings of the National Academy of Sciences of the United States of America*, 105(50), pp. 19732–7. doi: 10.1073/pnas.0810057105.

Erdel, F. and Rippe, K. (2018) ‘Formation of Chromatin Subcompartments by Phase Separation.’, *Biophysical journal*. The Biophysical Society, 114(10), pp. 2262–2270. doi: 10.1016/j.bpj.2018.03.011.

Ester, M. *et al.* (1996) ‘A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise’, in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. AAAI Press (KDD’96), pp. 226–231. Available at: http://dl.acm.org/citation.cfm?id=3001460.3001507.

Flemming, W. (1882) *Zellsubstanz, kern und zelltheilung*. Leipzig: Leipzig, F. C. W. Vogel. Available at: http://books.google.com/books?id=ndYcngEACAAJ&pgis=1 (Accessed: 22 October 2014).

Franklin, R. E. and Gosling, R. G. (1953) ‘Molecular configuration in sodium thymonucleate.’, *Nature*, 171(4356), pp. 740–1. doi: 10.1038/171740a0.

Ganai, N., Sengupta, S. and Menon, G. I. (2014) ‘Chromosome positioning from activity-based segregation.’, *Nucleic acids research*, 42(7), pp. 4145–59. doi: 10.1093/nar/gkt1417.

Genet, M. D., Cartwright, I. M. and Kato, T. A. (2013) 'Direct DNA and PNA probe binding to telomeric regions without classical in situ hybridization', *Molecular Cytogenetics*. Molecular Cytogenetics, 6(1), p. 1. doi: 10.1186/1755-8166-6-42.

Grigoryev, S. A. *et al.* (2009) 'Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions', *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 106(32), pp. 13317–13322. doi: 10.1073/pnas.0903280106.

Heinlein, T. *et al.* (2003) 'Photoinduced Electron Transfer between Fluorescent Dyes and Guanosine Residues in DNA-Hairpins', *The Journal of Physical Chemistry B*, 107(31), pp. 7957–7964. doi: 10.1021/jp0348068.

Heitz, E. (1928) 'Das Heterochromatin der Moose', *Jahrbücher für Wissenschaftliche Botanik*. Bornträger, (69), pp. 762–818. Available at: https://books.google.de/books?id=3JwgSQAACAAJ.

Hsieh, T.-H. S. *et al.* (2015) 'Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C.', *Cell*, 162(1), pp. 108–19. doi: 10.1016/j.cell.2015.05.048.

Hughes-Schrader, S. (1948) 'Cytology of coccids (Coccoïdea-Homoptera).', *Advances in genetics*, 35(2), pp. 127–203. Available at: https://www.sciencedirect.com/science/article/pii/S006526600860468X (Accessed: 7 September 2019).

Jackson, D. A. and Pombo, A. (1998) 'Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells.', *The Journal of cell biology*. Rockefeller University Press, 140(6), pp. 1285–95. doi: 10.1083/jcb.140.6.1285.

Jost, D. *et al.* (2014) 'Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains', *Nucleic Acids Research*, 42(15), pp. 9553–9561. doi: 10.1093/nar/gku698.

Joti, Y. *et al.* (2012) 'Chromosomes without a 30-nm chromatin fiber.', *Nucleus (Austin, Tex.)*. Taylor & Francis, 3(5), pp. 404–10. doi: 10.4161/nucl.21222.

Jungmann, R. *et al.* (2010) 'Single-Molecule Kinetics and Super-Resolution Microscopy by Fluorescence Imaging of Transient Binding on DNA Origami', *Nano Letters*, 10(11), pp. 4756–4761. doi: 10.1021/nl103427w.

Kishi, J. Y. *et al.* (2018) 'SABER enables highly multiplexed and amplified detection of DNA and RNA in cells and tissues', *bioRxiv*, p. 401810. doi: 10.1101/401810.

Kornberg, R. D. (1974) 'Chromatin structure: a repeating unit of histones and DNA.', *Science (New York, N.Y.)*, 184(4139), pp. 868–71. doi: 10.1126/science.184.4139.868.

Kornberg, R. D. and Thomas, J. O. (1974) 'Chromatin structure; oligomers of the histones.', *Science (New York, N.Y.)*, 184(4139), pp. 865–8. doi: 10.1126/science.184.4139.865.

Kossel, A. (1911) *Ueber die chemische Beschaffenheit des Zellkerns....* Available at: http://books.google.com/books?id=bunImAEACAAJ&pgis=1 (Accessed: 22 October 2014).

Lakadamyali, M. and Cosma, M. P. (2015) 'Advanced microscopy methods for visualizing chromatin structure.', *FEBS letters*. Federation of European Biochemical Societies. doi: 10.1016/j.febslet.2015.04.012.

Lieberman-Aiden, E. *et al.* (2009) 'Comprehensive mapping of long-range interactions reveals folding principles of the human genome.', *Science (New York, N.Y.)*, 326(5950), pp. 289–93. doi: 10.1126/science.1181369.

Lomvardas, S. *et al.* (2006) 'Interchromosomal Interactions and Olfactory Receptor Choice', *Cell*, 126(2), pp. 403–413. doi: 10.1016/j.cell.2006.06.035.

Losada, A., Hirano, M. and Hirano, T. (1998) 'Identification of Xenopus SMC protein complexes required for sister chromatid cohesion', *Genes & Development*, 12(13), pp. 1986–1997. doi: 10.1101/gad.12.13.1986.

Marbouty, M. *et al.* (2015) 'Condensin- and Replication-Mediated Bacterial Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging', *Molecular Cell*, 59(4), pp. 588–602. doi:

10.1016/j.molcel.2015.07.020.

Mateo, L. J. *et al.* (2019) 'Visualizing DNA folding and RNA in embryos at single-cell resolution', *Nature*. Springer US, 568(7750), pp. 49–54. doi: 10.1038/s41586-019-1035-4.

MBINFO (2017) *nucleosome-DNA-packaging.jpg (JPEG Image, 850 × 331 pixels)*, *uploads/06/*. MBINFO; This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. Available at: https://www.mechanobio.info/wp-content/uploads/2017/06/nucleosome-DNA-packaging.jpg (Accessed: 8 September 2019).

Miescher, J. F. (1871) 'Ueber die chemische Zusammensetzung der Eiterzellen', *Hoppe-Seyler's medicinisch-chemische Untersuchungen*. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Ueber+die+chemische+Zusammensetzung+der+Eiterzellen.#0 (Accessed: 22 October 2014).

Miron, E. *et al.* (2019) 'Chromatin arranges in filaments of blobs with nanoscale functional zonation', *bioRxiv*, p. 566638. doi: 10.1101/566638.

Moindrot, B. *et al.* (2012) '3D chromatin conformation correlates with replication timing and is conserved in resting cells', *Nucleic Acids Research*, 40(19), pp. 9470–9481. doi: 10.1093/nar/gks736.

Ni, Y. *et al.* (2017) 'Super-resolution imaging of a 2.5 kb non-repetitive DNA in situ in the nuclear genome using molecular beacon probes', *eLife*, 6. doi: 10.7554/eLife.21660.

Nielsen, P. E. *et al.* (1991) 'Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide', *Science*, 254(5037), pp. 1497–1500. doi: 10.1126/science.1962210.

Nir, G., Farabella, I., Pérez Estrada, C., *et al.* (2018) 'Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling', *PLOS Genetics*. Edited by G. P. Copenhaver. Public Library of Science, 14(12), p. e1007872. doi: 10.1371/journal.pgen.1007872.

Nir, G., Farabella, I., Estrada, C. P., *et al.* (2018) 'Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling', *bioRxiv*. Cold Spring Harbor Laboratory, p. 374058. doi: 10.1101/374058.

Nora, E. P. *et al.* (2017) 'Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization.', *Cell*, 169(5), pp. 930-944.e22. doi: 10.1016/j.cell.2017.05.004.

Olins, A. L. and Olins, D. E. (1974) 'Spheroid chromatin units (v bodies).', *Science (New York, N.Y.)*, 183(4122), pp. 330–2. doi: 10.1126/science.183.4122.330.

Ovesný, M. *et al.* (2014) 'ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging.', *Bioinformatics (Oxford, England)*, 30(16), pp. 2389–90. doi: 10.1093/bioinformatics/btu202.

Parelho, V. *et al.* (2008) 'Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms', *Cell*, 132(3), pp. 422–433. doi: 10.1016/j.cell.2008.01.011.

Parslow, A., Cardona, A. and Bryson-Richardson, R. J. (2014) 'Sample Drift Correction Following 4D Confocal Time-lapse Imaging', *Journal of Visualized Experiments*, (86). doi: 10.3791/51086.

Passarge, E. (1979) 'Emil Heitz and the concept of heterochromatin: Longitudinal chromosome differentiation was recognized fifty years ago', *American Journal of Human Genetics*, 31(2), pp. 106–115.

Phillips-Cremins, J. E. *et al.* (2013) 'Architectural protein subclasses shape 3D organization of genomes during lineage commitment.', *Cell*, 153(6), pp. 1281–95. doi: 10.1016/j.cell.2013.04.053.

Di Pierro, M. *et al.* (2016) 'Transferable model for chromosome architecture', *Proceedings of the National Academy of Sciences*, 113(43), pp. 12168–12173. doi: 10.1073/pnas.1613607113.

Di Pierro, M. *et al.* (2017) 'De novo prediction of human chromosome structures: Epigenetic marking patterns encode genome architecture', *Proceedings of the National Academy of Sciences*, 114(46), pp. 12126–12131. doi:

10.1073/pnas.1714980114.

Politi, A. Z. *et al.* (2018) ‘Quantitative mapping of fluorescently tagged cellular proteins using FCS-calibrated four-dimensional imaging.’, *Nature protocols*. Nature Publishing Group, 13(6), pp. 1445–1464. doi: 10.1038/nprot.2018.040.

Pope, B. D. *et al.* (2014) ‘Topologically associating domains are stable units of replication-timing regulation.’, *Nature*, 515(7527), pp. 402–405. doi: 10.1038/nature13986.

Rabl, C. (1885) ‘Über Zelltheilung. Morphologisches Jahrbuch. 10’, *Morphologisches Jahrbuch*, 10, pp. 214–336.

Rao, S. S. P. *et al.* (2014) ‘A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping’, *Cell*. Elsevier Inc., 159(7), pp. 1665–1680. doi: 10.1016/j.cell.2014.11.021.

Rao, S. S. P. *et al.* (2017) ‘Cohesin Loss Eliminates All Loop Domains’, *Cell*. Cell Press, 171(2), pp. 305-320.e24. doi: 10.1016/J.CELL.2017.09.026.

Richmond, T. J. *et al.* (1984) ‘Structure of the nucleosome core particle at 7 A resolution.’, *Nature*, 311(5986), pp. 532–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6482966 (Accessed: 23 October 2014).

Rivera-Mulia, J. C. and Gilbert, D. M. (2016) ‘Replicating Large Genomes: Divide and Conquer’, *Molecular Cell*. Cell Press, 62(5), pp. 756–765. doi: 10.1016/J.MOLCEL.2016.05.007.

Robinson, P. J. J. *et al.* (2006) ‘EM measurements define the dimensions of the “30-nm” chromatin fiber: Evidence for a compact, interdigitated structure’, *Proceedings of the National Academy of Sciences of the United States of America*, 103(17), pp. 6506–6511. doi: 10.1073/pnas.0601212103.

Rust, M. J., Bates, M. and Zhuang, X. (2006) ‘Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).’, *Nature methods*, 3(10), pp. 793–795. doi: 10.1038/nmeth929.

Schalch, T. *et al.* (2005) ‘X-ray structure of a tetranucleosome and its implications for

the chromatin fibre.', *Nature*, 436(7047), pp. 138–41. doi: 10.1038/nature03686.

Schindelin, J. *et al.* (2012) 'Fiji: An open-source platform for biological-image analysis', *Nature Methods*, 9(7), pp. 676–682. doi: 10.1038/nmeth.2019.

Schnitzbauer, J. *et al.* (2017) 'Super-resolution microscopy with DNA-PAINT', *Nature Protocols*. Nature Publishing Group, 12(6), pp. 1198–1228. doi: 10.1038/nprot.2017.024.

Schrader, F. (1921) 'The chromosomes of Pseudococcus nipæ', *Biological Bulletin*, 40(5), pp. 259–269. doi: https://www.jstor.org/stable/1536736.

Schwartz, Y. B. and Cavalli, G. (2017) 'Three-Dimensional Genome Organization and Function in *Drosophila*', *Genetics*, 205(1), pp. 5–24. doi: 10.1534/genetics.115.185132.

Schwarzer, W. *et al.* (2017) 'Two independent modes of chromatin organization revealed by cohesin removal', *Nature*. Nature Publishing Group, 551(7678), pp. 51–56. doi: 10.1038/nature24281.

Sharonov, A. and Hochstrasser, R. M. (2006) 'Wide-field subdiffraction imaging by accumulated binding of diffusing probes', *Proceedings of the National Academy of Sciences*, 103(50), pp. 18911–18916. doi: 10.1073/pnas.0609643104.

Simonis, M. *et al.* (2006) 'Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C)', *Nature Genetics*, 38(11), pp. 1348–1354. doi: 10.1038/ng1896.

Simpson, R. T. (1978) 'Structure of the Chromatosome, a Chromatin Particle Containing 160 Base Pairs of DNA and All the Histones', *Biochemistry*, 17(25), pp. 5524–5531. doi: 10.1021/bi00618a030.

Song, F. *et al.* (2014) 'Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units.', *Science (New York, N.Y.)*, 344(6182), pp. 376–80. doi: 10.1126/science.1251413.

Speicher, M. R. and Carter, N. P. (2005) 'The new cytogenetics: blurring the boundaries with molecular biology.', *Nature reviews. Genetics*, 6(10), pp. 782–92.

doi: 10.1038/nrg1692.

Splinter, E. *et al.* (2006) ‘CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus.’, *Genes & development*, 20(17), pp. 2349–54. doi: 10.1101/gad.399506.

Stedman, W. *et al.* (2008) ‘Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators’, *The EMBO Journal*, 27(4), pp. 654–666. doi: 10.1038/emboj.2008.1.

Sumara, I. *et al.* (2000) ‘Characterization of vertebrate cohesin complexes and their regulation in prophase.’, *The Journal of cell biology*. The Rockefeller University Press, 151(4), pp. 749–62. doi: 10.1083/jcb.151.4.749.

Thoma, F. and Koller, T. (1977) ‘Influence of histone H1 on chromatin structure’, *Cell*. Cell Press, 12(1), pp. 101–107. doi: 10.1016/0092-8674(77)90188-X.

Tóth, A. *et al.* (1999) ‘Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication.’, *Genes & development*. Cold Spring Harbor Laboratory Press, 13(3), pp. 320–33. doi: 10.1101/gad.13.3.320.

Tremethick, D. J. (2007) ‘Higher-order structures of chromatin: the elusive 30 nm fiber.’, *Cell*, 128(4), pp. 651–4. doi: 10.1016/j.cell.2007.02.008.

Tsunaka, Y. (2005) ‘Alteration of the nucleosomal DNA path in the crystal structure of a human nucleosome core particle’, *Nucleic Acids Research*, 33(10), pp. 3424–3434. doi: 10.1093/nar/gki663.

Vietri Rudan, M. *et al.* (2015) ‘Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture.’, *Cell reports*. Elsevier, 10(8), pp. 1297–309. doi: 10.1016/j.celrep.2015.02.004.

Walther, N. *et al.* (2018) ‘A quantitative map of human Condensins provides new insights into mitotic chromosome architecture’, *The Journal of Cell Biology*, 217(7), pp. 2309–2328. doi: 10.1083/jcb.201801048.

Wang, J. C. (1979) ‘Helical repeat of DNA in solution.’, *Proceedings of the National*

Academy of Sciences of the United States of America, 76(1), pp. 200–3. doi: 10.1073/pnas.76.1.200.

Watson, J. D. and Crick, F. H. (1953) 'Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.', *Nature*, 171(4356), pp. 737–8. doi: 10.1038/171737a0.

Wendt, K. S. *et al.* (2008) 'Cohesin mediates transcriptional insulation by CCCTC-binding factor', *Nature*, 451(7180), pp. 796–801. doi: 10.1038/nature06634.

Wilkins, M. H. F., Stokes, A. R. and Wilson, H. R. (1953) 'Molecular structure of deoxypentose nucleic acids.', *Nature*, 171(4356), pp. 738–40. doi: 10.1038/171738a0.

Würtele, H. and Chartrand, P. (2006) 'Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology', *Chromosome Research*, 14(5), pp. 477–495. doi: 10.1007/s10577-006-1075-0.

Wutz, G. *et al.* (2017) 'Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins.', *The EMBO journal*. John Wiley & Sons, Ltd, 36(24), pp. 3573–3599. doi: 10.15252/embj.201798004.

Xiang, W. *et al.* (2018) 'Correlative live and super-resolution imaging reveals the dynamic structure of replication domains.', *The Journal of cell biology*, 217(6), pp. 1973–1984. doi: 10.1083/jcb.201709074.

Yaroslavsky, A. I. and Smolina, I. V. (2013) 'Fluorescence imaging of single-copy DNA sequences within the human genome using PNA-directed padlock probe assembly.', *Chemistry & biology*, 20(3), pp. 445–53. doi: 10.1016/j.chembiol.2013.02.012.

Zhao, Z. *et al.* (2006) 'Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions', *Nature Genetics*, 38(11), pp. 1341–1347. doi: 10.1038/ng1891.

Zhou, B.-R. R. *et al.* (2015) ‘Structural Mechanisms of Nucleosome Recognition by Linker Histones’, *Molecular Cell*, 59(4), pp. 628–38. doi: 10.1016/j.molcel.2015.06.025.

List of figures

List of tables

List of supplementary tables

Supplementary tables

Supplementary Table 1: Biotin adapter sequence

Name	Biotin Docking Handle (5'-3')	Biotin Adapter Sequence (5'-3')
Biotin adapter	…TCGGTTGTACTGTGACCGATTC	Biotin-GAATCGGTCACAGTACAACCG

Supplementary Table 2: Primary probes targeting M13 phage scaffold

Probe Name	Staple Sequence (5'-3')	Extension (5'-3')
0[27]0[0]-biotin	CGGAGAGGGTAGCTATTTTTGAGAGATC	TCGGTTGTACTGTGACCGATTC
0[55]0[28]-biotin	TCAACCGTTCTAGCTGATAAATTAATGC	TCGGTTGTACTGTGACCGATTC
0[87]0[56]-E1	CGGAGACAGTCAAATCACCATCAATATGATAT	TTTCCTCAATTA
0[119]0[88]-E1	TGTAGGTAAAGATTCAAAAGGGTGAGAAAGGC	TTTCCTCAATTA
0[151]0[120]-E1	TCATATATTTTAAATGCAATGCCTGAGTAATG	TTTCCTCAATTA
0[183]0[152]	TTTCAACGCAAGGATAAAAATTTTTAGAACCC	
0[215]0[184]	ACCCTGTAATACTTTTGCGGGAGAAGCCTTTA	
0[247]0[216]-E0	AAAGCTAAATCGGTTGTACCAAAAACATTATG	TTATACATCTA
0[279]0[248]	TAGCAAAATTAAGCAATAAAGCCTCAGAGCAT	
0[311]0[280]-E0	ATCCAATAAATCATACAGGCAAGGCAAAGAAT	TTATACATCTA
0[343]0[312]	GCATCAATTCTACTAATAGTAGTAGCATTAAC	
0[375]0[344]-E0	TATTTTCATTTGGGGCGCGAGCTGAAAAGGTG	TTATACATCTA
0[407]0[376]	ATTTCGCAAATGGTCAATAACCTGTTTAGCTA	
0[439]0[408]-E0	GAACGAGTAGATTTAGTTTGACCATTAGATAC	TTATACATCTA
0[471]0[440]	CATTCCATATAACAGTTGATTCCCAATTCTGC	
0[503]0[472]-E0	ATATGCAACTAAAGTACGGTGTCTGGAAGTTT	TTATACATCTA
0[535]0[504]	TGAATATAATGCTGTAGCTCAACATGTTTTAA	
0[567]0[536]-E0	TCATTTTTGCGGATGGCTTAGAGCTTAATTGC	TTATACATCTA
0[599]0[568]	GAGTACCTTTAATTGCTCCTTTTGATAAGAGG	
0[631]0[600]-E0	CCGGAAGCAAACTCCAACAGGTCAGGATTAGA	
0[659]0[632]-biotin	TTTAATTCGAGCTTCAAAGCGAACCAGA	TCGGTTGTACTGTGACCGATTC
0[687]0[660]-biotin	GAAGCCCGAAAGACTTCAAATATCGCGT	TCGGTTGTACTGTGACCGATTC
0[719]0[688]-E2	AGCAAAGCGGATTGCATCAAAAAGATTAAGAG	TTACAATTTTCC

0[751]0[720]-E2	ATCAGGTCTTTACCCTGACTATTATAGTCAGA	TTACAATTTTCC
0[783]0[752]-E2	GTTCAGAAAACGAGAATGACCATAAATCAAAA	TTACAATTTTCC
0[815]0[784]	ATTCATTGAATCCCCCTCAAATGCTTTAAACA	
0[847]0[816]	ATAGCGTCCAATACTGCGGAATCGTCATAAAT	
0[879]0[848]-E0	CAGAGGGGGTAATAGTAAAATGTTTAGACTGG	TTATACATCTA
0[911]0[880]	ATAGCGAGAGGCTTTTGCAAAAGAAGTTTTGC	
0[943]0[912]-E0	CCCTCGTTTACCAGACGACGATAAAAACCAAA	TTATACATCTA
0[975]0[944]	TACGAGGCATAGTAAGAGCAACACTATCATAA	
0[1007]0[976]-E0	AACTAATGCAGATACATAACGCCAAAAGGAAT	TTATACATCTA
0[1039]0[1008]	TTCATCAGTTGAGATTTAGGAATACCACATTC	
0[1071]0[1040]-E0	CTAACGGAACAACATTATTACAGGTAGAAAGA	TTATACATCTA
0[1103]0[1072]	TTGGGAAGAAAAATCTACGTTAATAAAACGAA	
0[1135]0[1104]-E0	TTAAGAACTGGCTCATTATACCAGTCAGGACG	TTATACATCTA
0[1167]0[1136]	ACTTTAATCATTGTGAATTACCTTATGCGATT	
0[1199]0[1168]-E0	TAGTAAATTGGGCTTGAGATGGTTTAATTTCA	TTATACATCTA
0[1231]0[1200]	AGGCTTGCCCTGACGAGAAACACCAGAACGAG	
0[1263]0[1232]	TCAACGTAACAAAGCTGCTCATTCAGTGAATA	
0[1291]0[1264]-biotin	GACAAGAACCGGATATTCATTACCCAAA	TCGGTTGTACTGTGACCGATTC
0[1319]0[1292]-biotin	CTGGCTGACCTTCATCAAGAGTAATCTT	TCGGTTGTACTGTGACCGATTC
0[1351]0[1320]-E3	CAGATGAACGGTGTACAGACCAGGCGCATAGG	TTATTTTACACC
0[1383]0[1352]-E3	AGGGAACCGAACTGACCAACTTTGAAAGAGGA	TTATTTTACACC
0[1415]0[1384]-E3	TAGCCGGAACGAGGCGCAGACGGTCAATCATA	TTATTTTACACC
0[1447]0[1416]	GTGTCGAAATCCGCGACCTGCTCCATGTTACT	
0[1479]0[1448]	CAACGGAGATTTGTATCATCGCCTGATAAATT	
0[1511]0[1480]-E0	CCCAGCGATTATACCAAGCGCGAAACAAAGTA	TTATACATCTA
0[1543]0[1512]	AAAAGAATACACTAAAACACTCATCTTTGACC	
0[1575]0[1544]-E0	CTACGAAGGCACCAACCTAAAACGAAAGAGGC	TTATACATCTA
0[1607]0[1576]	TTTCCATTAAACGGGTAAAATACGTAATGCCA	
0[1639]0[1608]-E0	CTTTGAGGACTAAAGACTTTTTCATGAGGAAG	TTATACATCTA
0[1671]0[1640]	CATCGGAACGAGGGTAGCAACGGCTACAGAGG	

0[1703]0[1672]-E0	GCGGGATCGTCACCCTCAGCAGCGAAAGACAG	TTATACATCTA
0[1735]0[1704]	CTGAGGCTTGCAGGGAGTTAAAGGCCGCTTTT	
0[1767]0[1736]-E0	CATCGCCCACGCATAACCGATATATTCGGTCG	TTATACATCTA
0[1799]0[1768]	ATACCGATAGTTGCGCCGACAATGACAACAAC	
0[1831]0[1800]-E0	TTGCTTTCGAGGTGAATTTCTTAAACAGCTTG	TTATACATCTA
0[1863]0[1832]	AAAGGAGCCTTTAATTGTATCGGTTTATCAGC	
0[1895]0[1864]	TCACGTTGAAAATCTCCAAAAAAAAGGCTCCA	
0[1923]0[1896]-biotin	ACTAAAGGAATTGCGAATAATAATTTTT	TCGGTTGTACTGTGACCGATTC
0[1951]0[1924]-biotin	TTCAGCGGAGTGAGAATAGAAAGGAACA	TCGGTTGTACTGTGACCGATTC
0[1983]0[1952]-E4	GTATGGGATTTTGCTAAACAACTTTCAACAGT	TTTCTTATACAC
0[2015]0[1984]-E4	CGTCTTTCCAGACGTTAGTAAATGAATTTTCT	TTTCTTATACAC
0[2047]0[2016]-E4	CTCATAGTTAGCGTAACGATCTAAAGTTTTGT	TTTCTTATACAC
0[2079]0[2048]	ACTACAACGCCTGTAGCATTCCACAGACAGCC	
0[2111]0[2080]	TACCGTAACACTGAGTTTCGTCACCAGTACAA	
0[2143]0[2112]-E0	TTCAGGGATAGCAAGCCCAATAGGAACCCATG	TTATACATCTA
0[2175]0[2144]	GAACCGCCACCCTCAGAGCCACCACCCTCATT	
0[2207]0[2176]-E0	TAGTACCGCCACCCTCAGAACCGCCACCCTCA	TTATACATCTA
0[2239]0[2208]	CGGAATAGGTGTATCACCGTACTCAGGAGGTT	
0[2271]0[2240]-E0	GTGCCGTCGAGAGGGTTGATATAAGTATAGCC	TTATACATCTA
0[2303]0[2272]	TAGCGGGGTTTTGCTCAGTACCAGGCGGATAA	
0[2335]0[2304]-E0	AGGCTGAGACTCCTCAAGAGAAGGATTAGGAT	TTATACATCTA
0[2367]0[2336]	AACCTATTATTCTGAAACATGAAAGTATTAAG	
0[2399]0[2368]-E0	TATAAACAGTTAATGCCCCCTGCCTATTTCGG	TTATACATCTA
0[2431]0[2400]	AACGGGGTCAGTGCCTTGAGTAACAGTGCCCG	
0[2463]0[2432]-E0	GATGATACAGGAGTGTACTGGTAATAAGTTTT	TTATACATCTA
0[2495]0[2464]	TACCGTTCCAGTAAGCGTCATACATGGCTTTT	
0[2527]0[2496]	AAAGCCAGAATGGAAAGCGCAGTCTCTGAATT	
0[2555]0[2528]-biotin	GATATTCACAAACAAATAAATCCTCATT	TCGGTTGTACTGTGACCGATTC
0[2583]0[2556]-biotin	GGTTGAGGCAGGTCAGACGATTGGCCTT	TCGGTTGTACTGTGACCGATTC
0[2615]0[2584]-E5	CACCACCAGAGCCGCCGCCAGCATTGACAGGA	TTACTACTTATC

0[2647]0[2616]-E5	AGAGCCACCACCCTCAGAGCCGCCACCAGAAC	TTACTACTTATC
0[2679]0[2648]-E5	TCAGAGCCGCCACCCTCAGAACCGCCACCCTC	TTACTACTTATC
0[2711]0[2680]	CGGAACCAGAGCCACCACCGGAACCGCCTCCC	
0[2743]0[2712]	AGCGTTTGCCATCTTTTCATAATCAAAATCAC	
0[2775]0[2744]-E0	TCATCGGCATTTTCGGTCATAGCCCCCTTATT	TTATACATCTA
0[2807]0[2776]	GTTTGCCTTTAGCGTCAGACTGTAGCGCGTTT	
0[2839]0[2808]-E0	AGCAGCACCGTAATCAGTAGCGACAGAATCAA	TTATACATCTA
0[2871]0[2840]	AGGCCGGAAACGTCACCAATGAAACCATCGAT	
0[2903]0[2872]-E0	CAAAATCACCAGTAGCACCATTACCATTAGCA	TTATACATCTA
0[2935]0[2904]	ACCGACTTGAGCCATTTGGGAATTAGAGCCAG	
0[2967]0[2936]-E0	AAATTATTCATTAAAGGTGAATTATCACCGTC	TTATACATCTA
0[2999]0[2968]	CCGATTGAGGGAGGGAAGGTAAATATTGACGG	
0[3031]0[3000]-E0	ACCAGCGCCAAAGACAAAAGGGCGACATTCAA	TTATACATCTA
0[3063]0[3032]	TTGTCACAATCAATAGAAAATTCATATGGTTT	
0[3095]0[3064]-E0	GAAACGCAAAGACACCACGGAATAAGTTTATT	TTATACATCTA
0[3127]0[3096]	AAATACATACATAAAGGTGGCAACATATAAAA	
0[3159]0[3128]	TCCTTATTACGCAGTATGTTAGCAAACGTAGA	
0[3187]0[3160]-biotin	TACCCAAAAGAACTGGCATGATTAAGAC	TCGGTTGTACTGTGACCGATTC
0[3215]0[3188]-biotin	AAACCGAGGAAACGCAATAATAACGGAA	TCGGTTGTACTGTGACCGATTC
0[3247]0[3216]-E6	TAAGCAGATAGCCGAACAAAGTTACCAGAAGG	TTTAAATTTCCC
0[3279]0[3248]-E6	AGCTATCTTACCGAAGCCCTTTTTAAGAAAAG	TTTAAATTTCCC
0[3311]0[3280]-E6	ATAATAAGAGCAAGAAACAATGAAATAGCAAT	TTTAAATTTCCC
0[3343]0[3312]	AGAGATAACCCACAAGAATTGAGTTAAGCCCA	
0[3375]0[3344]	AAAGTCAGAGGGTAATTGAGCGCTAATATCAG	
0[3407]0[3376]-E0	TTAGACGGGAGAATTAACTGAACACCCTGAAC	TTATACATCTA
0[3439]0[3408]	CAGAGAGAATAACATAAAAACAGGGAAGCGCA	
0[3471]0[3440]-E0	TTTAACGTCAAAAATGAAAATAGCAGCCTTTA	TTATACATCTA
0[3503]0[3472]	TTATCCCAATCCAAATAAGAAACGATTTTTTG	
0[3535]0[3504]-E0	TTTGCCAGTTACAAAATAAACAGCCATATTAT	TTATACATCTA
0[3567]0[3536]	CCAACGCTAACGAGCGTCTTTCCAGAGCCTAA	

0[3599]0[3568]-E0	TGCACCCAGCTACAATTTTATCCTGAATCTTA	TTATACATCTA
0[3631]0[3600]	TTGAAGCCTTAAATCAAGATTAGTTGCTATTT	
0[3663]0[3632]-E0	GTTTTAGCGAACCTCCCGACTTGCGGGAGGTT	TTATACATCTA
0[3695]0[3664]	AAGGCTTATCCGGTATTCTAAGAACGCGAGGC	
0[3727]0[3696]-E0	ACCGCGCCCAATAGCAAGCAAATCAGATATAG	TTATACATCTA
0[3759]0[3728]	AAGCCGTTTTTATTTTCATCGTAGGAATCATT	
0[3791]0[3760]	TAAACCAAGTACCGCACTCATCGAGAACAAGC	
0[3819]0[3792]-biotin	CTTTCCTTATCATTCCAAGAACGGGTAT	TCGGTTGTACTGTGACCGATTC
0[3847]0[3820]-biotin	ATGTAGAAACCAATCAATAATCGGCTGT	TCGGTTGTACTGTGACCGATTC
0[3879]0[3848]-E7	AGAAAAATAATATCCCATCCTAATTTACGAGC	TTACTCTATTCA
0[3911]0[3880]-E7	CCTGTTTATCAACAATAGATAAGTCCTGAACA	TTACTCTATTCA
0[3943]0[3912]-E7	ATAAACAACATGTTCAGCTAATGCAGAACGCG	TTACTCTATTCA
0[3975]0[3944]	AAAGGTAAAGTAATTCTGTCCAGACGACGACA	
0[4007]0[3976]	AGCCAGTAATAAGAGAATATAAAGTACCGACA	
0[4039]0[4008]-E0	GCCAACATGTAATTTAGGCAGAGGCATTTTCG	TTATACATCTA
0[4071]0[4040]	GGGCTTAATTGAGAATCGCCATATTTAACAAC	
0[4103]0[4072]-E0	TCTTACCAGTATAAAGCCAACGCTCAACAGTA	TTATACATCTA
0[4135]0[4104]	AGCCTGTTTAGTATCATATGCGTTATACAAAT	
0[4167]0[4136]-E0	GAATAAACACCGGAATCATAATTACTAGAAAA	TTATACATCTA
0[4199]0[4168]	ACCGACCGTGTGATAAATAAGGCGTTAAATAA	
0[4231]0[4200]-E0	CATCTTCTGACCTAAATTTAATGGTTTGAAAT	TTATACATCTA
0[4263]0[4232]	GAAAACTTTTTCAAATATATTTTAGTTAATTT	
0[4295]0[4264]-E0	ATGCAAATCCAATCGCAAGACAAAGAACGCGA	TTATACATCTA
0[4327]0[4296]	AGGTTGGGTTATATAACTATATGTAAATGCTG	
0[4359]0[4328]-E0	GTCTGAGAGACTACCTTTTTAACCTCCGGCTT	TTATACATCTA
0[4391]0[4360]	AAGAGTCAATAGTGAATTTATCAAAATCATAG	
0[4423]0[4392]-E0	AACATAGCGATAGCTTAGATTAAGACGCTGAG	
0[4451]0[4424]-biotin	TAATTAATTTTCCCTTAGAATCCTTGAA	TCGGTTGTACTGTGACCGATTC
0[4479]0[4452]-biotin	TAACCTTGCTTCTGTAAATCGTCGCTAT	TCGGTTGTACTGTGACCGATTC
0[4511]0[4480]-E8	AAACAGTACATAAATCAATATATGTGAGTGAA	TTATCAATCTTC

0[4543]0[4512]-E8	AACAATTTCATTTGAATTACCTTTTTTAATGG	TTATCAATCTTC
0[4575]0[4544]-E8	CAAACATCAAGAAAACAAAATTAATTACATTT	TTATCAATCTTC
0[4607]0[4576]	TTCAATTACCTGAGCAAAAGAAGATGATGAAA	
0[4639]0[4608]	GTTACAAAATCGCGCAGAGGCGAATTATTCAT	
0[4671]0[4640]-E0	TAACGGATTCGCCTGATTGCTTTGAATACCAA	TTATACATCTA
0[4703]0[4672]	AGTAACAGTACCTTTTACATCGGGAGAAACAA	
0[4735]0[4704]-E0	AGATTTTCAGGTTTAACGTCAGATGAATATAC	TTATACATCTA
0[4767]0[4736]	TTGCACGTAAAACAGAAATAAAGAAATTGCGT	
0[4799]0[4768]-E0	GGAAGGGTTAGAACCTACCATATCAAAATTAT	TTATACATCTA
0[4831]0[4800]	TCCTGATTGTTTGGATTATACTTCTGAATAAT	
0[4863]0[4832]-E0	GATTATCAGATGATGGCAATTCATCAATATAA	TTATACATCTA
0[4895]0[4864]	ACCAGAAGGAGCGGAATTATCATCATATTCCT	
0[4927]0[4896]-E0	GTAACATTATCATTTTGCGGAACAAAGAAACC	TTATACATCTA
0[4959]0[4928]	TTGCCCGAACGTTATTAATTTTAAAAGTTTGA	
0[4991]0[4960]-E0	TACAAACAATTCGACAACTCGTATTAAATCCT	TTATACATCTA
0[5023]0[4992]	AATACATTTGAGGATTTAGAAGTATTAGACTT	
0[5055]0[5024]	TAACAACTAATAGATTAGAGCCGTCAATAGAT	
0[5083]0[5056]-biotin	AGGTTATCTAAAATATCTTTAGGAGCAC	TCGGTTGTACTGTGACCGATTC
0[5111]0[5084]-biotin	CAAATCAACAGTTGAAAGGAATTGAGGA	TCGGTTGTACTGTGACCGATTC
0[5143]0[5112]-E9	ATCAAACCCTCAATCAATATCTGGTCAGTTGG	TTTTTCTAAACC
0[5175]0[5144]-E9	AATCTAAAGCATCACCTTGCTGAACCTCAAAT	TTTTTCTAAACC
0[5207]0[5176]-E9	AGTGCCACGCTGAGAGCCAGCAGCAAATGAAA	TTTTTCTAAACC
0[5239]0[5208]	GTGAGGCGGTCAGTATTAACACCGCCTGCAAC	
0[5271]0[5240]	CGAACGAACCACCAGCAGAAGATAAAACAGAG	
0[5303]0[5272]-E0	TGATAGCCCTAAAACATCGCCATTAAAAATAC	TTATACATCTA
0[5335]0[5304]	TTTGAATGGCTATTAGTCTTTAATGCGCGAAC	
0[5367]0[5336]-E0	AAAGCGTAAGAATACGTGGCACAGACAATATT	TTATACATCTA
0[5399]0[5368]	CTGGCCAACAGAGATAGAACCCTTCTGACCTG	
0[5431]0[5400]-E0	CCAGTCACACGACCAGTAATAAAAGGGACATT	TTATACATCTA
0[5463]0[5432]	TCTGAAATGGATTATTTACATTGGCAGATTCA	

0[5495]0[5464]-E0	CATGGAAATACCTACATTTTGACGCTCAATCG	TTATACATCTA
0[5527]0[5496]	TTACCGCCAGCCATTGCAACAGGAAAAACGCT	
0[5559]0[5528]-E0	TATCGGCCTTGCTGGTAATATCCAGAACAATA	TTATACATCTA
0[5591]0[5560]	AACATCACTTGCCTGAGTAGAAGAACTCAAAC	
0[5623]0[5592]-E0	ACCGTTGTAGCAATACTTCTTTGATTAGTAAT	TTATACATCTA
0[5655]0[5624]	CGAGTAAAAGAGTCTGTCCATCACGCAAATTA	
0[5687]0[5656]	CTGAGAAGTGTTTTTATAATCAGTGAGGCCAC	
0[5715]0[5688]-biotin	TTTTAGACAGGAACGGTACGCCAGAATC	TCGGTTGTACTGTGACCGATTC
0[5743]0[5716]-biotin	GAGCTAAACAGGAGGCCGATTAAAGGGA	TCGGTTGTACTGTGACCGATTC
0[5775]0[5744]-E10	ATAACGTGCTTTCCTCGTTAGAATCAGAGCGG	TTTCAATATCTC
0[5807]0[5776]-E10	GGCGCGTACTATGGTTGCTTTGACGAGCACGT	TTTCAATATCTC
0[5839]0[5808]-E10	CCACACCCGCCGCGCTTAATGCGCCGCTACAG	TTTCAATATCTC
0[5871]0[5840]	GGCAAGTGTAGCGGTCACGCTGCGCGTAACCA	
0[5903]0[5872]	AAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCT	
0[5935]0[5904]-E0	GAAAGCCGGCGAACGTGGCGAGAAAGGAAGGG	TTATACATCTA
0[5967]0[5936]	TAAAGGGAGCCCCCGATTTAGAGCTTGACGGG	
0[5999]0[5968]-E0	TCGAGGTGCCGTAAAGCACTAAATCGGAACCC	TTATACATCTA
0[6031]0[6000]	GTGAACCATCACCCAAATCAAGTTTTTTGGGG	
0[6063]0[6032]-E0	AAAAACCGTCTATCAGGGCGATGGCCCACTAC	TTATACATCTA

Supplementary Table 3: Genomic coordinates in the GRCh37 reference genome of loop 1-10 primary probes

Chromosome	Start	End	Locus Name	Docking Handle Name
chr8	128581495	128586495	Loop1	E1
chr8	128591521	128596521	Loop2	E2
chr8	128601094	128606094	Loop3	E3
chr8	128610854	128615854	Loop4	E4
chr8	128621643	128626643	Loop5	E5
chr8	128632099	128637099	Loop6	E6
chr8	128641830	128646830	Loop7	E7
chr8	128650562	128655562	Loop8	E8
chr8	128660917	128665917	Loop9	E9
chr8	128671629	128676629	Loop10	E10

Supplementary Table 4: Primary probes and docking handles targeting MYC loop 1

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop1_1_E1	cttcctgttcaattccctagtctctattgcccctacaacataaatccaat	TTATACATCTACGG
MYC_Loop1_2_E1	agctcaaaatcttctgcacttagagaaacatcccaagtcccctaccacagc	TTATACATCTACGG
MYC_Loop1_3_E1	tcctctgccgaaacacagtttcttcaacatggtcacccctcatttcttcc	TTATACATCTACGG
MYC_Loop1_4_E1	agcacaggcattgctaaatatgctccttgatttgctatcacctgttgccc	TTATACATCTACGG
MYC_Loop1_5_E1	cccaggccatacatacagcaagcgtctcctgagtaagctgctctgggcgc	TTATACATCTACGG
MYC_Loop1_6_E1	agaaggtaccaccttctctttttctcccagttctgatccatgtctgggca	TTATACATCTACGG
MYC_Loop1_7_E1	tacctgctgagaggctcattcctgagccctgtatcagcaccaggactgc	TTATACATCTACGG
MYC_Loop1_8_E1	agaaagcatcctgccttggctaaatcacactcctcaaaccatttcctggc	TTATACATCTACGG
MYC_Loop1_9_E1	caacttgaaaggtgacaactatataggttaacccttgtctgcactccagc	TTATACATCTACGG
MYC_Loop1_10_E1	accttactcgagaacctccaattcctctttcctggagtttccaggggcta	TTATACATCTACGG
MYC_Loop1_11_E1	cagtcctttcctgactctctctatccggactctccttgctccattcctgt	TTATACATCTACGG
MYC_Loop1_12_E1	tgtggtcattctcccagcatgccataccctctaaagcctccacacctctg	TTATACATCTACGG
MYC_Loop1_13_E1	ttcccctctcctttgattaataagttacattcatctttcaagcctcagct	TTATACATCTACGG
MYC_Loop1_14_E1	cctctatgaagtatttcctgacacctccttcgagtgggttccagtccagg	TTATACATCTACGG
MYC_Loop1_15_E1	ctgggacaaattctatcatatatcggtagcattttatttcacttcataa	TTATACATCTACGG
MYC_Loop1_16_E1	acctacacaatctttatcctgactatcctgtccacccatcagctctctgc	TTATACATCTACGG
MYC_Loop1_17_E1	ctgtccctcagttatttatgtgaccatctccctcacctcctcctatccct	TTATACATCTACGG
MYC_Loop1_18_E1	tgtcccttctccaccttgcttttctctatagacttcatcaccatccaaca	TTATACATCTACGG
MYC_Loop1_19_E1	ttatttatctattttggttattgtccatattgcctcactacagcttcaga	TTATACATCTACGG
MYC_Loop1_20_E1	ttttgtctgtttacctcgttattagagactcaatgctttgaatagtgcct	TTATACATCTACGG
MYC_Loop1_21_E1	aggactgaatatttttttgaaggaatgaatagtatctctagcatctagcat	TTATACATCTACGG
MYC_Loop1_22_E1	catagtagacatccagtaagtaattgggaataatgaatgaaagaagtcag	TTATACATCTACGG
MYC_Loop1_23_E1	gagaagaaaggaaggaagaaaagacagaagggaaggagaaaagaagagag	TTATACATCTACGG
MYC_Loop1_24_E1	gaaggaagcaagaaagaaacagatacagtagggagacagaaatacaaaga	TTATACATCTACGG
MYC_Loop1_25_E1	ttatttgcctaaagtcctcagacaattgtgtgtgttgtgggcttagacag	TTATACATCTACGG

MYC_Loop1_26_E1	ggcaaaagtaattgctgtatttactgttaaaattaaagggaaaactgcaa	TTATACATCTACGG
MYC_Loop1_27_E1	ttagatttcctgggaaatttctctgccctcagcttctttttttttgcca	TTATACATCTACGG
MYC_Loop1_28_E1	cccagagatgaaaacatgaatgttttacaaggatgcttggtcattagttt	TTATACATCTACGG
MYC_Loop1_29_E1	tggcaaattctaaacttcagcttgctacctgggatacctgcatgcatgca	TTATACATCTACGG
MYC_Loop1_30_E1	cctttggccaagagctcttcacttaatattttaaagagatttttgttttt	TTATACATCTACGG
MYC_Loop1_31_E1	ccagccttaaaaagaatgcatttctttataattgtagaatacacacaaag	TTATACATCTACGG
MYC_Loop1_32_E1	cttcttatgtctggagttctactattgctggatacgtgcttaggacatag	TTATACATCTACGG
MYC_Loop1_33_E1	gcaggaaggtgtacactctggaaactgacatacttcagaccagtccaaac	TTATACATCTACGG
MYC_Loop1_34_E1	gtggctgaagaccctagttaaattacgtaagtaaagtgatgttcccccat	TTATACATCTACGG
MYC_Loop1_35_E1	actcagggactgtccgagtttacacctgctgtttcagtgtaattaataag	TTATACATCTACGG
MYC_Loop1_36_E1	aatgctcaaaatgtcccagtttaaacaataagttatatggtccttctatc	TTATACATCTACGG
MYC_Loop1_37_E1	gagctttgcttttctctactgtaaaaatatagaaaacaatatccatttcc	TTATACATCTACGG
MYC_Loop1_38_E1	tatctttgctgtgtaacacaaactacccccaaatttaatggcttaaaaca	TTATACATCTACGG
MYC_Loop1_39_E1	accttggggcttctcatgacatgtgaatgatccaagagagaaaaagaggg	TTATACATCTACGG
MYC_Loop1_40_E1	agctgccaatatttctaatctcaaaagtgagaacgcagcctggtgccgtg	TTATACATCTACGG
MYC_Loop1_41_E1	gtgagacctcgtctctgcgaaaaagagaactaaccaagtatggtggcatg	TTATACATCTACGG
MYC_Loop1_42_E1	aagtgagaaaccaccaattctgccacatagacccaccctggtacaatgtg	TTATACATCTACGG
MYC_Loop1_43_E1	cacagaacttaacgtaaatactgcttgtaaagtccctggaacatagcaga	TTATACATCTACGG
MYC_Loop1_44_E1	tgttagctacttcagagcaggatcagacaagaaaatcttctaaaaataat	TTATACATCTACGG
MYC_Loop1_45_E1	gtaagttcttcacagctaaggatttaaatattatgtctttcttataagcc	TTATACATCTACGG
MYC_Loop1_46_E1	tccacaaatatgtattatatacctactgcttgttattatgtccttgaagc	TTATACATCTACGG
MYC_Loop1_47_E1	cacaactaactactgagttgttgtagatccttaaactcattacttatata	TTATACATCTACGG
MYC_Loop1_48_E1	tcttaaacttctgaagatgtaatcttagcatgtcatttgagtgtcgctgt	TTATACATCTACGG

Supplementary Table 5: Primary probes and docking handles targeting MYC loop 2

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop2_1_E2	tcttccttcagggaaatcagtctttttctataaaggccttcaactgattg	TTTCTTCATTAGCG
MYC_Loop2_2_E2	tccactgatttaaatgtctacctcatctaaatgagagaggtgtagaataa	TTTCTTCATTAGCG
MYC_Loop2_3_E2	atatccaggtctcatggcctagtcaaattgaacacataaaattaacagtc	TTTCTTCATTAGCG
MYC_Loop2_4_E2	tgagactagccctcatcttctcccttctcaccacttcttcctaacaaatt	TTTCTTCATTAGCG
MYC_Loop2_5_E2	tctctaacagtatcctgtgctctaaactagtggcttccaacttttggctgc	TTTCTTCATTAGCG
MYC_Loop2_6_E2	tggaatgccacttgcaattcatagtttcctacaactataattggtagcat	TTTCTTCATTAGCG
MYC_Loop2_7_E2	caaaaggtccagggcaactctaggcatgattggtacttaacaaaaagtta	TTTCTTCATTAGCG
MYC_Loop2_8_E2	atgaaatcagaatttctgcaggtaagtcccatgcatttagtttttaactcc	TTTCTTCATTAGCG
MYC_Loop2_9_E2	agcagaaggtcagatatctcatctgcttctatctagtgatggattacatg	TTTCTTCATTAGCG
MYC_Loop2_10_E2	gagtttcctatagaaacaacaccctgattaatacacaatacatgacccaa	TTTCTTCATTAGCG
MYC_Loop2_11_E2	cccagacagaccagtgggtgaagggtgtaatttaggatactatcccgcga	TTTCTTCATTAGCG
MYC_Loop2_12_E2	cagaagctcacctgcttgaaggatcagccttactcatgccctatgtgctc	TTTCTTCATTAGCG
MYC_Loop2_13_E2	catattcatgagccatggatgtcaatggcagcaagaaagaagagaaaata	TTTCTTCATTAGCG
MYC_Loop2_14_E2	ccagtcggatctgctgagtcaatcatggaatccaatgagacacgaattca	TTTCTTCATTAGCG
MYC_Loop2_15_E2	taaacaccagtagtaataagaaccatcgttaattgggtgtcttgccagac	TTTCTTCATTAGCG
MYC_Loop2_16_E2	ggctgcccaattaagaaaaccagaatatgatttgcaaagcatcttcacat	TTTCTTCATTAGCG
MYC_Loop2_17_E2	gccattttactaagtattcatttcacagacaaaaacccaagatgccaata	TTTCTTCATTAGCG
MYC_Loop2_18_E2	gacttatcccaagccacatacccggttagtcatgcaaccaggactcatgc	TTTCTTCATTAGCG
MYC_Loop2_19_E2	cagaaagttgaactgctgttatctggctctttattaaggtatattaatat	TTTCTTCATTAGCG
MYC_Loop2_20_E2	caccaacttaaaataactcccctaactttcttcacaattcctcctccccc	TTTCTTCATTAGCG
MYC_Loop2_21_E2	tgggctggctcccaacgactacattcgtacaatatgataccccaaccctc	TTTCTTCATTAGCG
MYC_Loop2_22_E2	aaaccaacaatggatctctgactgagaatgaaccaatcagattcttttcc	TTTCTTCATTAGCG
MYC_Loop2_23_E2	aaattacactgagtgtttctttctccatgtggtcctagctgtaacacatg	TTTCTTCATTAGCG
MYC_Loop2_24_E2	tagtagtaggtaacaataaccattggtaatgcttgccaaatgtgcccagc	TTTCTTCATTAGCG
MYC_Loop2_25_E2	gtttgtgagcagatgtgacctgtgcacttctagacaggatcacatcaaaa	TTTCTTCATTAGCG
MYC_Loop2_26_E2	ctctctacctctctcttacgaatcagaaatctttgaggtggtggctgctt	TTTCTTCATTAGCG

MYC_Loop2_27_E2	acattggacatgtaatgtgagccaagaataaaatccttatttttgtaaac	TTTCTTCATTAGCG
MYC_Loop2_28_E2	gggggttgttcgttaatatggcatagtctagctcattctgactgatagtt	TTTCTTCATTAGCG
MYC_Loop2_29_E2	agtaaacccttataaaccaatagggtttgggaattataactacagcatag	TTTCTTCATTAGCG
MYC_Loop2_30_E2	aattggcaaagtatagtctgtgggctatagttttataaactacatgcaat	TTTCTTCATTAGCG
MYC_Loop2_31_E2	gactgattacaacaaaaataagaacaatagcaaatacttgtgtagtgacc	TTTCTTCATTAGCG
MYC_Loop2_32_E2	ttgacacacattaactcactcaatccacacatcaaccagtttccctgcct	TTTCTTCATTAGCG
MYC_Loop2_33_E2	cattttcacataagcaaatcaaaccaagtgagttttcaatcattcagtg	TTTCTTCATTAGCG
MYC_Loop2_34_E2	cttatctcctacctgttataaaaatttatccaaattccttatctctgtgc	TTTCTTCATTAGCG
MYC_Loop2_35_E2	catctgtaaagcaggggcgatatcagtaactacctcatagacttattaga	TTTCTTCATTAGCG
MYC_Loop2_36_E2	gctgcaaagtttctgttgcctagtgatttgttacaactgcctacagtatt	TTTCTTCATTAGCG
MYC_Loop2_37_E2	tgtaagtaccctactctctacgaggttcccacaaagacagaattgcttga	TTTCTTCATTAGCG
MYC_Loop2_38_E2	tcagaacaagtccccatgatcaagcaacacatgactgtattactcttaat	TTTCTTCATTAGCG
MYC_Loop2_39_E2	caacactgaaagcagactcaagtatctttatttaacaaatgagaatattc	TTTCTTCATTAGCG
MYC_Loop2_40_E2	taataccaatacatagtagggactcaatgagtatgcatgcaataaatgaa	TTTCTTCATTAGCG
MYC_Loop2_41_E2	ctcatacctgacactgacttctaggatccttgctaggggacatcccctac	TTTCTTCATTAGCG
MYC_Loop2_42_E2	tttctcctgccagataaatgcagttgagtgattccttctgaagccatttc	TTTCTTCATTAGCG
MYC_Loop2_43_E2	cgctataaagaaggtcattcagcaaactagatggtagttactgtgtcact	TTTCTTCATTAGCG
MYC_Loop2_44_E2	gatattttgaattccagagaaaatggtgttaggtaaagagtgacatttta	TTTCTTCATTAGCG
MYC_Loop2_45_E2	tctcatttacttttaaattgaacttaggtgtgtttgaacttgtccctgaa	TTTCTTCATTAGCG
MYC_Loop2_46_E2	tcaaaacacaaacacacagtttgtaaagaaagtttgcctcctgtgagacc	TTTCTTCATTAGCG
MYC_Loop2_47_E2	gtaatgaagatttttcttcagacacactttctccttctgtcactgcttag	TTTCTTCATTAGCG
MYC_Loop2_48_E2	tgtacctacattctctctgaaatgttacctctgtgaatctcatcttgccc	TTTCTTCATTAGCG

Supplementary Table 6: Primary probes and docking handles targeting MYC loop 3

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop3_1_E3	atcaaggcaattggatgccaggtctgagttcttcatttcagaaccattct	TTTCAATGTATGGC
MYC_Loop3_2_E3	gcttagtttaagaacaaactttctgacagaaatggcatcccacaaagcac	TTTCAATGTATGGC
MYC_Loop3_3_E3	caggctgcgggaggttaagctgaggctctagttcctctgatgccagctct	TTTCAATGTATGGC
MYC_Loop3_4_E3	acacatttgtggagtgtaaacacactttcatgttctaaaggccactgtca	TTTCAATGTATGGC
MYC_Loop3_5_E3	tgggctcaatcagtgaccgcacaagaatgtgggcatgctgaatccctgcg	TTTCAATGTATGGC
MYC_Loop3_6_E3	gcaccggggattggaagagataagccagagaccagcaagattcaaagcag	TTTCAATGTATGGC
MYC_Loop3_7_E3	ttggtcagaggaagctgttgagcagccattgaatgggaagcaggtgaag	TTTCAATGTATGGC
MYC_Loop3_8_E3	gttaaagttgaaagcattcttctccctccttaatagagatgtggagatca	TTTCAATGTATGGC
MYC_Loop3_9_E3	ggctgaactatttgcaggctataactgacggatccaggtttcaaacccag	TTTCAATGTATGGC
MYC_Loop3_10_E3	ttcactgcctttacttaaagtgaccgcttttgtaagcaatccattcattt	TTTCAATGTATGGC
MYC_Loop3_11_E3	actaaggataaaagaaatagtacttacctggtgggaataagatgatatag	TTTCAATGTATGGC
MYC_Loop3_12_E3	cacctggttcctggcaccaaacaaaagcttggtgagagctaaatataagt	TTTCAATGTATGGC
MYC_Loop3_13_E3	atttgtttttttttctccaactagattctaagttcatgtttggagacacac	TTTCAATGTATGGC
MYC_Loop3_14_E3	tattttgcatgccctacaatgcctagaacaatactttctatctgtaaggg	TTTCAATGTATGGC
MYC_Loop3_15_E3	tgcactgcacaaaaacagtctttactagactgcaaaatctgaaagaatcc	TTTCAATGTATGGC
MYC_Loop3_16_E3	taaggctttacattgagaaactggtgttgttttaaaacctcagcccctgc	TTTCAATGTATGGC
MYC_Loop3_17_E3	cagaaccagactctgtctcagaagaatgaaagaaaaagagagagagagaa	TTTCAATGTATGGC
MYC_Loop3_18_E3	aggctaaactggcagggttgtcacaaaagctggaaattgtatccgtttca	TTTCAATGTATGGC
MYC_Loop3_19_E3	cacatctcttagatgctgcttggtttaaggtggttgattcacttttgtgtg	TTTCAATGTATGGC
MYC_Loop3_20_E3	gacttctgcctgttcctaaattgtgaaaccctagtaactcgttctattat	TTTCAATGTATGGC
MYC_Loop3_21_E3	catccctttaaacacaccttaaaatgtagggaatgaagtctctcaagaaa	TTTCAATGTATGGC
MYC_Loop3_22_E3	ggtggctggtcttctgaatatcttatcacagttaaagtaggaacatcaga	TTTCAATGTATGGC
MYC_Loop3_23_E3	cagtgaccaaagtaccccatatttgtctagctccagcacatatacggttt	TTTCAATGTATGGC
MYC_Loop3_24_E3	aaccataacaccggagaaatgtccttgttagacccatcgtacaaaagagg	TTTCAATGTATGGC
MYC_Loop3_25_E3	caactccaaaccaaagttcataccttaggtgttattgcagcaactttatt	TTTCAATGTATGGC
MYC_Loop3_26_E3	ctttcctatctgaccgacgattgatctctttcatattgtaaagttcaaac	TTTCAATGTATGGC

MYC_Loop3_27_E3	caatgagtttcaccttcctcttaatccccaacaattacttttt cactttg	TTTCAATGTATG GC
MYC_Loop3_28_E3	ttcagccctgctgaacaagatgcttccttcttagaatacccca ggtgaga	TTTCAATGTATG GC
MYC_Loop3_29_E3	ggattctcgtgtgagcactaatgttaatctgctacacggctca cttctcc	TTTCAATGTATG GC
MYC_Loop3_30_E3	cccaacccttaaccataccatcgggttgttctcaaactgagat aacagac	TTTCAATGTATG GC
MYC_Loop3_31_E3	tgaaaacgttacaagcactatacaaatgtaggcatcattatta ccaaggc	TTTCAATGTATG GC
MYC_Loop3_32_E3	ttatcacactgataaatggtgagtatgtttctgtgcctctgta gtttcca	TTTCAATGTATG GC
MYC_Loop3_33_E3	aattgcagttggcagataaatccaaggcagttaaccaaaccca ggagtat	TTTCAATGTATG GC
MYC_Loop3_34_E3	cacaagttggccaaatccaaatgattttaaaacacgctataaa aataaac	TTTCAATGTATG GC
MYC_Loop3_35_E3	ataataattccagtttcgtcaatccttcctgctacttggagca atgtagc	TTTCAATGTATG GC
MYC_Loop3_36_E3	caattacccttttggatccttttaatcccagctgatcattttca gctttaa	TTTCAATGTATG GC
MYC_Loop3_37_E3	ggtttgaatgttgttgtagttcaaatgtagccaattagcatgg cctcagc	TTTCAATGTATG GC
MYC_Loop3_38_E3	aaagaacactccttaaaatccctggtggactttgtcccagatt gtaccag	TTTCAATGTATG GC
MYC_Loop3_39_E3	cacaccctcaggtgcttattaagtagaaaaaaaggccagcaga ggctcag	TTTCAATGTATG GC
MYC_Loop3_40_E3	agatcaagttttcataatgctggtgtcaaaatgtggtccagga tgcctgg	TTTCAATGTATG GC
MYC_Loop3_41_E3	tgattaagcagatggcacaggagagattcacacatctggatgg ggctggg	TTTCAATGTATG GC
MYC_Loop3_42_E3	ggtggcctttgtgatcacaaagatttctctctccccccttgg gtgtctg	TTTCAATGTATG GC
MYC_Loop3_43_E3	cttggcatttaggatgtataggctcacatcctggttctgcaat ttactgg	TTTCAATGTATG GC
MYC_Loop3_44_E3	aagagaatcttacaggtgtgaacagggaacatcgtttcagggc ttattct	TTTCAATGTATG GC
MYC_Loop3_45_E3	aatgattattacttaactcttctgtggcttaatttcctcatct ggaaaat	TTTCAATGTATG GC
MYC_Loop3_46_E3	tagcattggcttcatgcagttgttctgaatattaaatgagcaa gcacatg	TTTCAATGTATG GC
MYC_Loop3_47_E3	gcacagtgtctgctatacataactgttcagttccttcagccgc ctcgtca	TTTCAATGTATG GC
MYC_Loop3_48_E3	attctctctttcggtggtgtcataactactgttcctgactcct tcctcgc	TTTCAATGTATG GC

Supplementary Table 7: Primary probes and docking handles targeting MYC loop 4

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop4_1_E4	tgcaatttagataggcatatcctctttctaatcatcattcgtccatcagc	TTAAAAAGTTCGAG
MYC_Loop4_2_E4	agttgagtttaatcttcctcgatggctgtacttttggtttcagtctctca	TTAAAAAGTTCGAG
MYC_Loop4_3_E4	cctcctgtaacttccattcagcatgcaatgagagaggcatgtgtctagag	TTAAAAAGTTCGAG
MYC_Loop4_4_E4	acggtggttctgctttgttaaaacaaaaaaggcccattgtaacactggtt	TTAAAAAGTTCGAG
MYC_Loop4_5_E4	gtcagccttgtctacacttccacaaatttacccacttgaatacgctgtag	TTAAAAAGTTCGAG
MYC_Loop4_6_E4	aaactcacaatgtttctgcagaaaggactgtttacatggaggttttcagc	TTAAAAAGTTCGAG
MYC_Loop4_7_E4	gcatgatgtttggatcctggaaatcatctatttttatgtgtaaataaac	TTAAAAAGTTCGAG
MYC_Loop4_8_E4	gagaaggcatccatacaattttaatgaacaagattcccatcagtcatcaa	TTAAAAAGTTCGAG
MYC_Loop4_9_E4	aggtagcttaaagggcacaaggaagctattcttctagtgatttctctgct	TTAAAAAGTTCGAG
MYC_Loop4_10_E4	ggaaaaggagtggcttctaaattaaaaaaaaagaaaatcatgctctagta	TTAAAAAGTTCGAG
MYC_Loop4_11_E4	ttgtctttgatcttgcatttcctaagaaagaagcacttaacctctagtat	TTAAAAAGTTCGAG
MYC_Loop4_12_E4	tggaaatgcacgctttattttaagatccttttctaacacatggaaactct	TTAAAAAGTTCGAG
MYC_Loop4_13_E4	agtgaacttaacatgattcaaacagattgcattccacaaccaacgtgctt	TTAAAAAGTTCGAG
MYC_Loop4_14_E4	gctgaattcaggcataaagctggtttttttcatcaccactaatgaattat	TTAAAAAGTTCGAG
MYC_Loop4_15_E4	caagtcttttccaaataaaagttacttaactaaattcgagccatgatgtt	TTAAAAAGTTCGAG
MYC_Loop4_16_E4	tggacactacccaagtggattctccttgcaccaaaatggcagattctcca	TTAAAAAGTTCGAG
MYC_Loop4_17_E4	cccttcccagtgaatgaagtgtattatgttggcagaaggatactccatca	TTAAAAAGTTCGAG
MYC_Loop4_18_E4	tattcatctgcatgtatcatgcacttgagctgcatgactgatgaactccg	TTAAAAAGTTCGAG
MYC_Loop4_19_E4	gaaagggaaatgcaccaagacaaatgtcacctttcaatgcttctccaaca	TTAAAAAGTTCGAG
MYC_Loop4_20_E4	gttagatttaaacaaagtcagtaatttcatgctttccaactggtgcctga	TTAAAAAGTTCGAG
MYC_Loop4_21_E4	tggcccctactctgtttaagggtgctaatacccaggttggttagttgagt	TTAAAAAGTTCGAG
MYC_Loop4_22_E4	ctactagagaacagtcaacaatttctacttcatatgcttattgaaataca	TTAAAAAGTTCGAG
MYC_Loop4_23_E4	acacttccccaggtatttatgcacgttggataatttaaaatttgggcagc	TTAAAAAGTTCGAG
MYC_Loop4_24_E4	cataggtatagctcttttgaaaaataatttggcactatgcgataaaaatg	TTAAAAAGTTCGAG
MYC_Loop4_25_E4	agcgtctgttttcttttgatgaggggaaattccacatcccagaagatact	TTAAAAAGTTCGAG
MYC_Loop4_26_E4	gaagaatgctcatacatatgttgggctaattttctttctccatgctccat	TTAAAAAGTTCGAG

MYC_Loop4_27_E4	gtttggtctgtactgaacaataagagatgtgaaattcaattcagaaaaaa	TTAAAAAGTTCGAG
MYC_Loop4_28_E4	attggtatacctgataccttattctgtatctaactggccttaacatctct	TTAAAAAGTTCGAG
MYC_Loop4_29_E4	atcagatcttgtgagagaaagagaaaggtttcatgtcttagcccttatca	TTAAAAAGTTCGAG
MYC_Loop4_30_E4	gcttcagacttaaagatgtccactgtgaaattatttgttacctaacaatg	TTAAAAAGTTCGAG
MYC_Loop4_31_E4	tgagcaaaatccactcagtgtgatgcataacactaggcagccactacaga	TTAAAAAGTTCGAG
MYC_Loop4_32_E4	caatgcctagaatatacttgataatattgaagcaaaaaacagaaggatag	TTAAAAAGTTCGAG
MYC_Loop4_33_E4	aaacattgataattatgaaaaaaggttggtcaagcttatcagagcgtgca	TTAAAAAGTTCGAG
MYC_Loop4_34_E4	tggatcctaggaggggaatacccagtctcaccactctgacatgttctttg	TTAAAAAGTTCGAG
MYC_Loop4_35_E4	gaaaacatacatagtttaatataagccgtaatgtaaagtatgtgacaagg	TTAAAAAGTTCGAG
MYC_Loop4_36_E4	caacccatacattcaacaacaatgatgataccaaaaatatccattaaaac	TTAAAAAGTTCGAG
MYC_Loop4_37_E4	aatgcttgtgtattgctgaatacaaaatggaaaacccaggacacaaaaat	TTAAAAAGTTCGAG
MYC_Loop4_38_E4	aaatgatgagagtatgttctgaattgggtattactgctagtatctgattc	TTAAAAAGTTCGAG
MYC_Loop4_39_E4	ggtcctaacgatatgtgaaatttgatgggaattacatttaaaattgtact	TTAAAAAGTTCGAG
MYC_Loop4_40_E4	caaagaattagagggatgttaaaatgctgggattatgacttctgttatgt	TTAAAAAGTTCGAG
MYC_Loop4_41_E4	ttggaagtgtatccttgatcatttgactatggaacaagaattctgtgcaa	TTAAAAAGTTCGAG
MYC_Loop4_42_E4	tcttagggagtgagtccttaggccaggtaaatatcttttatgtctttggt	TTAAAAAGTTCGAG
MYC_Loop4_43_E4	atgaggaaattaaggccagaaatttgatgaagtcctatcattagtagaac	TTAAAAAGTTCGAG
MYC_Loop4_44_E4	tttttctgtttgaggaaacacttataatgttgactccattgccctttatt	TTAAAAAGTTCGAG
MYC_Loop4_45_E4	cctagtccctaatacttagtagatacccaataaatgaacatttgagttct	TTAAAAAGTTCGAG
MYC_Loop4_46_E4	ttgtcctctgattgcttcagagacctttgacgtgatttcacagcaagaat	TTAAAAAGTTCGAG
MYC_Loop4_47_E4	tggcagccaggatgactttggccaatattcttttagcatcttcccaacac	TTAAAAAGTTCGAG
MYC_Loop4_48_E4	gactaggtacactaccaatgaattgattgacagttgaagagaaaatttaa	TTAAAAAGTTCGAG

Supplementary Table 8: Primary probes and docking handles targeting MYC loop 5

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop5_1_E5	aaaaagaaaatcactcataaaacacttttttcattgcaaaacccctcccaa	TTTAGTTAGAGCCC
MYC_Loop5_2_E5	cattaagtaattattgagtatccactatgtggcatgcactacatagtttt	TTTAGTTAGAGCCC
MYC_Loop5_3_E5	acacggtccctaccctccatggactttataaagtagtggaactaaattcc	TTTAGTTAGAGCCC
MYC_Loop5_4_E5	aacagtgctcactggggtaataactactaacataaaaagtcatggttgac	TTTAGTTAGAGCCC
MYC_Loop5_5_E5	tgtagctaattacttaagttgatctccacaggtatggtcatcaacccatc	TTTAGTTAGAGCCC
MYC_Loop5_6_E5	aaatgagaatcaagttcaggactatcttcctgagattattagtatctaca	TTTAGTTAGAGCCC
MYC_Loop5_7_E5	tttccaaatatcctttgatgtattagcccaatttagatgcaataacaaag	TTTAGTTAGAGCCC
MYC_Loop5_8_E5	cccagtggcttaaaaaaacaatttctttctagctcatgttgcatgtggat	TTTAGTTAGAGCCC
MYC_Loop5_9_E5	agcagctgtagaatctgcaccacctgtcttatgctgagatccagcctgga	TTTAGTTAGAGCCC
MYC_Loop5_10_E5	ccatccgtgagtcaggagttgtatcctcctccctcctatgggagacactg	TTTAGTTAGAGCCC
MYC_Loop5_11_E5	ggtcacgtgtagtcctcttaaaaggaagggacaggccgtccgggaggtga	TTTAGTTAGAGCCC
MYC_Loop5_12_E5	tgacctggggaaaggaagaacagccttgtatggggttatgacaatggact	TTTAGTTAGAGCCC
MYC_Loop5_13_E5	cccgtctgaatttgaatccaggcctgctagttattagctgagtggccttg	TTTAGTTAGAGCCC
MYC_Loop5_14_E5	aaaacattataatcctcagtcccttcgtctataaaatgggaacatcagta	TTTAGTTAGAGCCC
MYC_Loop5_15_E5	aggccgcagtgaagatttaatgagataaagcttaagttcttagcacagag	TTTAGTTAGAGCCC
MYC_Loop5_16_E5	aaaatattcagtgaatgacaaatataataataccagagtgaggccgggtg	TTTAGTTAGAGCCC
MYC_Loop5_17_E5	tcttcaaactagatctcatgagcacctatgattctcaccactcctggtgc	TTTAGTTAGAGCCC
MYC_Loop5_18_E5	tccctcggtaagacagcctctctcaacctggaaaatgccactggtccac	TTTAGTTAGAGCCC
MYC_Loop5_19_E5	ggaagaaaaacaaagaaaatggtaagagctgacagtttcgacaagctgtt	TTTAGTTAGAGCCC
MYC_Loop5_20_E5	ttctcaatcatggacaagtcaactctggtgacacatgaaaatgatgtgcc	TTTAGTTAGAGCCC
MYC_Loop5_21_E5	aatataccttagactcctgatatctaaatctggctttggaaaaagaatac	TTTAGTTAGAGCCC
MYC_Loop5_22_E5	ggaggacaggcaattctcaccttctctactgtaggaggaaaagcagccct	TTTAGTTAGAGCCC
MYC_Loop5_23_E5	ttcctctttgataaggggactagtttcagcaacatgacccctgccctc	TTTAGTTAGAGCCC
MYC_Loop5_24_E5	ttagcaacgatagaggtaacctcaccatcttcagtaattacccattggat	TTTAGTTAGAGCCC
MYC_Loop5_25_E5	cagttccttggtacctaggaagctggataagaagggtttcagggagccct	TTTAGTTAGAGCCC
MYC_Loop5_26_E5	aatgtctgccataagaaggagctgaattctaaggcagccttactccagcc	TTTAGTTAGAGCCC

MYC_Loop5_27_E5	acagagactgatggttgtgaatacagattcttgtcagccaacctggctca	TTTAGTTAGAGCCC
MYC_Loop5_28_E5	ttgagttttgatttccccatcagtaaaaatggggatgttaatataagtta	TTTAGTTAGAGCCC
MYC_Loop5_29_E5	ttcgtccttcagccattgtcacaaatctcatgatcaccctctcatttcag	TTTAGTTAGAGCCC
MYC_Loop5_30_E5	atacatcattcactacattgagtcttcacctggggtagtttcttttaaatg	TTTAGTTAGAGCCC
MYC_Loop5_31_E5	tatcagagaattacttcttgatgtttgcatgtgagccttatattacataa	TTTAGTTAGAGCCC
MYC_Loop5_32_E5	catttacccaaattgattgtcacatctgtggattgttatattattgcatt	TTTAGTTAGAGCCC
MYC_Loop5_33_E5	ctggactgttatattattgtttagaaacatttgttaccatcttaggtttc	TTTAGTTAGAGCCC
MYC_Loop5_34_E5	gtttattatatgtcagtagaaagttttctaaaatacaagttgccatgtcc	TTTAGTTAGAGCCC
MYC_Loop5_35_E5	ttctgccatttctcgctttcttctgacacatgaatacccgtgcccagatg	TTTAGTTAGAGCCC
MYC_Loop5_36_E5	ttccaaagagacgagaccatggagcagagtcatagctggcctccaaccaa	TTTAGTTAGAGCCC
MYC_Loop5_37_E5	aataagaaataaatgtggttcttataaaccgctgagatgtggatgtggtt	TTTAGTTAGAGCCC
MYC_Loop5_38_E5	catcacttagcctaaactatacaacaactgtgaggtatggaaattgataa	TTTAGTTAGAGCCC
MYC_Loop5_39_E5	cacagatgagaaactgcagctgtgatttgcagacactgcaaagctggaga	TTTAGTTAGAGCCC
MYC_Loop5_40_E5	agcctcctgcgcatgcttttgtttccaaaccttgtactcatcctactacg	TTTAGTTAGAGCCC
MYC_Loop5_41_E5	ccaacaaaatttggttcctttgtagaacttctcatagttcagatttttat	TTTAGTTAGAGCCC
MYC_Loop5_42_E5	ttttatttttttgtcagactcacccatcagtctaagctccatgaaggcag	TTTAGTTAGAGCCC
MYC_Loop5_43_E5	tttcaacattttattctcagttctgagtacatattagacgctattccttg	TTTAGTTAGAGCCC
MYC_Loop5_44_E5	aatgatcaaatgaatgaatgtatctgcccattgaataacttgattctttc	TTTAGTTAGAGCCC
MYC_Loop5_45_E5	gggaagaaaaaaagtgaccacaaaacaagtggaaattttgcattttaatt	TTTAGTTAGAGCCC
MYC_Loop5_46_E5	tgtaaccctcatgcaaataataaacattgttcccaataggtattttttca	TTTAGTTAGAGCCC
MYC_Loop5_47_E5	tgggaaatgtttagattcaaagtggacattagacccagtggcccagctgg	TTTAGTTAGAGCCC
MYC_Loop5_48_E5	atgacacccttaccgtccaactcattaactttctgtaagtattactacca	TTTAGTTAGAGCCC

Supplementary Table 9: Primary probes and docking handles targeting MYC loop 6

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop6_1_E6	ccttttgtaaacattaccttcatgacttatttacaaggcactgggtagag	TTTTGATGATAGCC
MYC_Loop6_2_E6	tggagggaagagaggaaagataaggtcacagagagagcactcagaggcaa	TTTTGATGATAGCC
MYC_Loop6_3_E6	accctcagaatatttgagggaggggctatgtacttctcccacttttattc	TTTTGATGATAGCC
MYC_Loop6_4_E6	acttctccaacattcatactggctttctgtttctcccttccctctccttc	TTTTGATGATAGCC
MYC_Loop6_5_E6	gggtacaggtgggtacagatggaggaaaattgtacacaagggtgcaagaa	TTTTGATGATAGCC
MYC_Loop6_6_E6	tcacgctaggggcagaaactctattttctcccaaaagtaacaatacagga	TTTTGATGATAGCC
MYC_Loop6_7_E6	aggaaactgaggaagaattgccgagtggaataagagcaagagctggtcag	TTTTGATGATAGCC
MYC_Loop6_8_E6	tctgcacagtcgtgtggcatcctccaatagtgctcagaagccggagtgct	TTTTGATGATAGCC
MYC_Loop6_9_E6	ctgtagcagtgaccacatttggggtttggctgaatgcttctggcaggagg	TTTTGATGATAGCC
MYC_Loop6_10_E6	cagaagcaggagcatcaataagagaattattttggtgatcaaccaagtgg	TTTTGATGATAGCC
MYC_Loop6_11_E6	ggcagacaggaggaaactggtgtcaaagcaagggatattgaatgcgaagt	TTTTGATGATAGCC
MYC_Loop6_12_E6	aaaccaggcagtgctgacattcgccagagtatggctacagggtagggtgg	TTTTGATGATAGCC
MYC_Loop6_13_E6	aaaaagcaatgatgctactgtgtacagtcaatgttctcagggacactgag	TTTTGATGATAGCC
MYC_Loop6_14_E6	agacaaaaaaaaaaagtgaggtaagagaaagatactgagccaggacccaa	TTTTGATGATAGCC
MYC_Loop6_15_E6	ggaaaaaagcaacaaaagaccaaaagctgcaatgaggagctgtgagggag	TTTTGATGATAGCC
MYC_Loop6_16_E6	tttttgtcaaaaaatcaaaatcacaacccaattgtcccatccaccaatga	TTTTGATGATAGCC
MYC_Loop6_17_E6	ctgtatgattccacttaaaggatacacctagactagtcaaattcgtagag	TTTTGATGATAGCC
MYC_Loop6_18_E6	agatcaacaagttctgaaagtagatggtggtgattgttgtacattgcaaa	TTTTGATGATAGCC
MYC_Loop6_19_E6	ccagtgaaatatacacttaaaacagtttagatggtaaattgtgtgttaca	TTTTGATGATAGCC
MYC_Loop6_20_E6	tgccaatttatagggaaaaaaaggagacagagtaacatgttagatgacac	TTTTGATGATAGCC
MYC_Loop6_21_E6	caatcagtaaagtttatgatttgagaaacttaaccaacaacaaggaaact	TTTTGATGATAGCC
MYC_Loop6_22_E6	gaaaccctaattaatagggacttaaaagacatatccataatgaatcctaa	TTTTGATGATAGCC
MYC_Loop6_23_E6	gaagtggtatgaggtctaggacttgcttaattgaaaacattcagggcagg	TTTTGATGATAGCC
MYC_Loop6_24_E6	gtcataagtgaacaagcgtcttcacaaaacaaggtcgtccatgagctgaa	TTTTGATGATAGCC
MYC_Loop6_25_E6	ttcatgaaggtgatgggtttatgcagattcattatatcactattctctct	TTTTGATGATAGCC
MYC_Loop6_26_E6	agttggaagcataccataataaaaagtagggagaaaatggacttttggtg	TTTTGATGATAGCC

MYC_Loop6_27_E6	agtagaagatattttaatgggcatatatataaatcctatgtttgggctca	TTTTGATGATAGCC
MYC_Loop6_28_E6	ggataccaggtgtcaagaatttacacaaaaatgatctgggatttagttgg	TTTTGATGATAGCC
MYC_Loop6_29_E6	gcatgaactagtaaagtggggtatctactacaaaagttaatcccacctta	TTTTGATGATAGCC
MYC_Loop6_30_E6	caacaggatcctctgtgtttaggtttgctaaggaatgtctggaatagtgt	TTTTGATGATAGCC
MYC_Loop6_31_E6	aaccccagaatttatagaagacattgtcaaactataaactattttactgcc	TTTTGATGATAGCC
MYC_Loop6_32_E6	taaatacagggatttgtgccatatcaggaagtagcatgagatatcacata	TTTTGATGATAGCC
MYC_Loop6_33_E6	ttttcactccattatcccgaagagcagagaaggaggataaatatgtcatc	TTTTGATGATAGCC
MYC_Loop6_34_E6	cacagactggaaaagtctcagaaatttgtctagtctttaaaaattatatg	TTTTGATGATAGCC
MYC_Loop6_35_E6	ctttaagctggtaaagatgatgttctagaagtatgcagtcggtcaggggc	TTTTGATGATAGCC
MYC_Loop6_36_E6	gtgtaaaaggaaacttactccagggagagacacataaaggggtgaataga	TTTTGATGATAGCC
MYC_Loop6_37_E6	atgtacctaataggtttaatgtatggatgtcagattctctacagaacaat	TTTTGATGATAGCC
MYC_Loop6_38_E6	tttattgagcatttgctatgaatcctataaagtcacttctgttatgctgc	TTTTGATGATAGCC
MYC_Loop6_39_E6	acgtgaaattgttccagcacaattgatacaataggaaacaattttgagcac	TTTTGATGATAGCC
MYC_Loop6_40_E6	catgtttgcttatgcaatatttcatctgtgagaaaaactaggtgcaccca	TTTTGATGATAGCC
MYC_Loop6_41_E6	ttatgtaggagtacacaaaacacacctcaaacatctcccagtcacctcag	TTTTGATGATAGCC
MYC_Loop6_42_E6	tatattcagcataatacacaaaggtgtgtgttatgaaccacaccccatcca	TTTTGATGATAGCC
MYC_Loop6_43_E6	tgatttcaaataactcttctatcacaacttctcaataactcacaaggtac	TTTTGATGATAGCC
MYC_Loop6_44_E6	cactcacttatacaagtaaactttgaattattcctagtaagataaagtgc	TTTTGATGATAGCC
MYC_Loop6_45_E6	ggaaaactgtgctaccattttatttggggtatatatatatatatttgagg	TTTTGATGATAGCC
MYC_Loop6_46_E6	tatgtatataactcgtgacatttttgagtattttccccttaccccattct	TTTTGATGATAGCC
MYC_Loop6_47_E6	ccgtggcttttgtgtgacttttaggaacacatgtgaagtactatagcaga	TTTTGATGATAGCC
MYC_Loop6_48_E6	gaagcaatttgcatgctttcagttgtcccaataactcagggaaatggatg	TTTTGATGATAGCC

Supplementary Table 10: Primary probes and docking handles targeting MYC loop 7

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop7_1_E7	gatgtcagtgcctcaggagagaaggcacaaatgactggcagcaacttgca	TTATAAAGTGTCCA
MYC_Loop7_2_E7	ctgtgctgggatgacctaaaggctgttgatcaaaggcctctccccggatcc	TTATAAAGTGTCCA
MYC_Loop7_3_E7	gccctttctatatgcacctcagggttccgagaacaagtgttccagcaaca	TTATAAAGTGTCCA
MYC_Loop7_4_E7	tgcatggctttttatgacctcacttcagaagtcccagtttcatttttggt	TTATAAAGTGTCCA
MYC_Loop7_5_E7	gtcaaagcagtcagaattccacttaaattcaaaggaaggggaccccagagc	TTATAAAGTGTCCA
MYC_Loop7_6_E7	ataaacacaattattaataacaaggtggttaagtaacttgcccacgccat	TTATAAAGTGTCCA
MYC_Loop7_7_E7	gcaaaggtgaagtctggtttcttaaccactgtacccttctcccttcctgac	TTATAAAGTGTCCA
MYC_Loop7_8_E7	tacactgctttcttttctctatagcactgtttacttcttggtcatcaatc	TTATAAAGTGTCCA
MYC_Loop7_9_E7	aatggaagttttgtgaagaccaggacttcagttcattcactgttatgtcc	TTATAAAGTGTCCA
MYC_Loop7_10_E7	gacaatgtgtggcgtgtaatctatactctaaaagaaaattgtcaaataaa	TTATAAAGTGTCCA
MYC_Loop7_11_E7	tctgctcaactacaaaccaaggtgaaaattataatccctacacatacttt	TTATAAAGTGTCCA
MYC_Loop7_12_E7	ctacagattattggatttgtctcgaggaggcattaggtagcttttgctct	TTATAAAGTGTCCA
MYC_Loop7_13_E7	gaggccacatatttggtagttgtacatttttactagtgatatagtaaaaa	TTATAAAGTGTCCA
MYC_Loop7_14_E7	caaactccttgcagttattttacttgcagacctttcctgaaatgtctttc	TTATAAAGTGTCCA
MYC_Loop7_15_E7	tcttgtcatcctcaaattaagcatggctcatgaagtgggggctgaatggt	TTATAAAGTGTCCA
MYC_Loop7_16_E7	atatgatatcctaagccctggacctgtgcatgttactttatatggtaaag	TTATAAAGTGTCCA
MYC_Loop7_17_E7	taacaacacgttgatgttagctcattgaaactgactttgaaattctggcc	TTATAAAGTGTCCA
MYC_Loop7_18_E7	gaaagaataaatttctgtcttttcaaaccaccatgtttgtgcaaatttgt	TTATAAAGTGTCCA
MYC_Loop7_19_E7	gaaactagcatgcctcctctagaaagcctgttctaactacttttcctcc	TTATAAAGTGTCCA
MYC_Loop7_20_E7	actcgtccacttccaagaaccagcaaataacatcaatttgtgaaatgagt	TTATAAAGTGTCCA
MYC_Loop7_21_E7	caaaccaaagggcatatgtcgctgcaaagcaccaagaaagagataggaac	TTATAAAGTGTCCA
MYC_Loop7_22_E7	gatgggtgaaggatatggaaggtaatttactaaagaaggtaaccacccaa	TTATAAAGTGTCCA
MYC_Loop7_23_E7	catgtgaagagtgctcaaactcattagtcattagaaaaaataaacataac	TTATAAAGTGTCCA
MYC_Loop7_24_E7	aaagctggaaaacaccaagagttggtgggacttgcagatttggaacatc	TTATAAAGTGTCCA
MYC_Loop7_25_E7	gatgggggtgaaaatgctgaaatgaaaaatcagttatgaatgcaccacac	TTATAAAGTGTCCA
MYC_Loop7_26_E7	gtagggaagcatccatccacgtgaccatctttggaggagggaataagact	TTATAAAGTGTCCA

MYC_Loop7_27_E7	atgaacggagtagccaacagcagttagaaaccatatgtatacacaacaac	TTATAAAGTGTCCA
MYC_Loop7_28_E7	tcaagtaccaggctgagtccagaaaataagaaacagaatgagatatctga	TTATAAAGTGTCCA
MYC_Loop7_29_E7	tttgtaaattaaactctacacgtacaaagcactatggatcttaccagaac	TTATAAAGTGTCCA
MYC_Loop7_30_E7	gaaagataaatgtcaatacattagaatcttgcctataggaagaggaagat	TTATAAAGTGTCCA
MYC_Loop7_31_E7	gagaaataaatagaagcaaagataaatgagacgggatcatttatctaatg	TTATAAAGTGTCCA
MYC_Loop7_32_E7	aggataacactgtgtaattggctgaagagttttattaactcaaccatctg	TTATAAAGTGTCCA
MYC_Loop7_33_E7	ttaaaaaaaaaaaaacatgctcacctatgctgtttggagatttttggagag	TTATAAAGTGTCCA
MYC_Loop7_34_E7	tcaaattctttgacacttcttccaccaaagcaatagaagtaatgttgcat	TTATAAAGTGTCCA
MYC_Loop7_35_E7	ctgggccataaaaggaatttttcctcccagaatagttcctctcatactca	TTATAAAGTGTCCA
MYC_Loop7_36_E7	agccccaagctgttagagtacctcctaagagtcttcctggctcaggcccc	TTATAAAGTGTCCA
MYC_Loop7_37_E7	aaatctgtgagcataataagatagctgtttttttccactaagctttgggg	TTATAAAGTGTCCA
MYC_Loop7_38_E7	actgctgtaggcaccaaatcatgccttatctaaaggagcaattacttcct	TTATAAAGTGTCCA
MYC_Loop7_39_E7	actgtttccccatggaaataagcactcaggttagctaaatattcttttt	TTATAAAGTGTCCA
MYC_Loop7_40_E7	ggaaaaatccagattgtcttctgatttttcaatgtcagcaactattaaaa	TTATAAAGTGTCCA
MYC_Loop7_41_E7	atagtgaagaagaaacaaatctcatttgcacttcctaattggccttggag	TTATAAAGTGTCCA
MYC_Loop7_42_E7	gtggctggtgtagggcttgtttcaaacaactgtgtgtctaaagaagcaat	TTATAAAGTGTCCA
MYC_Loop7_43_E7	gaggtgatacaaaaatgtactggtcaaaatggtggtccccaaaaagatat	TTATAAAGTGTCCA
MYC_Loop7_44_E7	gacacagagaaaaagacaatgtgaagatcgaggcagagactgaagtgatg	TTATAAAGTGTCCA
MYC_Loop7_45_E7	ttttgcacttctggtcttcagagttgtgaaagaaataaatttctgttgtt	TTATAAAGTGTCCA
MYC_Loop7_46_E7	tagtacccataggaaattaataaaagattaagtcggtgcaaaagtaattg	TTATAAAGTGTCCA
MYC_Loop7_47_E7	ctgaaaccactacatggtccttaatccataataactgtgaagtcaccacc	TTATAAAGTGTCCA
MYC_Loop7_48_E7	gttacttaaatcttgcattattgtttaatattctcatttccctcctggtg	TTATAAAGTGTCCA

Supplementary Table 11: Primary probes and docking handles targeting MYC loop 8

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop8_1_D8	caatccaaacttgattccaaggtcaatcttctcctgttccctgggtaagg	TTATATGATCTCCG
MYC_Loop8_2_D8	cagacttctccttcattgagttactaagtggtttctgccactatggagtt	TTATATGATCTCCG
MYC_Loop8_3_D8	gaaattaaggcaaggggcacatgaaaatatttgcttgcctggtattcttg	TTATATGATCTCCG
MYC_Loop8_4_D8	tatattctttgagcacagcagatattcaaaaatgcctctttaagagacat	TTATATGATCTCCG
MYC_Loop8_5_D8	ttctgtttcggctgatggtttattccttccacatctcccaggaatgtggc	TTATATGATCTCCG
MYC_Loop8_6_D8	ctgggctccactctaggcaactgtacggaacggacagactacaatcccac	TTATATGATCTCCG
MYC_Loop8_7_D8	tgcatatagaacatggaaagagtgttctatttttatccctgaaaagctct	TTATATGATCTCCG
MYC_Loop8_8_D8	ctgcacagtgtagccatctctcctttggatcttgcaacagcagctgacca	TTATATGATCTCCG
MYC_Loop8_9_D8	gtccctcaaatgtctcgggtgggagtcaggcttaagtccactgcaacctg	TTATATGATCTCCG
MYC_Loop8_10_D8	tcttgctgaaacccttctaggtagactccaatgagggacatgttccctgc	TTATATGATCTCCG
MYC_Loop8_11_D8	gagctcagtggagaagcaacctctaccaggaagtcctctttcttctaaat	TTATATGATCTCCG
MYC_Loop8_12_D8	ctgacactttcctatcctggtgtaggtgaggcacttaagagggagctaac	TTATATGATCTCCG
MYC_Loop8_13_D8	actgacttttcacaagttcttctgggtggtcaatctcacaacacattttg	TTATATGATCTCCG
MYC_Loop8_14_D8	tctgttgtcataggatctcagccaaaacatccatgttaacttcctttcat	TTATATGATCTCCG
MYC_Loop8_15_D8	ctactctgcaactttcccaagagtgaaatccctgtttttttttgtttttt	TTATATGATCTCCG
MYC_Loop8_16_D8	cagtattcaagagtcattgcttgaatgaatgaataaactaatagatactg	TTATATGATCTCCG
MYC_Loop8_17_D8	aaagcagggtcctcaaaattaaagatatcagcttgactgaagaagagagt	TTATATGATCTCCG
MYC_Loop8_18_D8	aggatcaagatctaaataggaaggaagggcttagataattgaggcagaaa	TTATATGATCTCCG
MYC_Loop8_19_D8	gaccatgaagcatagtggaaaaaaagataggtttcagggacagatggatc	TTATATGATCTCCG
MYC_Loop8_20_D8	tcagctctgccattaaagttgactctggaccaggagcaaggttcttaacc	TTATATGATCTCCG
MYC_Loop8_21_D8	cagcattctcactgtaaaataaatatcatcaatgctatctcataagattg	TTATATGATCTCCG
MYC_Loop8_22_D8	gacagtgtatgtaaaacacctggaatagggcttatcacacagtgggtggc	TTATATGATCTCCG
MYC_Loop8_23_D8	aataatgactttaaatcatatttgagtacatcttgtgggtcaggactgt	TTATATGATCTCCG
MYC_Loop8_24_D8	tttagcacaaagaattaaagttaaatgccatgctcaaagtcacagagctg	TTATATGATCTCCG
MYC_Loop8_25_D8	atctctcttcctacagagcctgctgtatgttaccctgacatctcatgcat	TTATATGATCTCCG
MYC_Loop8_26_D8	gatcattgtgtagctggcattaagagactctaggaaagaaatggtagcag	TTATATGATCTCCG

MYC_Loop8_27_D8	caacatgcttcagtctcacgctaaaagcaattataattttcttttcttaa	TTATATGATCTCCG
MYC_Loop8_28_D8	aaatttctcaggaaataaaacatgttagtatgattggagtacacactgtg	TTATATGATCTCCG
MYC_Loop8_29_D8	acagtgagggtcagctgacagaatgtgtgaaaacgcttggaaacctataa	TTATATGATCTCCG
MYC_Loop8_30_D8	aatccatctctccccaccaaggtcttcctcccactaaaagaacaagaatg	TTATATGATCTCCG
MYC_Loop8_31_D8	atgaagagtcaagagaggtgagaatactgtgggaatgcacctcaccaact	TTATATGATCTCCG
MYC_Loop8_32_D8	gttttatctgacactgggaccaaagcaaatgtgacctgaatattcaaccc	TTATATGATCTCCG
MYC_Loop8_33_D8	caggacacattattaacaagtagattgtgtacacctaggaatataaatgg	TTATATGATCTCCG
MYC_Loop8_34_D8	tacctccactggagcctcggagctttaaatagacccatatcctacttcag	TTATATGATCTCCG
MYC_Loop8_35_D8	ttaacccctaacacaaccctgtagagtagatatttgaatcaccagcataa	TTATATGATCTCCG
MYC_Loop8_36_D8	caggtctatgcaatttctaaatcaatgttcatgataactatggtatctta	TTATATGATCTCCG
MYC_Loop8_37_D8	agctaaactggaggccaagatgaagatataaacagagaatggaaaggtgt	TTATATGATCTCCG
MYC_Loop8_38_D8	aaaaaaggaaacagatactgggctctgaaagtggagtgccccacctctat	TTATATGATCTCCG
MYC_Loop8_39_D8	tctattccttcagtaagttatgcatgtcctcagttttttcatctataaaa	TTATATGATCTCCG
MYC_Loop8_40_D8	tacgattgtcgtaaggatgtgatgagataataacccgcaaagcccttaga	TTATATGATCTCCG
MYC_Loop8_41_D8	catgtaagaagtgccattttagctacaattactctgaagtgtgggtgggc	TTATATGATCTCCG
MYC_Loop8_42_D8	tttgaaactatcttcctgactgccaagaccagatggtttccctaggagct	TTATATGATCTCCG
MYC_Loop8_43_D8	ctacagtgaactggaagaaactaacattggtttacccacagcaaggcttg	TTATATGATCTCCG
MYC_Loop8_44_D8	tttcttaatagggattcaaactgggaatgttctaggcataacagatggtg	TTATATGATCTCCG
MYC_Loop8_45_D8	gccattaaggtctcctgtgatttttttccacatttattggttttatttta	TTATATGATCTCCG
MYC_Loop8_46_D8	tgggtttttatagaggtttcactacttaagcatgattgattaaaccactg	TTATATGATCTCCG
MYC_Loop8_47_D8	cacagtccactccctggttttcgacatggatcccttgcatcaaaagaata	TTATATGATCTCCG
MYC_Loop8_48_D8	taataattagtccactccatcatattgcatgaatatcttcccagggtgag	TTATATGATCTCCG

Supplementary Table 12: Primary probes and docking handles targeting MYC loop 9

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop9_1_E9	gtcttaataagaaaaaccaacaggagcctgaatgaccgtcagcccacagg	TTTATTAAGCTCGC
MYC_Loop9_2_E9	gaaacactgaaggtaacacttgacctcaagaagtttcccctctggctggg	TTTATTAAGCTCGC
MYC_Loop9_3_E9	tcttttcaagtcttatcttagtggctttctttacaactttgtccctagta	TTTATTAAGCTCGC
MYC_Loop9_4_E9	cttcctgaaactcacctcgcctttaatttccagggccctactctcaacac	TTTATTAAGCTCGC
MYC_Loop9_5_E9	tttttttttcattctctcttcttggttcttcttccatctctcttttagttg	TTTATTAAGCTCGC
MYC_Loop9_6_E9	cagagttccttggtactcttgtcttactgcacatagctccctgagccaca	TTTATTAAGCTCGC
MYC_Loop9_7_E9	gtgctaacctgctacactaccttccaggtaacagctcctgggtacccccca	TTTATTAAGCTCGC
MYC_Loop9_8_E9	cttcaatgcagaacccattgttttcccttcaaacctattcctccttcagg	TTTATTAAGCTCGC
MYC_Loop9_9_E9	ttaatgatattactatccaattagctgcccaagttctgtttgctaagtga	TTTATTAAGCTCGC
MYC_Loop9_10_E9	tgattaaaagcacaagtgctagaactgagctgtctgttttcagattccag	TTTATTAAGCTCGC
MYC_Loop9_11_E9	cttcaatagtgcctactgtattgagttgttgtaaagatagaataagttaa	TTTATTAAGCTCGC
MYC_Loop9_12_E9	tcgatttctcctactccctcaaacctctgcaaatggtttgccctcacctt	TTTATTAAGCTCGC
MYC_Loop9_13_E9	tcctatcaggaatctcacttctccatcccaattgctgctgcctgtgttca	TTTATTAAGCTCGC
MYC_Loop9_14_E9	ttcaaagccacatctacccagttcatgtccccagagttgaggccctggtg	TTTATTAAGCTCGC
MYC_Loop9_15_E9	tatctaatgccaagacttagctctcagattctgtggttccagaaccctgg	TTTATTAAGCTCGC
MYC_Loop9_16_E9	gggcagcatgagttcatgtagagtcagcgaagtttccccctccacactca	TTTATTAAGCTCGC
MYC_Loop9_17_E9	caaaattccactatctcttcagcattcctcctactttgcattcctttctt	TTTATTAAGCTCGC
MYC_Loop9_18_E9	tatttcctggtccctttagtccacagacaggtaggaagccacaagcagc	TTTATTAAGCTCGC
MYC_Loop9_19_E9	attgataagggtgggaacattcaagctaccgataagacctctgcacccag	TTTATTAAGCTCGC
MYC_Loop9_20_E9	gaaaacagagaatcactttatggtttcctgtaagcagccttctccaaagc	TTTATTAAGCTCGC
MYC_Loop9_21_E9	attattacaacttgggggatgagtgctgaaggcatccagtggatagtggc	TTTATTAAGCTCGC
MYC_Loop9_22_E9	caatgcatagaacagtctagcaagcaaagaattatctggctcaacacatc	TTTATTAAGCTCGC
MYC_Loop9_23_E9	ggttcagaaacgctggtttagaaaatggagagaagaggaatggaaagaga	TTTATTAAGCTCGC
MYC_Loop9_24_E9	ccatgaatataaagagtgaggcaccatggcatttttgctctgctgctctt	TTTATTAAGCTCGC
MYC_Loop9_25_E9	gtatccacacttcaaagagatgtcaaaaatccgtcatggttctaaacagt	TTTATTAAGCTCGC
MYC_Loop9_26_E9	gtaagatctatttgtctccttgttttacatgagctgacaaactcacttga	TTTATTAAGCTCGC

MYC_Loop9_27_E9	aattcatggttcatggtgcccagagaccttataggtcctggctgggtagc	TTTATTAAGCTCGC
MYC_Loop9_28_E9	tgggaacagtgaattctctgataccttagaaaaggaagggctgacatcag	TTTATTAAGCTCGC
MYC_Loop9_29_E9	ctggattgtgcactggccttggaaagaagagatatctgagcccaggacag	TTTATTAAGCTCGC
MYC_Loop9_30_E9	tcaataaatcgaggtgggaaagagggttttaagagaaccacacaagcaaa	TTTATTAAGCTCGC
MYC_Loop9_31_E9	gagataaaatgtgtgtctgacccattgtagttgctcataaatgccggctc	TTTATTAAGCTCGC
MYC_Loop9_32_E9	agaaaactgaggtgagcagatgaagcttactctactgggacccaacccta	TTTATTAAGCTCGC
MYC_Loop9_33_E9	tgacagcgtccttagttacggaaaccttagggccttcagtgatacttcag	TTTATTAAGCTCGC
MYC_Loop9_34_E9	tgggagaaagcaccagaaccatctacttgagtgattcaacacattttctg	TTTATTAAGCTCGC
MYC_Loop9_35_E9	cctttccaggataatttcctcaaaattatactgtttttaataggatcag	TTTATTAAGCTCGC
MYC_Loop9_36_E9	gtgtaagtgaatgtgttttagtcttctcaaatggtgtttatgtactttttt	TTTATTAAGCTCGC
MYC_Loop9_37_E9	gattgtttcacttggtaatatgcactgaaggttcttccatgactttttat	TTTATTAAGCTCGC
MYC_Loop9_38_E9	catttatgtttagtgctgaataatattgcattttctggatggaccacagt	TTTATTAAGCTCGC
MYC_Loop9_39_E9	gaagagaaacagatggattagtttcctggagagattctgcaggctccttg	TTTATTAAGCTCGC
MYC_Loop9_40_E9	aacacatcatgtagcagatcatgcaaaccaagcttcaagagggaagatga	TTTATTAAGCTCGC
MYC_Loop9_41_E9	cccagagttctttactgacagattcaagatgatagccctgggtccccacc	TTTATTAAGCTCGC
MYC_Loop9_42_E9	atctaaatctcaatgtccttctctagattaggggcacgtcacatccaccc	TTTATTAAGCTCGC
MYC_Loop9_43_E9	tcagtacgactgtttctatgtcatactattcagtggagttgaagaagagt	TTTATTAAGCTCGC
MYC_Loop9_44_E9	tccaaacagttaagtcccagaaggactttgctagacctgaggcagtgaag	TTTATTAAGCTCGC
MYC_Loop9_45_E9	ttcaatcctttgcttaacaaatattttggaatgtctacttatgccagaca	TTTATTAAGCTCGC
MYC_Loop9_46_E9	gaaatatcattcaacgtctctgagtgttagtttcttcatttgcaaaatga	TTTATTAAGCTCGC
MYC_Loop9_47_E9	ctaccctaaaacagaactgagacaacatctgcacatctctcagtcggctg	TTTATTAAGCTCGC
MYC_Loop9_48_E9	gtgagcatgcagtacaagttacctactgatgtcgccaggctaaggcagag	TTTATTAAGCTCGC

Supplementary Table 13: Primary probes and docking handles targeting MYC loop 10

Probe Name	Genomic Sequence (5'-3')	Docking Handle Sequence (5'-3')
MYC_Loop10_1_E10	gtactttgaaggtgagtcctcaacccaaatttccaacagggttgtatgta	TTTTAAAACAGCCT
MYC_Loop10_2_E10	cccacaatgctgctaaattgcctaaacattactggctaacattataaagt	TTTTAAAACAGCCT
MYC_Loop10_3_E10	agcatatttgtagccatatgtgagagtttagctaaatcctgttttatgtt	TTTTAAAACAGCCT
MYC_Loop10_4_E10	aaatgtattaaagtattaaagaggtaaaaagtagcgtggaggccgggcgc	TTTTAAAACAGCCT
MYC_Loop10_5_E10	cagcaagtatttactgaaaacgtactcaatatacaaccctgtgctgggca	TTTTAAAACAGCCT
MYC_Loop10_6_E10	acactctccttttcctcaatgtgttttgacaccagtgagaacagcaaggg	TTTTAAAACAGCCT
MYC_Loop10_7_E10	agagcacaaactttgggcttgacaaacttaggtttgcattcttgctcctc	TTTTAAAACAGCCT
MYC_Loop10_8_E10	cttgggaaaagttattccatttcttttaagttcagtttcttcatctgtaa	TTTTAAAACAGCCT
MYC_Loop10_9_E10	cctcatagactggtttcaaggatttcattcaatcaaaacttaaaacataa	TTTTAAAACAGCCT
MYC_Loop10_10_E10	tgcaaggactcaattcatgtggttgaaaaatattcagccactgtgaaaaa	TTTTAAAACAGCCT
MYC_Loop10_11_E10	gaagttatcatgtgatccagtaagttcatccctaagtatatacccaaggt	TTTTAAAACAGCCT
MYC_Loop10_12_E10	tttgtgcacacaaaaccttgtatgatagtgttcatagcagcattattcac	TTTTAAAACAGCCT
MYC_Loop10_13_E10	acacatggtagaatattcagcaagaaaaaggaatgaagtcctgatatatg	TTTTAAAACAGCCT
MYC_Loop10_14_E10	caggaacatgatgttaagctgaagaaaccaatcacagaggattacatatt	TTTTAAAACAGCCT
MYC_Loop10_15_E10	tttttataaaatgtccagatggcaaatccagactggggagaacaccacag	TTTTAAAACAGCCT
MYC_Loop10_16_E10	cactgcttaatggggacaggatctcctttggggtaatgaaaatgttctag	TTTTAAAACAGCCT
MYC_Loop10_17_E10	gggtgtactaaatgccatgaattattcgttttaaaatggttgattttatg	TTTTAAAACAGCCT
MYC_Loop10_18_E10	ttttggcactgaatgaatgaaagcacatcatgcttttcctggtcctaggc	TTTTAAAACAGCCT
MYC_Loop10_19_E10	tttatgttggcttgaaattttgtgcattgattctgtgtcctgaaacttca	TTTTAAAACAGCCT
MYC_Loop10_20_E10	tcctctgtagtcctcaggttctagcagaaattcagtagacagttcttgcc	TTTTAAAACAGCCT
MYC_Loop10_21_E10	tatagatggtaggagtgttcctaagctctgaagataagttaacaccttgt	TTTTAAAACAGCCT
MYC_Loop10_22_E10	gcaacctctgtgcaaaacattgtaaagctctgagtctcagtttcttcctc	TTTTAAAACAGCCT
MYC_Loop10_23_E10	ttctctctctatctcaaatctccttctacccgtctctttcaccccaggctg	TTTTAAAACAGCCT
MYC_Loop10_24_E10	gcctagccagataatctaggataatcttatctaaagattttaccttaat	TTTTAAAACAGCCT
MYC_Loop10_25_E10	ttactctttcaggaattagtacataaacatatctcttgatgggggcgggg	TTTTAAAACAGCCT
MYC_Loop10_26_E10	aattcaacccactgttgaagagataatcctggtaagtgctttctctggcac	TTTTAAAACAGCCT

MYC_Loop10_27_E10	cagatatgttcttatgatcatcgtcatcattgtcttctttattattctgt	TTTTAAAACAGCCT
MYC_Loop10_28_E10	tgtatgtctctaggacagagtctaactgaaatcttcatctcactggggct	TTTTAAAACAGCCT
MYC_Loop10_29_E10	acagccttccatccttttttaggatgctggagccttttctcgtctctccc	TTTTAAAACAGCCT
MYC_Loop10_30_E10	cccggtccaaattcaattctcactgggattatgcagctcttctccctcaa	TTTTAAAACAGCCT
MYC_Loop10_31_E10	atttctgcgcccacccactattctgagattctgggatcttaatcctgttt	TTTTAAAACAGCCT
MYC_Loop10_32_E10	ctatatgcactggactcttcaggtgcccctacagagctgaccctccgcag	TTTTAAAACAGCCT
MYC_Loop10_33_E10	tgagatcccacctctacaaaagaagtaaaaactagccatgtgtggtgctg	TTTTAAAACAGCCT
MYC_Loop10_34_E10	cccaagaatggcattgctccatttctaagcactggaaatttcaagatgt	TTTTAAAACAGCCT
MYC_Loop10_35_E10	ctattgcttactaaacagttctggagcagaaagagttttatgggggaaag	TTTTAAAACAGCCT
MYC_Loop10_36_E10	gtgataggcttggaatcaaatcttgccacttattaagctgtgtggccctg	TTTTAAAACAGCCT
MYC_Loop10_37_E10	ggagttaataatacttatctcacagaactggtgtgcagaagtcatgaata	TTTTAAAACAGCCT
MYC_Loop10_38_E10	cagtctctgcaccataataaggaagactcaaattccttctcctttccctc	TTTTAAAACAGCCT
MYC_Loop10_39_E10	agtttaaatcaaattgcagacaccaaaagtgtaccttctaccaaaggttt	TTTTAAAACAGCCT
MYC_Loop10_40_E10	gaaagggttaattgtttagtttagagtcttagatatgtgttcgctagatt	TTTTAAAACAGCCT
MYC_Loop10_41_E10	cagaaatccactcccagatgattaaattagggtgaggaaaaacggtgagg	TTTTAAAACAGCCT
MYC_Loop10_42_E10	gtggcttcgggctacttgatggtaatttgactaggttaacacaaactgca	TTTTAAAACAGCCT
MYC_Loop10_43_E10	tttggcaggctacaaaagctggtagacgtgaaagttctggtgagcttttt	TTTTAAAACAGCCT
MYC_Loop10_44_E10	atgctcccagcattgtctctgcattttgccttttgcaggaacaccctgtg	TTTTAAAACAGCCT
MYC_Loop10_45_E10	gcctttcccttacattttccatctttttaccttccattagcagtttgca	TTTTAAAACAGCCT
MYC_Loop10_46_E10	tgcaattctcctccccgatgtattattccattctcacgttgttataaaga	TTTTAAAACAGCCT
MYC_Loop10_47_E10	gaagcaaacacattcttctttacgtgatgtcaggagagagaaatgccaag	TTTTAAAACAGCCT
MYC_Loop10_48_E10	ctaataaagacagagtcatctataatgtctctcacacatcttttcctcat	TTTTAAAACAGCCT

Contributions

Dr. Carl Barton wrote a light version of ChromoTrace, analysed exchange data, and contributed to the downstream data analysis. He also contributed by converting my ImageJ and R based spot detection scheme into more efficient programming languages such as C#.

Prof. Dr. Ralf Jungmann and his colleagues Orsolya Kimbu Wade, Dr. Maximilian Strauß, and Thomas Schlichthärle helped with the *in vitro* DNA-Exchange-PAINT experiment that I performed in their lab. Florian Schüder designed the bifunctional docking handles and imaged FISH probes with a spinning disk confocal microscope.

Dr. Christian Tischer developed a plugin for FIJI that automatically extracts metadata from images and saves them in both human and computer readable *yaml* formats, while Dr. Jean-Karim Hériché defined the core structure of this format. These contributions deserve particular acknowledgement and saved me countless hours of work.

Technical support was provided by Bianca Nijmeijer, Nathalie Daigle and Dr. Andrea Callegari, who helped me with culturing cells, running FISH protocols, performing immunofluorescence staining and much more.

Dr. Nike Walther was an equal contributor to the invention and establishment of the EdU-PAINT protocol.

Dr. Jean-Karim Hériché contributed with suggestions for the analysis of image data and the statistical analysis of chromatin traces as well as general programming guidance.

Dr. M. Julius Hossain contributed with 3D nuclear segmentation. This was especially important for separating agglomerated nuclei. He also gave advice on how to segment spots detected in confocal images of FISH probes.

Dr. Antonio Z. Politi contributed to the planning and analysis of FCS-calibrated imaging data.

Prof. Magda Bienko provided excellent training and shared the first set of FISH probes.

Anniken Waage Fougner contributed by drawing line profiles through spots in the DNA-PAINT experiments.

Acknowledgements

Firstly, I would like to thank Dr. Jan Ellenberg for giving me the opportunity to work in his lab and allowing me to pursue this challenging project. I am very grateful for all his advice and both critical and constructive advice and questions. I have enjoyed working in an environment where I had the freedom to define the direction of my project while knowing that there is support and guidance available when needed. It has also been tremendously helpful that Jan introduced me to many external and internal collaborators and eagerly sent me to visit collaborators in order to learn new techniques and perform experiments. I would also like to thank Dr. Stephanie Alexander for her encouraging comments and her continuous support.

Secondly, I would like to thank all former and current members of the Ellenberg lab for their willingness to share their knowledge and for all the good moments we had both at and outside work. I would particularly like to thank Dr. Nike Walther for her friendship, scientific support, good moments in short coffee breaks, and for all the shared ice cream.

I would also like to thank my thesis advisory committee - Dr. Jan Korbel, Prof. Dr. Dirk-Peter Herten, Dr. Jonas Ries, and Prof. Dr. Rein Aasland - for advice and comments.

Thank you to Dr. Marko Lampe, Dr. Jonas Ries and Ulf Matti for showing me how to operate the various super-resolution microscopes as well as thank you to Nathalie Daigle who showed me how to operate confocal microscopes.

I would also like to thank Dr. Stephanie Alexander, Dr. Franziska Kundel, Dr. Natalia Rosalia Morero, Marjolein van der Boon, Adele Rickerby, Nathalie Daigle and Dr. Nike Walther for proofreading this thesis.

I would furthermore like to thank Prof. Dr. Rein Aasland who introduced me to science in my Bachelor studies in Bergen and has remained a mentor and inspiration throughout my scientific career.

Finally, I would like to thank Anniken for her patience, support and countless sacrifices to help me reaching this point.