A genetic memory initiates the epigenetic loop necessary to preserve centromere position

Sebastian Hoffmann

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à Institut Curie

A genetic memory initiates the epigenetic loop necessary to preserve centromere position

Une mémoire génétique initie la boucle épigénétique nécessaire pour préserver la position du centromère

Soutenue par

Sebastian HOFFMANN

Le 01 octobre 2020

Ecole doctorale n° 474

Frontières de l'Innovation en Recherche et Éducation (FIRE)

Spécialité

Biologie cellulaire et biologie du développement

Composition du jury :

Sylvia, ERHARDT Prof. Dr., ZMBH University of Heidelberg/ Karlsruher Institute of Technology	*Présidente*
Lars, JANSEN PhD associate Prof., University of Oxford	*Rapporteur*
Ines Anna, DRINNENBERG CR, PhD, HDR, Institut Curie	*Examinatrice*
Daniele, FACHINETTI CR, PhD, HDR, Institut Curie, UMR144	*Directeur de thèse*

institutCurie | PSL

TABLE OF CONTENTS

1. ACKNOWLEDGEMENTS

Pursuing a PhD is often a challenging endeavor (see for example E. Pain, Science, doi:10.1126/science.caredit.a1700028). Fortunately, I've received support in many ways from numerous people to whom I would like to express my utmost gratitude!

I'd like to thank Daniele (Dani) Fachinetti! I met Post-doc Dani during my master studies in San Diego in 2015, where he worked in the lab next door. In SD, Dani stood out to me for two reasons: He has always been very helpful and he showed a good sense of humor! Luckily, Dani gave me the opportunity to start my PhD in his brand-new lab at Institut Curie. During my PhD, I then had the joy to experience Dani's remarkable scientific knowledge, benefit from his lab skills and scientific expertise as well as his great personality! Dani is absolutely passionate about our work and never afraid to go on new paths. I am very grateful for all the things I learned from Dani, for his sedateness, for all of our open and engaging discussions and for his continuous believe in me and the project. Thanks Dani!

I'd like to thank all past and present members of the lab! Riccardo Gamba (Ric) greatly contributed to the success of this project. He is an incredible, much valued scientist and a true gem for the lab in so many ways. Thank you, Ric! The success of the young Fachinetti lab is linked to one person in particular, which is our lab manager, Marie Dumont! Marie is the reason why the lab runs smoothly. She truly is the heart of the lab. Thank you, Marie! Special past members of the Fachinetti lab are our biochemists: Florian Chardon and Stéphane Gangnard. These guys really spice things up (in the case of Stéphane of course not literally). Thank you so much for your help and for giving the lab the right vibes. I'd further like to thank Veer Keizer, especially for performing single molecule microscopy experiments with me. The Fachinetti is defined by collaborative spirit. This also holds true for Solène Hervé. Thank you for all your help, especially on my quantifications! I'd like to thank Andrea Scelfo. He provided important input for my project and is a very insightful and helpful desk neighbor! Thanks a lot, Andrea! I'd like to thank Catalina Salinas-Luypaert and Therese Wilhelm as well as Viviana Barra and Pierre Simonin. Thank you for your suggestions and help, the great atmosphere in the lab, for fun lunches and especially coffee talks. Laurent Chen helped me as a master student to progress my project with great effort and Gabrielle (Gabi) Bersano-Marchisio contributed greatly to cell line generation, which I was able to benefit from during the revision of the project. Many thanks!

Next, I'd like to thank all other collaborators of this project: Special thanks go to Helena Izquierdo and Nicolas Manel who crucially contributed to this project, in particular with their expertise on T cells: Thanks for this fruitful and fun collaboration! Further, I thank our external collaborators Shannon McNulty and Beth Sullivan for their important contributions on this project!

I'd like to express my gratitude to Cristina Bartocci for proofreading not only the EMBO journal manuscript but also other important manuscripts during my PhD. Special thanks also go to Aruni Senaratne for proofreading the introduction of this thesis. Thank you very much for going through all this!

I would like to take the opportunity to thank the members of my thesis jury for accepting to read this work and/or judge my defense.

During my PhD, I received scientific advice and/or reagents from the following people that I'd like to thank; Past and present members of my thesis advisory committee: P. Tran (IC, FR), S. Marcand (CEA, FR), G. Orsi (LBMC, FR), U. Böckelmann (ESPCI); The Basto and Drinnenberg team members for joined lab meetings; Other outstanding scientists: A. Coulon (IC, FR), B. Black and G. Logsdon (UPENN, US), I. Cheeseman (MIT, US), D. Cleveland and A. Desai (UCSD, US) P. Maddox (UNC, US), D. Foltz (Northwestern University, US), L. Jansen (University of Oxford, UK), A. Straight (Stanford, US) and G. Almouzni (IC, FR). I thank the Flow cytometry platform, the Cell and Tissue Imaging facility, the antibody facility platform and the sequencing platform at Institut Curie. I thank my funding bodies: the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 666003 and the French foundation for medical research (FRM). I'd further like to acknowledge the Perez and Houdousse labs for the great atmosphere on our floor and of course for sharing reagents, equipment and knowledge with us!

I've been very lucky to be part of the international PhD program of the Institut Curie (IC3i) and I thank the organizers of this great program under the direction of Graça Raposo. A great gift of this program is the international community of PhD students which comes with it! I thank all the lovely PhD students of the program and also many other people from Curie, whom I befriended over the years! This community was of great importance for my personal life during the past 4 years! Merci to all of you! In particular, I'd like to mention Deepanjan Ghosh. He is a hoot! Some corny words for him, because I know he likes them: Deep has been the greatest joy of my life in Paris. I am eternally grateful for having met him! We went through quite a lot together and I do not want to know how this would have been without him! A person with tremendous influence on my personal life in Paris and personal development is Lexie Bucholtz. She inspired me, opened my eyes on many

things outside of the lab in Paris! Thanks Lexie for an unforgettable time! I thank my friends Ralfs Buks, Ana Loncar, Linda Silina, Pablo Canales, Fadi Jebbawi, Vincent Galand, Guilherme Nader, Mohamad Kamal for countless wonderful nights, (table-) tennis matches and/or other moments during different phases of the past 4 years in Paris. I would like to give special thanks to Julia Torné. A very special and wonderful woman that I thank very sincerely for all her support and the time we spend together.

Finally, I would like to thank some special people from Germany (a few words in German):

Ein ganz besonderer Dank kommt meiner Familie und vor allem meinen Eltern zu. Ich danke euch von ganzen Herzen für alles was ihr für mich getan habt, für eure Unterstützung und Liebe! Das Gleiche gilt für meine Oma, auf die ich mich blind verlassen kann und die mich immer unterstützt! Ich danke meinem Bruder Alex, der hinter mir steht, wenn nötig! Danke an die Jungs: Gerrit, Finn, Chenming, Marius, Flo, Lars, Kevin, Robert! Die Gruppe ist ein enorm geschätzter Teil meines Lebens trotz der Distanz und meiner leider langen, nicht nur physischen Abwesenheit. Das Gleiche gilt auch für Christian Meyer. Ich bin sehr dankbar für unsere enge Verbindung und wirklich inzwischen jahrzehntelange Freundschaft. Ich danke auch meinen Ex-Potsdam Kommilitonen/innen und vielen weiteren Menschen, die mir auf dem Weg zum Doktor in vielerlei Hinsicht geholfen haben. Dann ist da noch Justus Ninnemann: Ca. 30 Jahre Freundschaft sprechen für sich! Merci!

2. RÉSUMÉ

Au cours de la mitose, les cellules de mammifères doivent conserver un centromère unique afin d'assurer une ségrégation correcte des chromatides sœurs dans chaque cellule. Malgré la présence de séquences d'ADN répétitives dans la plupart des organismes, les centromères sont marqués épigénétiquement par une histone H3 spécifique appelée CENP-A (CENtromeric Protein A) via un mécanisme d'autoréplication. Il n'est pourtant pas certain que CENP-A soit l'unique marqueur de la position du centromère chez l'Homme ou si d'autres facteurs, tels que la séquence d'ADN, y contribuent également.

Afin d'évaluer cette hypothèse, j'ai récemment développé le système CENP-A$^{OFF/ON}$ qui permet d'éliminer l'identité du centromère dans les cellules humaines en quelques minutes. En utilisant ce système unique, je peux identifier le(s) mécanisme(s) qui permet(tent) la réintégration *de novo* de CENP-A à sa position initiale et la formation du centromère qui en résulte, en temps réel, sur des chromosomes humains. Il est intéressant de noter que c'est la présence de la protéine CENP-B liée à l'ADN répétitif qui va permettre la réintégration de CENP-A, de nouveau exprimée, à la position initiale. En revanche, l'absence de CENP-B et d'autres marqueurs épigénétiques CENP-A-dépendants au centromère, empêchent toute nouvelle déposition de CENP-A ce qui entraîne la formation d'un néocentromère. La spécification du centromère liée à CENP-B a lieu, en partie, indépendamment de CENP-C, composant clé du centromère connu pour son rôle dans le recrutement de CENP-A. Cet ensemble de résultats démontrent que le mécanisme bien établi d'auto-assemblage de CENP-A n'est pas essentiel pour le maintien de l'identité du centromère et que CENP-B joue un rôle clé dans le maintien de la position du centromère. Ainsi, ceci permet d'accroître nos connaissances, actuellement restreintes, concernant les procédés (épi)génétiques qui contrôlent l'identité du centromère, essentiel pour la transmission correcte du matériel génétique. Enfin, dans le cadre de cette thèse, nous avons identifié une population de cellules T CD4+ au repos CENP-A-négatives, CENP-B/C-positives, capables de ré-exprimer et de réassembler le CENP-A à l'entrée du cycle cellulaire, ce qui démontre l'importance physiologique de la mémoire génétique.

Mots-clés : CENP-B, centromère identité, CENP-A, CENP-C, l'ADN alpha satellite, la stabilité du génome

3. ABSTRACT

Centromeres are built on repetitive DNA sequences (CenDNA) and a specific chromatin enriched with the histone H3 variant CENP-A, the epigenetic mark that identifies centromere position. During my thesis project, I interrogated the importance of CenDNA in centromere specification by developing a system to rapidly remove and re-activate CENP-A (CENP-A$^{OFF/ON}$). Using this system, I define the temporal cascade of events necessary to maintain centromere position. I unveil that CENP-B bound to CenDNA provides memory for maintenance on human centromeres by promoting *de novo* CENP-A deposition. Indeed, the lack of CENP-B favors neocentromere formation under selective pressure. Occasionally, CENP-B triggers centromere re-activation initiated by CENP-C, but not CENP-A, recruitment at both ectopic and native centromeres. This is then sufficient to initiate the CENP-A-based epigenetic loop. Finally, in the frame of this thesis, we identified a population of CENP-A-negative, CENP-B/C-positive resting CD4^{+} T cells capable to re-express and reassembles CENP-A upon cell cycle entry, demonstrating the physiological importance of the genetic memory.

Keywords: CENP-B, centromere identity, CENP-A, CENP-C, alpha satellite DNA, genome stability

4. INTRODUCTION

I. The Nucleus

The genomic DNA of eukaryotes is organized within the nucleus making the nucleus the main repository of genetic information of the cell. A two-lipid bilayer membrane, the nuclear envelope, separates the inside of the nucleus (nucleoplasm) from the cytosol. The nuclear envelope is structurally supported by a network of intermediate filaments called the nuclear lamina, altogether protecting and structuring the nucleus (Gruenbaum and Foisner, 2015).

The human genome contains a set of 22 autosome pairs and two sex chromosomes (XX or XY). Each human diploid cell has around 6 billion base pairs. Given an approximate length of around 0.34 nm for each base pair (bp) the total length of extended genomic DNA spans around 2 meters (Annunziato, A. 2008). This length is extraordinary considering the ~20 µm diameter of an average human nucleus. Storage of DNA inside such a small compartment, while still allowing for vital DNA related life processes such as transcription, is achieved via a dynamic DNA packaging system. In nearly all eukaryotes, DNA is wrapped around architectural proteins called histones in a first level of DNA packaging (Annunziato, A. 2008). Compaction is further increased by folding DNA and histones into higher order structures. All levels of compaction must be reversible to allow for protein or RNA access whenever this is necessary. The dynamic packaging concept is conserved even in bacteria where HU proteins can fulfill histone functions demonstrating the essentiality of the principle for life in general (Talbert and Henikoff, 2010).

I.1. Histones, Nucleosomes and Chromatin

Bulk DNA packaging is mostly mediated by the four canonical core histones H2A, H2B, H3 and H4. An octamer comprising two H2A-H2B heterodimers and two H3-H4 heterodimers is loaded with 146/7 bp DNA. This DNA-protein complex forms the so-called nucleosome core particle. The DNA wraps around the histone octamer in 1.7 turns in a left-handed conformation in which the H3/H4 histones associate with the ends of the DNA regardless of the underlying DNA sequence (Henikoff and Smith, 2015). Nucleosome core particles are interconnected by a linking DNA stretch of approximately up to 80-bp, which collectively is called the nucleosome. The core histones contain

a histone fold domain and characteristic histone tails that can protrude from the nucleosome core particle. Histone tails are major target sites of diverse post-translational modifications (PTMs) which can lead to a change of chromatin compaction or contribute to other histone related functions (Talbert and Henikoff, 2010).

Histones are involved in many fundamental and essential processes like gene expression and DNA damage protection. Non-canonical histone variants can replace canonical histones in the nucleosome and are implicated in a wide range of processes for example DNA double strand repair (H2A.X), histone replacement during transcription (H3.3), promotor insulation (H2A.Z) or sex chromosome inactivation (phospho-H2A.X, macroH2A, and H2A.B) (Weber and Henikoff, 2014). A special histone H3 variant is CENtromere Protein A (CENP-A) that marks the centromere position and is essential for the assembly of the centromere. Structure, function, assembly and maintenance of CENP-A will be discussed in depth in the centromere chapter of this introduction (**section 2**).

Another type of histone proteins are linker histones. These proteins are not part of the nucleosome core particle but commonly bind to a small part of the DNA in between two nucleosomes. It is generally considered that this binding protects the linking DNA and contributes to the structural arrangement of nucleosomes by influencing the relative angle between the nucleosomes and/or increasing the DNA flexibility by neutralizing negative DNA backbone charge (Cutter and Hayes, 2015). Linker histones are involved in the formation and stabilization of higher order arrangement of nucleosomes (Lyubitelev et al., 2016).

An array of nucleosomes in a so-called “beads-on-a-string” conformation is called a 10-nm chromatin filament based on its thickness. These filaments can arrange to 30-nm chromatin fibers supported by linker histones. Fibers in turn can fold into higher order structures providing a higher level of DNA compaction. More recent advances in chromatin imaging techniques also suggest that rather than orderly folded, chromatin mainly exist in disordered chromatin chains with different chromatin densities that in turn allow faster re-organization of chromatin (Ou et al., 2017). Hence, despite great progress, the overall chromatin organization and its dynamics is not fully understood.

Depending on its accessibility, chromatin is commonly categorized in two classes: heterochromatin and euchromatin (**see section 1.3**). Most of the DNA in interphase is heterochromatic. Chromatin compaction can be regulated by protein complexes that bind and wind chromatin loops such as condensins. Especially upon entry into mitosis or meiosis, condensins further increase compaction by 2-3-fold and dramatic structural changes transform chromatin into individual rod-shaped chromosomes. A triggering event for this massive rearrangement is a PTM on

histone H3 (H310P). Chromatin hyper-condensation is a critical function for faithful chromosome segregation (**see section 1.4**).

I.2 Histone Assembly

Histone deposition into chromatin or eviction as well as histone transfers, stabilization and storage are mediated by different histone chaperones and/or assembly factors. Histone assembly can occur during DNA replication (RC, Replication Coupled) or in a Replication Independent (RI) manner. During DNA replication, histones are disassembled from the parental DNA strand and subsequently re-assembled on the daughter strand. For rapid histone assembly on daughter strands, DNA replication machinery and RC assembly are tightly linked (Annunziato et al., 1982; McKnight and Miller, 1977; Sogo et al., 1986). Nucleosomes are either assembled *de novo* or parental histones are inherited to daughter strands as H2A-H2B or H3-H4 tetramers, which ensures the conservation of epigenetic marks through replication. A variety of histone chaperones were shown to be involved in RC histone assembly such as CAF-1 (Chromatin Assembly Factor 1), ASF1 (Anti Silencing Function) and FACT (FAcilitates Chromatin Transcription). In the first step of RC chromatin assembly the H3-H4 tetramers are deposited by the histone chaperone complexes CAF-1/ASF-1 which can interact with the DNA sliding clamp protein PCNA (Proliferating Cell Nuclear Antigen) [reviewed in (Loyola and Almouzni, 2004)]. H2A-H2B tetramers rapidly follow H3-H4 to complete the nucleosome. This process is likely supported by NAP1 (Nucleosome Assembly Protein 1) or FACT chaperones. The mechanism for histone inheritance during replication is not fully understood but again the PCNA in complex with CAF-1, as well as the MCM2-7 (Mini Chromosome Maintenance) helicase complex in connection with the chaperones ASF1, FACT or HJURP (Holliday Junction Recognizing Protein) seem to be important players in this process (Alabert and Groth, 2012; Burgess and Zhang, 2013; Gérard et al., 2006; Ransom et al., 2010; Shibahara and Stillman, 1999; Zasadzińska et al., 2018). After DNA replication, the two daughter strands become tied to each-other via cohesin, a connection that is removed during mitosis (see section 1.4).

The histone variant H3.3 is deposited genome-wide in a RI mechanism involving the HIRA chaperone or alternatively, involving DAXX (Death Domain Assoicated Protein) chaperone (involving ATRX) for telomeric and pericentromeric deposition. CENP-A deposition also occurs in a RI fashion and follows a unique highly regulated mechanism which is discussed in greater detail in section 2.2.

I.3 Chromatin and Epigenetics

The chromatin state is crucial for many nuclear functions. The term epigenetics - although differently defined over time - describes the chromatin adaptations that can be inherited through DNA replication, mitosis and even transgenerationally through meiosis (Bird, 2007). As indicated in the Greek prefix “epi” (above), epigenetic states and events do not evoke or require DNA sequence alterations (Bird, 2007; Waddington, 1954). Epigenetic mechanisms are most commonly associated with the establishment of gene expression pattern of a cell, a process that is especially important during development and allows organisms to adapt to environmental stimuli (Bird, 2007). Epigenetic mechanisms also include non-gene regulatory functions as they impact the overall chromatin structure. The epigenetic identification of the centromere position which is of greater interest for this work is another example of an epigenetic mechanism (**see section 2.3**).

On a molecular basis chromatin states adopt due to the establishment and action of epigenetic marks. These include DNA methylation, PTMs of histones, histone variants and non-coding RNAs as well as binding of structural and regulatory proteins that can identify chromatin states and/or remodel the three-dimensional chromatin architecture and accessibility (Bird, 2007).

While euchromatic regions are less compacted and transcriptionally active, heterochromatin instead is further condensed, and considered transcriptionally inactive. Heterochromatin can be sub-grouped into facultative and constitutive states. The functional relevance of the latter is not completely understood but it likely contributes to protect the genome from damage by for example regulating nuclear stiffness in response to environment changes and can permanently silence transposons (Haaf and Schmid, 1991; Nava et al., 2020). A dense layer of constitutive heterochromatin can be found in particular at the nuclear periphery in most mammalian cells (Buchwalter et al., 2018; Steensel and Belmont, 2017). The existence of constitutive heterochromatin domains is further proposed to rely on an intrinsic demixing phenomenon called phase separation (Strom et al., 2017).

Facultative heterochromatin are chromatin regions which are potentially active (Haaf and Schmid, 1991). Since chromatin states are convertible, transient intermediate states of neither “open” euchromatin nor “closed” heterochromatin could exist. Indeed, in reality different eu- and heterochromatin states can be found harboring a mix of “open” and “close” chromatin marks not only as transient appearances (**see section 2.2d** for the special chromatin state of the centromere).

I.4 Chromosome Segregation

Proliferating cells undergo a cell cycle comprising a division phase (M-phase) and an interphase which is subdivided into a G1-, S- and G2-phase. Cell division requires DNA duplication during interphase which normally occurs in S-phase. In M-phase, duplicated genomic DNA is then equally transmitted to both daughter cells. This is achieved via a fascinating, highly complex segregation process of the sister chromatids during mitosis or of the bivalents during meiosis (**Figure 1**). The centromere plays a pivotal role in this process. I will here briefly summarize some of the key events and regulatory processes that coordinate faithful chromosome segregation with a focus on human mitosis.

Following chromosome condensation and, in many species such as human, nuclear envelope breakdown during prophase chromosomes attach to the spindle apparatus. The spindle apparatus is formed by microtubule fibers which are hollow protofilament polymers that display a dynamic instability (Mitchison and Kirschner, 1984). This property allows lengthening and shortening of the fibers by rapidly assembling and disassembling microtubule subunits called alpha and beta tubulin (Mitchison and Kirschner, 1984). Spindle fibers have an inherent polarity and emanate from microtubule organizing centers which are found on opposite cell poles. The plus ends of the spindle fiber attach to chromosomes (Maddox et al., 1999) (**Figure 1**).

Two structural units at the chromosome, easily discriminated by electron microscopy during mitosis, link the chromosome to spindle fibers: centromeres, which define the spindle to chromosome interaction region along the chromosome, and kinetochores that assemble on top of the centromeres and directly interact with spindle fibers (then called Kinetochore fibers, K-fibers). After spindle fiber attachment, chromosomes congress at the equatorial plane of the cell (Maiato et al., 2017).

At this point the sister chromatids or bivalents are solely connected at their (peri)centromeric regions. This crucial connection is maintained by the protection of cohesin, a ring-shaped protein complex that holds the sister chromatids tightly together after DNA replication (Peters et al., 2008). While cohesin removal at chromosome arms is initiated already in early prophase by the cohesin release factor WAPL (Wing APart-Like) (indirectly activated via several mitotic kinases), the activity of WAPL is inhibited at centromeric regions via Shugosin (Haarhuis et al., 2014; Uhlmann,

2000). Only when all chromosomes are properly attached and aligned at the cell equator is cohesin finally removed from the centromeric region by separase activity, a protease that can hydrolyze a cohesin subunit (Silkworth et al., 2011). Subsequently, the sister chromatids start to migrate to opposite poles mainly driven by the depolymerization of k-fibers (anaphase) (Maddox et al., 2002) (**Figure 1**).

The kinetochore has been described as a functional versatile protein complex that not only forms load-bearing spindle fiber attachments but also mediates sensor and regulatory functions. The core of the kinetochore is built by a 10-subunit protein complex called the KMN network (Musacchio and Desai, 2017).

The KMN network is named after three functional distinct sub-complexes – the **K**NL1 (Kinetochore Null Protein 1), **M**is12 (MIS-segregation 12) and the **N**dc80 (Nuclear division cycle 80) sub-complex [reviewed in (Musacchio and Desai, 2017)]. The Mis12 complex is at the root of the kinetochore and binds to the centromere on one hand (**see section 2**) and to the Ndc80 and KNL1 complexes on the other hand. The Ndc80 complex forms major load-bearing attachments with spindle fibers (Alushin et al., 2010) and together with the KNL1 complex has regulatory function for cell cycle progression [reviewed in (London and Biggins, 2014)]. Assembly of the KMN network occurs at the onset of mitosis (Hara and Fukagawa, 2018; Hori et al., 2003). After mitosis, kinetochore proteins are disassembled while centromere components either remain chromatin associated and/or are replenished in the following interphase [reviewed in (Hara and Fukagawa, 2018)].

During mitosis, kinetochores possess regulatory function to ensure proper chromosome to spindle attachment. Chromosomes are wrongly attached to the spindle apparatus when both kinetochores of the same chromosome connect to spindle fibers from the same cell pole (called synthelic attachment) or when a single kinetochore is attached to fibers from both spindle poles (merotelic attachment) [reviewed in (Murray, 2011)]. Microtubule-derived pulling forces generate tension across sister-kinetochores when a chromosome is correctly aligned. This tension intrinsically stabilizes spindle-kinetochore interactions (Akiyoshi et al., 2010; Cheeseman et al., 2008). In addition, in absence of tension the protein kinase Aurora B phosphorylates kinetochore proteins. The Ndc80 complex is one of the major targets of Aurora B. Phosphorylation destabilizes spindle fiber

attachment which allows for a new correct attachment of the sister kinetochores (Cheeseman et al., 2008).

Kinetochores also indirectly mediate mitotic arrest in the absence of spindle fiber attachments. In such a case, the unattached kinetochore activates the so-called Spindle Assembly Checkpoint (SAC) [reviewed in (Lara-Gonzalez et al., 2012)]. A major driver of SAC signaling is the MPS1 (MonoPolar Spindle 1) kinase. MPS1 binds to unattached kinetochores and phosphorylates proteins of the KMN network. In brief, through a cascade of events MPS1 kinetochore phosphorylation lead to the formation of a protein complex known as the mitotic checkpoint complex (MCC; components: Mad2, Cdc20, Bub3, Mad3) [reviewed in (London and Biggins, 2014)]. The MCC mainly sequesters a protein called Cdc20 which is an activator of the APC/C (Anaphase Promoting Complex/Cyclosome) ubiquitin ligase (Hoyt, 2001; Murray, 2011). APC/C activity is required for targeting securin and cyclin B for proteosomal degradation. In human, securin inhibits separase and hence cohesin removal at the centromere region and cyclin B activates CDK1 (cyclin dependent kinase 1). Anaphase onset requires degradation of both proteins to allow sister chromatid separation and CDK1 inactivation which terminates checkpoint activity and crucially stabilizes spindle fiber-kinetochore attachments.

Error correction and SAC are intertwined mechanisms indicated by the involvement of several proteins (e.g. MPS1) in both processes. The molecular mechanism underlying the SAC and error correction pathways are just briefly touched in this introduction. The complex matter is further subject of ongoing research.

In the last phase of mitosis (telophase) separated chromosome de-condense and the nuclear envelope is reformed (**Figure 1**). Cell division is terminated in a process called cytokinesis in which the cytoplasm of the parental cell is divided [reviewed in (Srivastava et al., 2015)]. To this end, cells form a protein ring structure under the cell membrane in between the two separated chromosome masses. The contraction of this ring forms a cleavage furrow in between the daughter cells and finally separates the cytoplasm of the divided cell.

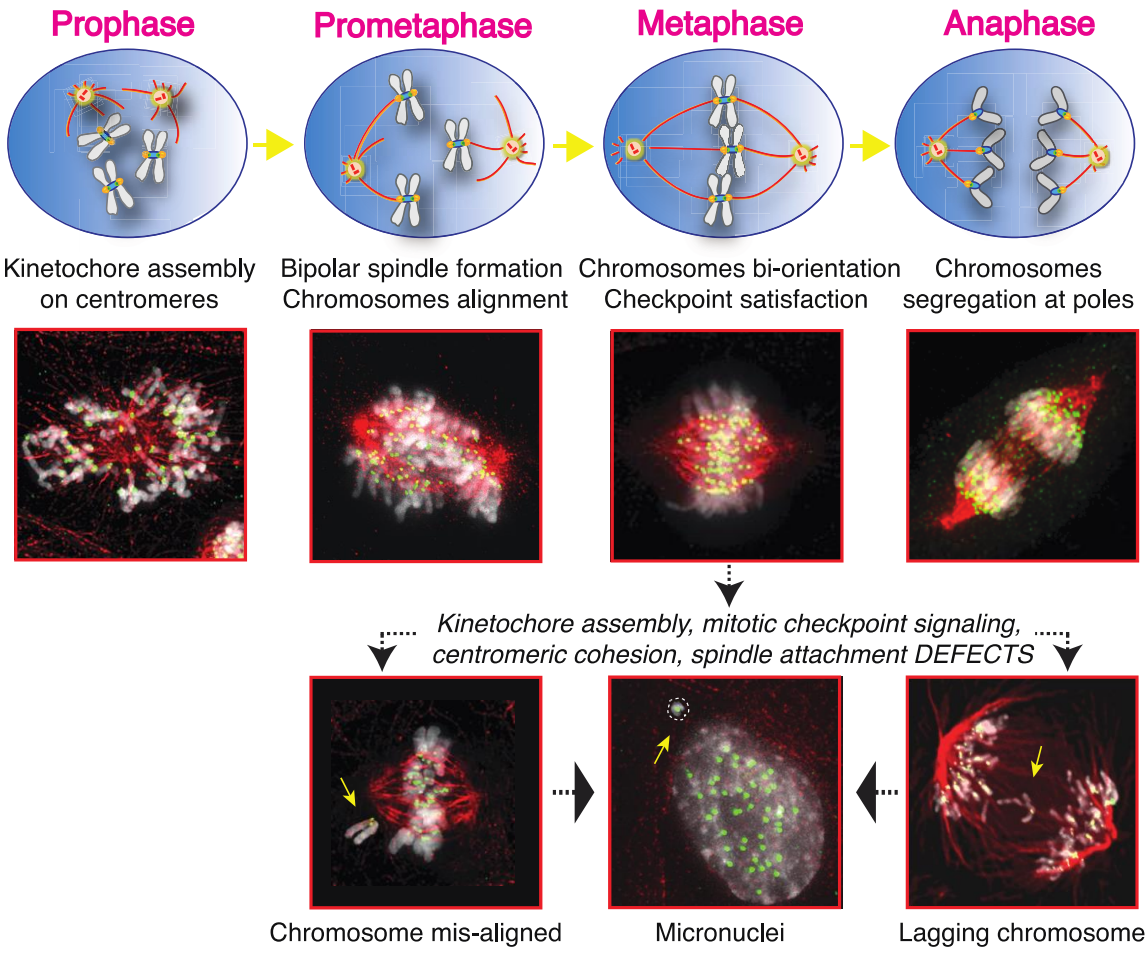


Figure 1: Chromosome segregation.

Top panels show a schematic illustration of chromosome segregation during mitosis. Middle panels show representative immunofluorescence images of cells in mitotic phases as indicated in the top panels. Centromeres are stained in green, α-tubulin in red and DNA in grey. After chromosome condensation kinetochores interact with the spindle apparatus to congress all chromosomes in the equatorial plane (metaphase) before next the sister chromatids are equally separated and pulled to the cell poles. Finally, chromosomes decondense in telophase (not shown). Lower panel show mitotic errors (left and right images) and a cell possessing a micronucleus as a consequence of a mitotic error (middle panel). Images from D. Fachinetti.

I.5 Chromosome Segregation Error

The spindle assembly checkpoint and error correction mechanism are important mechanisms of the cell to control proper chromosome segregation. Defects in their function can lead to errors in

chromosome segregation and as a consequence to aberrant chromosome number (numerical aneuploidy) or gain or loss of chromosomes parts (structural aneuploidy) in the daughter cells [reviewed in (Santaguida and Amon, 2015)] (**Figure 1**). In healthy human cells a surveillance system based on the tumor suppressor p53 is activated when cells become aneuploid which stops cells from proliferating and can induce senescence or cell death. Aneuploidy is indeed a hallmark of cancer cells and found in genomic disorders although causality and consequence of chromosome segregation defects and disease remain still elusive. Aneuploidy can cause genomic instability (increased chromosome missegregation, increased mutation rates, replication stress, defects in DNA repair) (Santaguida and Amon, 2015).

If anaphase onset occurs with unattached or incorrectly attached centromeres these chromosomes will not be efficiently segregated and will lag behind during anaphase (**Figure 1**). During cytokinesis, lagging chromosomes can form chromosome bridges between the daughter cells and can break in the cleavage furrow (Crasta et al., 2012; Ganem et al., 2009; Janssen et al., 2011; Umbreit et al., 2020). Moreover, when cells reassemble the nuclear envelope at mitotic exit, lagging chromosomes can be excluded from the main nucleus as a consequence of their physical distance. They form a separate small nuclear entity, called micronucleus, harboring a fragile nuclear envelope in about half of the cases (**Figure 1**). Chromosomes within micronuclei undergo extensive DNA damage and massive structural rearrangements, a process called chromothripsis (Crasta et al., 2012; Hatch et al., 2013; Ly et al., 2017; Stephens et al., 2011).

One prerequisite for preventing segregation errors during mitosis is the proper formation of the centromere since it provides the basis for kinetochore to spindle connections (**section 2**). Moreover, the centromere in many species must be restricted to a unique position. Indeed, gain or loss of centromeres can lead to chromosome segregation errors. Chromosomes with two centromeres (dicentric chromosomes) can arise when two chromosomes are fused as a consequence of DNA breakage events (Hill and Bloom, 1987; Stimpson et al., 2012). During mitosis, dicentric chromosomes tend to form chromosome bridges during anaphase which in turn can lead to another breakage and chromosome fusion event. These Breakage-Fusion-Bridge (BFB) cycles were first observed by Barbara McClintock in maize (McClintock, 1941, 1939). Similar events can also occur when neocentromeres are formed (**see section 2.6**).

II The Centromere

Humans and many other eukaryotic species possess one spatially delimited centromere per chromosome which defines them as monocentric species. Ensuring the maintenance of a unique centromere is critical in these monocentric organisms since loss of the centromere or gain of an additional centromere has been found to cause defects in chromosome segregation during cell division as described in the previous section.

The first descriptions of centromeres date back to the work of Walter Flemming at the end of the 19th century who first recognized primary constrictions on chromosomes (Flemming, 1882). 54 years later Cyril Darlington introduced the term centromere (Darlington 1936). This name could wrongly imply that the centromere is found in the center of each chromosome while in reality centromere positions vary among different chromosomes. Depending on the relative centromere position chromosomes are categorized in three types: Metacentric or submetacentric when the centromere is found roughly in a central position of the chromosome; acrocentric when the centromere is very close to the tip of a chromosome (telomere); and telocentric when the centromere is found in very close proximity to the telomere. Humans do not possess telocentric chromosome. The concept of possessing a single centromere per chromosome is not universal in eukaryotic organisms. Holocentric organisms like the model organism *Caenorhabditis elegans*, several insects or lower plants present exceptions where the centromere is extended along the entire chromosome (Melters et al., 2012).

In the following chapters, I will summarize major aspects of centromere function, structure, maintenance, DNA and *de novo* formation with a focus on organisms possessing one centromere per chromosome (monocentric) and a special focus on the human case.

II.1 The "Meshwork" and its "Blueprint" - Organization and Function of the Centromere Protein Network

The centromere consists of special chromatin (**see section 2.2**) and 16 chromatin-associated proteins which form the Constitutive Centromere Associated Network (CCAN) (**Figure 2A**). In contrast to kinetochore proteins CCAN components, as indicated by the name, are constitutively present at the centromere throughout the cell cycle. The replenishment of the CCAN during the cell

cycle relies mostly on its underlying centromeric chromatin bound by CENP-A (**see section 2.2a/b**). In turn, in many organisms studied so far, the maintenance of CENP-A is also dependent on components of the CCAN. This interdependency crucially preserves a unique genomic location of the centromere during each cell cycle (**see section 2.3**). The following section will introduce the CCAN and its assembly.

a) CCAN organization and assembly

The 16 proteins of the CCAN build the foundation for the kinetochore providing an absolute vital function for chromosome segregation (**section 1.4 – 1.5**). Most CCAN components were initially identified by mass spectrometry analysis (Foltz et al., 2006; Izuta et al., 2006; Okada et al., 2006). Extensive *in vivo* studies using knock-down and conditional knock-out as well as *in vitro* reconstitution approaches have further shed light on the structure and protein interactions present in the network.

A central protein of the CCAN is CENP-C which acts as an assembly hub, provides CCAN stability and directly connects the centromeric chromatin and the KMN network (Carroll et al., 2010; Fukagawa et al., 1999; Klare et al., 2015; Milks et al., 2009). In addition to CENP-C four other CCAN sub-complexes consisting of the proteins CENP-L/-N, CENP-H/-I/-K/-M, CENP-T/-W/-S/-X and CENP-O/-P/-Q/-R are described (Musacchio and Desai, 2017) (**Figure 2B**). The ability of CENP-C to bind to CENP-A nucleosomes and to act as an assembly hub for many CCAN proteins raises the question if CENP-C is - after CENP-A - upstream in the assembly pathway of the CCAN complex or if there is no clear hierarchy in the CCAN assembly.
Indeed, CENP-C is not the only CCAN protein that makes direct contacts with centromeric chromatin. CENP-L/-N was also found to directly interact with CENP-A nucleosomes. Moreover, the hetero-tetramer CENP-T/-W/-S/-X forms a nucleosome-like structure that binds to 80-100 bp DNA. All of these CCAN-chromatin connections likely increase the stability of a fully assembled CCAN complex (Fukagawa et al., 2001; Hoffmann et al., 2016; Kwon et al., 2007; Nagpal et al., 2015). Indeed, CENP-C depletion leads to a strong reduction of CCAN proteins but not necessarily to an immediate and complete loss of directly or indirectly CENP-C associated CCAN proteins (Kwon et al., 2007). However, despite a certain weak CENP-C-independent stability of the CCAN especially *in vitro* reconstitution experiments of the CCAN unveiled a CENP-C based hierarchy of

the CCAN assembly suggesting that CENP-C functions as “the blueprint” of the network (Klare et al., 2015; Pesenti et al., 2018).

A powerful tool to decipher the sequence of assembly events *in vivo* are ectopic tethering assays. In human cell culture, a frequently used system is a U-2 OS cell line developed by Janicki et al. 2004 in which an array of 200 copies of a 256×LacO/96×tetracycline-responsive element (TRE) was integrated into a sub-telomeric location of chromosome 1 (Janicki et al., 2004). The array allows tethering of LacI or tetR tagged protein constructs at this non-centromeric location. *In vivo,* using the LacO/LacI tethering system, CENP-C recruitment to a LacI-CENP-A was found to strictly precede CENP-T (Tachiwana et al., 2015) and CENP-T recruitment in human required both CENP-C and the N-terminal tail of CENP-A (Logsdon et al., 2015). The CENP-H/-I/-K/-M complex is essential for CENP-T centromere localization and indirectly links CENP-T with CENP-C (Basilico et al., 2014; McKinley et al., 2015) (**Figure 2B**). The CENP-T and CENP-H/-I/-K/-M interaction may be weak which could explain why CENP-T alone is unable to provide enough stability for CENP-H/-I/-K/-M and in turn CENP-C recruitment (Gascoigne et al., 2011; Pesenti et al., 2018).

The second chromatin associated complex CENP-L/-N requires, despite its direct interaction with CENP-A, the presence of both CENP-A (Hoffmann et al., 2016) and other CCAN proteins (likely CENP-C or HIKM) for its centromere recruitment during mid to late S-phase (Carroll et al., 2009; Chittori et al., 2017; Fang et al., 2015; Pentakota et al., 2017; Tian et al., 2018). This indicates that CENP-L/-N is assembled downstream of CENP-A and CENP-C similarly to CENP-T/-W/-S/-X.

In contrast to CENP-L/-N and the CENP-T complex, human CENP-A alone is sufficient for CENP-C recruitment as demonstrated by a variety of *in vitro* and *in vivo* approaches (Barnhart et al., 2011; Carroll et al., 2010; Fachinetti et al., 2013; Falk et al., 2016, 2015; Guse et al., 2011; Hoffmann et al., 2016; Kato et al., 2013; Logsdon et al., 2015; Tachiwana et al., 2015). The regulation of this interaction remains not fully understood. CENP-A is indeed also found in greater numbers at non-centromeric locations (**see section 2.2a**) and it was reported that not all centromeric CENP-A nucleosomes associate with CENP-C (Kyriacou and Heun, 2018; Melters et al., 2019). On the other hand, non-centromeric CENP-A tethered to a LacO site is able to recruit CENP-C *de novo* (Barnhart et al., 2011; Logsdon et al., 2015; Tachiwana et al., 2015). Importantly, this system provides a massive local CENP-A accumulation far beyond the level found at native centromeres (Barnhart et al., 2011; Tachiwana et al., 2015) but some studies even failed to observe CENP-A induced centromere formation (Gascoigne et al., 2011; Hooser et al., 2001). Moreover, CENP-A overexpression also leads to elevated CENP-A incorporation along chromosome arms and is

accompanied by CENP-C mis-localization. However, neocentromere formation as a consequence of CENP-A overexpression is not observed in human (Lacoste et al., 2014) (further discussed in **sections 2.3** and **2.6**). In conclusion, efficient CCAN formation may rely on the additional presence of supporting factors that are uniquely found at the centromere (**section 2.3d**).

The CENP-O/-P/-Q/-U/-R complex most likely assembles further downstream of all other CCAN sub-complexes. Like many components of the network, also this complex connects to at least two other CCAN sub-complexes CENP-H/-I/-K/-M and CENP-L/-N and interestingly was recently found to be able to connect to microtubules *in vitro* (Pesenti et al., 2018). CENP-O/-P/-Q/-U/-R would be the only direct connection of the constitutive network with microtubules identified so far. While the functional relevance of this interaction remains untested a speculative functional implication for this is the establishment of stable end-on spindle attachments once the chromosome is bi-oriented (Pesenti et al., 2018). Proteins of the CENP-O/-P/-Q/-U/-R complex are indeed found to be important for proper segregation (Hori et al., 2008b) in chicken and moreover CENP-U in chicken was also suggested to suppress drifting of the centromere but the molecular mechanism of this remains poorly understood (Hori et al., 2016).

Overall multiple cross-connection between CCAN proteins indicate that the complex forms a stable protein meshwork (Klare et al., 2015; McKinley et al., 2015). The assembly of the CCAN seems to derive from a single axis based on CENP-A and CENP-C at least in human (Logsdon et al., 2015; Tachiwana et al., 2015). Moreover, cross-connections between individual CCAN units likely exist (e.g. through dimerization of CENP-C or other unknown interactions (Carroll et al., 2010; Cohen et al., 2008). During the cell cycle, CCAN architecture was further described to undergo conformational changes (Allu et al., 2019; Nagpal et al., 2015) but overall CCAN dynamics during the cell cycle and regulations of this still remain only very partially uncovered (Watanabe et al., 2019). Mitotic centromeres are described as more stable compared to interphase kinetochores suggesting a centromere maturation process during the cell cycle (Ribeiro et al., 2010).

b) Inner kinetochore (CCAN) to outer kinetochore connections

Indirect connection of the CCAN with microtubules are mediated via kinetochore proteins. In human, although derived from a single axis (see previous section), two branches for kinetochore

assembly which involve CENP-C and CENP-T have been described (Gascoigne et al., 2011) (**Figure 2B**).

CENP-C interacts with the Mis12 complex, a four-subunit kinetochore complex comprising MIS12, PMF1, DSN1 and NSL1 that in turn bind the microtubule binding Ndc80 complex (Gascoigne et al., 2011; Przewloka et al., 2011; Richter et al., 2016; Screpanti et al., 2011). The CENP-C-Mis12 complex interaction is enhanced by Aurora B phosphorylation of Dsn1 (Kim and Yu, 2015; Rago et al., 2015). A crystal structure of human Mis12 complex bound to a CENP-C$^{1-71}$ construct revealed the molecular details of CENP-C-Mis12 complex interactions (Petrovic et al., 2016): CENP-C interacts via a ~45-residue amino-terminal domain with all subunits of the Mis12 complex although the main interaction site is a "head" structure formed by Mis12 and PMF1. DSN1 phosphorylation was found to be a negative regulator of the CENP-C-Mis12 complex interaction but was not directly involved in the interaction (Petrovic et al., 2016). Possibly, Aurora B regulation is important to restrict CENP-C-Mis12 interaction exclusively to centromeric region (Musacchio and Desai, 2017).

CENP-T in contrast to CENP-C can directly interact with components of the Ndc80 complex circumventing the requirement of the Mis12 complex. CENP-T possesses a long unstructured N-terminal domain and a C-terminal histone fold domain (HFD) (Musacchio and Desai, 2017). While the HFD connects CENP-T to chromatin the N-terminal long unstructured regions can interact with subunits of the Ndc80 complex (SPC24/25). In human, this interaction requires the phosphorylation of two threonine residues on CENP-T by CDK1 (Rago et al., 2015). Once phosphorylated, one CENP-T molecule can interact with two Ndc80 complexes. Interestingly, a third CDK1 phosphorylation site on CENP-T was described. This site is involved in a CENP-C independent interaction of CENP-T with the Mis12 complex. Hence, via this indirect path, one CENP-T molecule could recruit even a third Ndc80 complex (Musacchio and Desai, 2017). Thus, theoretically, the current state of the art suggests that one centromere/kinetochore unit could connect to microtubules using five interfaces (three via CENP-T, one via CENP-C and one via CENP-O/-P/-Q/-U/-R). In reality, not all these possibilities seem to be occupied. A quantitative approach in human suggest a total of ~244 Ndc80 complexes per kinetochore (and around 14 per individual microtubule) (Suzuki et al., 2015). Using individual CENP-T/C knockdown this analysis suggests that out of ~215 CENP-C molecules present per centromere only ~80 assemble a KMN complex and each CENP-T molecule (~70 per centromere) could contribute to 2 Ndc80 complexes (total >140) (Suzuki et al., 2015). Since CENP-C has multiple functions at the centromere it is tantalizing to speculate that different CENP-C molecules serve different purposes or are unoccupied with KMN to function as connecting units.

Individual depletion of the domain that connects CENP-C to the KMN network is not compensated by CENP-T and vice versa although it remains unclear if this is a dosage effect by the resulting Ndc80 shortage (Gascoigne et al., 2011; Nishino et al., 2013). Interestingly, in other organisms such as *Drosophila*, CENP-T and many other proteins of the centromere were not identified which suggest that, at least in this species, a specific type of centromere kinetochore interaction (here via CENP-C) could be sufficient (Musacchio and Desai, 2017).

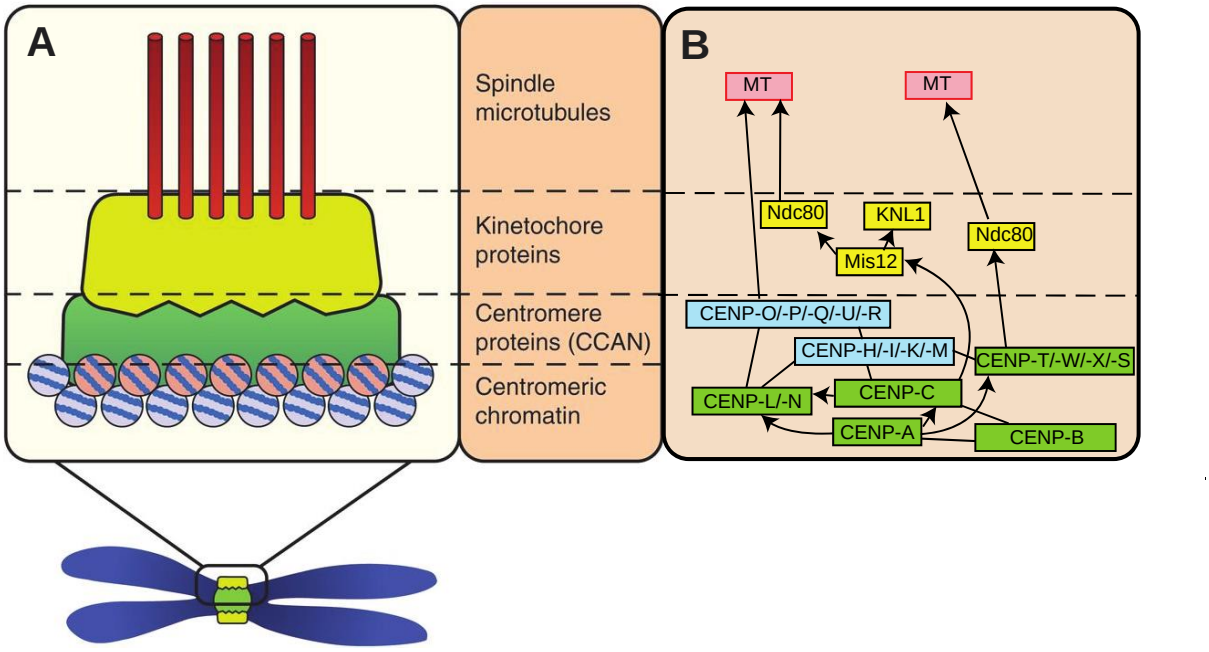


Figure 2: Centromere architecture and organization.

*(**A**) Schematic overview of the centromere/kinetochore complex during mitosis (adapted from (Westhorpe and Straight, 2015)).*

*(**B**) Schematic illustration of centromere/kinetochore organization of mitotic chromosomes. Protein/complexes with direct connection to DNA or chromatin (green), constitutive chromatin associated proteins (green & blue) and outer kinetochore proteins (yellow) are depicted.*

II.2 A region like no other? - Centromeric Chromatin/Epigenetic Environment

In addition to the complex network of centromere proteins described in the previous section, the centromere is built on a unique chromatin region distinctive from the rest of the genome. A key

feature of centromeric chromatin is the accumulation of nucleosomes possessing the histone H3 variant CENP-A.

a) The Histone H3 variant CENP-A

CENP-A together with CENP-B and CENP-C is one of the first centromere proteins identified in human using an anti-centromere-autoantibody (ACA) discovered in patients with a variant of the scleroderma autoimmune disease which is called CREST (Calcinosis/Raynaud's phenomenon/ Esophageal dysmotility/Sclerodactyly/ Telangiectasia) (Earnshaw and Rothfield, 1985; Guldner et al., 1984; Moroi et al., 1980). In all monocentric organisms studied so far, CENP-A molecules accumulate at the centromere (Earnshaw and Rothfield, 1985; Henikoff et al., 2000; Palmer et al., 1987; Stoler et al., 1995). As found in many species, CENP-A is known to be the epigenetic mark required to maintain centromeres at unique locations (**see section 2.3**).

The number of CENP-A molecules found at a single centromere differs between species and even individual centromeres (Bodor et al., 2014). In the allegedly simple case of *Saccharomyces cerevisiae* (budding yeast) point centromeres, chromatin immunoprecipitation (ChIP) analysis initially revealed that CENP-A (called Cse4 in budding yeast) binds to a single nucleosome at the centromere (Furuyama and Biggins, 2007). This leads to the assumption that there is a maximum of two Cse4 molecules per centromere in an octameric Cse4 nucleosome. However, conflicting reports regarding the actual molecule number of Cse4 and the nucleosome conformation in this context exist, ranging from one molecule to eight Cse4 molecules per centromere [reviewed in (Stankovic and Jansen, 2017)]. Higher numbers of centromeric Cse4 molecules may be related to detection errors or to the existence of more difficult to define variable Cse4 molecules that surround the centromere core (Haase et al., 2013).

Despite a greater uncertainty in the determination of the absolute molecule number of centromeric CENP-A, species with regional centromeres show generally higher centromeric CENP-A levels compared to budding yeast (Stankovic and Jansen, 2017) E.g. fission yeast: 26 (Lando et al., 2012) *Drosophila*: 84-336 (Schittenhelm et al., 2010) chicken DT40 cells: 25-62 (Johnston et al. 2010; Ribeiro et al 2010) human: ~400 (Bodor et al., 2014).

Using a 3-D imaging strategy Bodor et al. (2014) assessed the number of internally tagged CENP-A-YFP at the centromere of stable diploid human retinal pigment epithelium cells (hRPE-1). On

average in this analysis a centromere was found to contain ~400 CENP-A molecules (Bodor et al., 2014). In comparison to the total average number of CENP-A molecules found in this cell line (91000 molecules), this number reflects only about 0,44 % of all cellular CENP-A molecules. Summed across all 46 centromeres hence only ~20% of total CENP-A molecules are centromeric. Even more surprisingly, over half of the CENP-A molecules (66%) are non-centromeric bound and the residual CENP-A molecules are soluble. However, taking into consideration the centromere length, there is still a 50-fold local enrichment of centromeric CENP-A compared to non-centromeric chromatin (Bodor et al., 2014). It remains unclear to what extent the generation of such a distinct local enrichment of CENP-A above a certain threshold is a feature required for CENP-A's function as the epigenetic mark of the centromere (**see section 2.3**).

b) The CENP-A Nucleosome

Tachiwana et al. (2011) solved the crystal structure of a nucleosome containing two CENP-A molecules which replaced the canonical histone H3 and possessed left handedly wrapped DNA, which overall is structurally very similar to the canonical nucleosome but with certain distinct differences as discussed below (Tachiwana et al., 2011). The question if the octameric conformation is the sole conformation of CENP-A containing nucleosomes has been a controversial issue in the past (Black and Cleveland, 2011). Various approaches in different model systems suggest that several CENP-A nucleosome adaptations could exist. Among these proposed conformations (neglecting the handedness of DNA wrapping) are CENP-A/H4 tetrameric nucleosomes deprived of H2A and H2B (Dalal et al., 2007; Shivaraju et al., 2012), hemisomes with one CENP-A/H4 and one H2A H2B heterodimer (Henikoff et al., 2014), hexameric nucleosomes (Mizuguchi et al., 2007; Xiao et al., 2011) or the octameric CENP-A nucleosome with two CENP-A molecules (Hasson et al., 2013; Nechemia-Arbely et al., 2017). In human and yeast, data generated using atomic force microscopy or cell imaging in particular proposed that CENP-A nucleosomes can change their conformation in different cell cycle stages (Bui et al., 2012; Shivaraju et al., 2012). Studies from the Black and the Cleveland lab however revealed that *in vivo,* the octamer conformation at human native centromeres and neocentromeres is the major form of CENP-A nucleosomes in human and that any conformational changes could – if present at all - just present a very minor subpopulation or very short-lived transient appearances (Hasson et al., 2013; Nechemia-Arbely et al., 2017). These studies were carried out using a variety of approaches, which importantly, include importantly Micrococcal Nuclease (MNase) digestion in combination with native ChIP in combination with

sequencing (ChIP-Seq). Most importantly, three size classes of MNase-protected CENP-A particles were initially identified with the smallest particle still protecting a DNA fragment that is ~30-50 bp longer than what would be expected for a tetrameric/hemisome conformation (Hasson et al., 2013). Moreover, particle size variation was found to be a consequence of partial DNA unwrapping which is conferred by the physical properties of CENP-A (Hasson et al., 2013; Nechemia-Arbely et al., 2017). In corroboration with data from *in vitro* generated octameric nucleosomes and sequencing approaches, there is thus overall strong evidence arguing for the homotypic (two CENP-A molecules per nucleosomes) octameric core particle conformation as the major form of CENP-A nucleosomes *in vivo* at human centromeres (Hasson et al., 2013; Nechemia-Arbely et al., 2017). Indeed, partial DNA unwrapping and the resulting higher flexibility of the DNA ends of the nucleosome were suggested to be important features for CENP-A's function (Roulland et al., 2016). Interestingly, in a H3-CENP-A-H3 tri-nucleosome sequence the CENP-A nucleosome is more accessible as the tri-nucleosome is less twisted compared to a purely H3 tri-nucleosome which may be important for CCAN nucleation (Takizawa et al., 2020). CENP-C binding to the CENP-A nucleosomes was found to exacerbate the unwrapped "open" configuration of CENP-A nucleosomes (Ali-Ahmad et al., 2019, p.) and at the same time, in collaboration with CENP-N, CENP-C binding stabilizes CENP-A nucleosomes (Falk et al., 2016, 2015; Guo et al., 2017)

Considering an absolute number of approximately 400 CENP-A molecules at the centromere there are on average around 200 CENP-A nucleosomes at the centromere. This also translates to an approximate ratio of CENP-A to H3 nucleosome of 1 to 25 at the centromere compared to 1 to 1200 in non-centromeric regions (Bodor et al., 2014; Stankovic and Jansen, 2017)

Histone H3 variants commonly share very high sequence similarity in their histone folding domains (HFD). CENP-A shows only ~60% sequence homology in this region (Sullivan et al., 1994). CENP-A-H4 pre-nucleosomal tetramers are more compact and show higher conformational rigidity compared to H3-H4 tetramers *in vitro* resulting in significant structural differences (Black et al., 2004). These special physical properties were found to be conveyed by a special domain inside the HFD. Remarkably substituting this domain into H3 is sufficient to target this H3 chimera to the centromere and therefore the domain was called CENP-A Targeting Domain (CATD) (Black et al., 2007, 2004). Importantly, CENP-A interacts with its chaperone HJURP via the CATD which specifically directs the assembly of the pre-nucleosomal CENP-A/H4 complex at the centromere (**see**

section 2.3a) (Hu et al., 2011; Shuaib et al., 2010). The CATD domain spans loop 1 and the alpha 2 helix of CENP-A (see section 2.3, **Figure 5**). Once assembled into chromatin, a specific RG-loop (Arg80/Gly81) within the CATD domain extends from the histone nucleosome and mediates CENP-A interaction with CENP-N (Carroll et al., 2010; Fang et al., 2015; Guo et al., 2017; Sekulic et al., 2010; Tachiwana et al., 2011).

Important differences to other H3 variants are also found in both N-terminal and C-terminal tail regions of CENP-A. Both tails are involved in centromere assembly: The N-terminal tail interacts with CENP-T (Folco et al., 2015; Logsdon et al., 2015) and CENP-B (indirect connection to CENP-C) (Fachinetti et al., 2015) and the C-terminal tail interacts with CENP-C (Carroll et al., 2010; Fachinetti et al., 2013; Kato et al., 2013; Logsdon et al., 2015; Tachiwana et al., 2015; Westhorpe et al., 2015).

c) Post Translational Modifications of CENP-A nucleosomes at the centromere

Since CENP-A molecules are found in a greater number outside of the centromere region it is tempting to speculate that CENP-A PTMs found exclusively at the centromere could restrict CENP-A's function to centromeric regions. In the case of the point centromeres of budding yeast for example the methylation of K37 is indeed described to be a key PTM for CCAN formation (Samel et al., 2012).

In species with regional centromeres a CENP-A methylation site with similar importance has not been described to my knowledge. However, in human, several CENP-A modifications (phosphorylation, acetylation, methylation, ubiquitination) have been reported and functional relevancies of these PTMs have been proposed. Phosphorylation of S68, acetylation/ubiquitination of K124 were suggested to be important for CENP-A deposition at the centromere (in the case of ubiquitination this mark was even proposed to be propagated by CENP-A dimerization at the centromere) and S7 phosphorylation for CENP-C recruitment or Aurora B recruitment to the centromere in human (Eot-Houllier et al., 2018; Goutte-Gattat et al., 2013; Kunitoku et al., 2003; Niikura et al., 2019, 2016, 2015; Yu et al., 2015; Zhao et al., 2016). Neither K124 nor S68 are part of the CENP-A CATD domain nor is S7 even close to the binding site for CENP-C and importantly, all functional relevancies of these modifications were substantially contradicted by our lab and the Black lab (Barra et al., 2019; Fachinetti et al., 2017). There is an ongoing controversy for the

importance of K124 ubiquitination, which can be compensated by a protein tag in a recent publication from the Kitagawa lab (Niikura et al. 2019). Phosphorylation of the residues S16 and S18 are crucial and present in pre-nucleosomal CENP-A. Their absence or hyper- phosphorylation leads to chromosome missegregation (Bailey et al., 2013; Barra et al., 2019).
Additionally, in the same report that described S16 and S18 phosphorylation of CENP-A, amino-terminal tri-methylation of glycine was discovered (glycine is the N-terminal amino acid since Methionine, translated from the start codon, is post-translationally removed from CENP-A) (Bailey et al., 2013). Loss of N-terminal tri-methylation of CENP-A leads to a reduction of CENP-T and CENP-I at the centromere and causes chromosome segregation errors. (Sathyan et al., 2017).
Overall, little is known about the regulation of CENP-A PTMs and only S16/18 phosphorylation have clear functional relevance before CENP-A is deposited at the centromere. Whether there are any functional implications of PTMs on centromeric CENP-A remains unclear.

A special importance has been attributed to mono-methylation of Lysine 20 of H4 which is present in CENP-A nucleosomes but not in a pre-nucleosomal CENP-A/H4 complex (Bailey et al., 2016). Importantly, this PTM is found to be essential for correct kinetochore formation (Hori et al., 2014). Recent structural studies suggested that CENP-C binding induces conformational changes in the CENP-A nucleosome that facilitate H4K20 mono-methylation (Ali-Ahmad et al., 2019) which could suggest that H4 is methylated after CENP-C is bound and might contribute to the stability of the nucleosome or kinetochore. Monomethylation of H4 is also proposed to be facilitated by the CATD domain of CENP-A itself suggesting that the epigenetic mark has a direct effect on the presence of this modification (Arimura et al., 2019).

d) Centrochromatin beyond CENP-A

Centromeric chromatin is exceptional due to the accumulation of CENP-A nucleosomes but the vast majority of nucleosomes present at the centromere do not contain CENP-A (Nechemia-Arbely et al., 2017). Other epigenetic features adjacent to CENP-A nucleosomes contribute substantially to the establishment of a special epigenetic landscape also called centrochromatin which is embedded into pericentromeric heterochromatin (García Del Arco and Erhardt, 2017; Sullivan and Karpen, 2004).

Acetylation of histones is commonly found in regions with transcriptionally permissive chromatin states. At the centromere however, certain histone acetylations are either absent or low in

abundance suggesting that centrochromatin is heterochromatic. The usage of the chromatin fiber technique first revealed the presence of "open" chromatin marks (H3K4me1/2, H3K36me2/3) at the centromeres of flies and human (García Del Arco and Erhardt, 2017; Sullivan and Karpen, 2004; Lam et al., 2006). In addition, using an approach to purify di- or tri-nucleosomes combined with mass spectrometry it was possible to identify histone modification of H3 and H4 histones closely associated with CENP-A nucleosomes (Bailey et al., 2016). This approach additionally identified heterochromatic marks (H3K9me2, H3K27me2) and also low levels of acetylation marks (H4K5Ac and H4K12Ac) with multiple PTMs being present on single nucleosomes (Bailey et al., 2016). Altogether, centromeric chromatin displays a complex blend of modified histones and moreover distinct methylation patterns [reviewed in (Scelfo and Fachinetti, 2019)] ultimately resulting in a chromatin state that is weakly transcriptionally permissive (Erliandri et al., 2014; Mravinac et al., 2009; Sullivan and Karpen, 2004).

Indeed, RNA polymerase II was found to localize at the centromere (and pericentromeric regions) (Bergmann et al., 2012; Chan and Wong, 2012) and centromere-specific transcripts are synthesized (Blower, 2016; Bury et al., 2020; Chan and Wong, 2012; McNulty et al., 2017; Saffery et al., 2003; Wong et al., 2007).

Low level of centrochromatin transcription in multiple species has been proposed to play diverse roles in centromere assembly or stability (association with CENP-C, CENP-A and CENP-B) and function (Blower, 2016; Bobkov et al., 2018; Bury et al., 2020; Du et al., 2010; Ferri et al., 2009; Grenfell et al., 2016; Ideue et al., 2014; McNulty et al., 2017; Quénet and Dalal, 2014; Rošić et al., 2014; Wong et al., 2007; Zhu et al., 2018). Centromere transcription continues even during M-phase where most of the transcription is normally silenced. This is likely caused by the association of RNA Polymerase II with cohesin which is retained at the centromere until anaphase onset (Perea-Resa et al., 2020).

Since the chromatin state seems to be important for centromere functions and/or identity, the question arises how the epigenetic status of centrochromatin is regulated. An important tool to study this are human (or mammalian) artificial chromosomes (HAC) (Barra et al., 2019; Basu et al., 2005; Harrington et al., 1997; Nakashima et al., 2005; Okada et al., 2007; Okamoto et al., 2007; Tachiwana et al., 2013). HAC formation requires the generation of a centromere on a naked large DNA fragment. This can be best induced by targeting centromere proteins to a LacO or tetR array present on the DNA or by the presence of specific centromere sequences (**see section 2.5** and **2.6b**). At the same time, chromatin modifying proteins can be recruited to the same arrays which allows to alter

the epigenetic environment of the artificially generated centromere and test for functional relevance of these perturbations. Hypomethylation of a HAC for example, was found to negatively impact the recruitment of the histone chaperone that mediates CENP-A assembly (**see section 2.3a**) (Bergmann et al., 2011). Recruitment of histone methyltransferases to an active centromere on a HAC was insufficient to silence centromere transcription indicating that centromeres establish mechanisms to prevent heterochromatin spreading (Martins et al., 2016). Similarly, the H3K4me2 mark found in centromeric regions was proposed to be not only important to enable transcription but likely also to somehow establish acetylation of H3K9, which could prevent a spreading of the adjacent pericentromeric heterochromatin (Molina et al., 2016). Such a heterochromatin spreading would cause centromere inactivation. A fine tuning of H3K9me3 and H3K9ac seems to be important on native human centromeres (Molina et al., 2016; Nakano et al., 2008). Mechanistically, this is perhaps regulated by interactions of histone modifiers with centromere proteins (**see section 2.3**) (Nakano et al., 2008; Ohzeki et al., 2016, 2012). Other studies also found that the origin recognition complex 2 (ORC2) when sumoylated recruits the histone H3K4 demethylase KDM5A to the centromere which further ensures low level of transcription (Huang et al., 2016). In *Drosophila*, the trithorax group (Trx-G) has been identified as another important regulator of centrochromatin (Piacentini et al., 2019) and a screen for CENP-A assembly and stability further found a number of chromatin modifying proteins affecting centromeric CENP-A level (Mitra et al., 2020).

DNA methylation might regulate centromeric transcription as well (Scelfo and Fachinetti, 2019). The importance of methylation in the centromeric (and pericentromeric) DNA is strikingly disclosed in the context of diseases (e.g. several tumors and the Immunodeficiency, Centromeric region instability, Facial anomalies (ICF) syndrome, a genetic disease), where hypomethylation of centromeric DNA is associated with centrochromatin (and pericentromeric) decondensation. This further demonstrates that a certain loss of compaction at the centromere is detrimental and accompanies several centromere dysfunctions and genomic rearrangements (Scelfo and Fachinetti, 2019). Furthermore, a functional interplay between the centromere proteins CENP-B as well as CENP-C with DNA methylation has been proposed but the detailed effects of aberrant DNA methylation at the centromere and how this is prevented remains very poorly understood (Scelfo and Fachinetti, 2019).

Altogether this shows that regulation of a certain centromere chromatin environment is a crucial task to ensure centromere function.

Centromeric chromatin undergoes changes during the cell cycle (**see section 2.3**). Most remarkable during DNA replication CENP-A nucleosomes are diluted to both daughter strands as their replenishment is uncoupled from DNA replication (Jansen et al., 2007a; Shelby et al., 2000) (**see section 2.3**). The histone variant H3.3 has been described as the placeholder for CENP-A nucleosomes indicating a special importance of H3.3 at the centromeric region (Dunleavy et al., 2011). During mitosis moreover, the mitotic kinase haspin phosphorylates H3 at the centromere (H3T3) specifically in the inner centromeric regions. This is important to recruit the chromosomal passenger complex to the centromere and has thus functional relevance for proper chromosome congression and error correction (Dai and Higgins, 2005; Kelly et al., 2010; Wang et al., 2010; Yamagishi et al., 2010).

e) Pericentromeric chromatin

Centrochromatin is embedded into a highly heterochromatic region formed on the so-called pericentromere region which is rich in histones with PTMs associated with heterochromatin and DNA methylation (Peters et al., 2008; Rice et al., 2003; Scelfo and Fachinetti, 2019). The transcriptionally repressive marks H3K9me3 and H3K27me3 are very abundant at the pericentromere and cause the recruitment of the heterochromatin protein 1 (HP1). This in turn creates a positive feedback-loop for heterochromatin formation as it recruits methyltransferases (e.g. SUV39h) which trimethylate H4K20, an additional repressive mark and promote the spreading of heterochromatin (Schotta et al., 2004). Functionally, pericentromeric heterochromatin is well known to be important for the recruitment of cohesin to the centromere region which is important for sister chromatid cohesion prior to anaphase onset (**see section 1.4**) (Sakuno and Watanabe, 2009; Watanabe, 2005).

Changing the epigenetic landscape of the pericentromere by for example, hypomethylation is associated with genomic instability but is also found in special tissue like paternal or maternal pronuclei (Déjardin, 2015; Scelfo and Fachinetti, 2019). Among other effects hypomethylation for example, promotes the recruitment of Polycomb proteins to the pericentromere which are important transcription regulatory proteins involved in gene silencing. The deposition of Polycomb at the pericentromere could be a triggering event for gene expression alteration in tissue and disease (Déjardin, 2015).

f) Centromere nucleosome sequence and centromere architecture

An important question to better understand the structural organization of the centromere is the identification of the nucleosome sequence and the 3-D organization of centrochromatin. ChIP-sequencing and chromatin fiber staining techniques have been applied to investigate the distribution of H3 and CENP-A nucleosomes at the centromere. In the chromatin fiber analysis, centromeric chromatin is physically stretched to 50-100 times of its normal interphase length after cell lysis providing a resolution of approximately 20 kb. Using this chromatin unfolding technique, CENP-A molecules and H3 nucleosomes were found to be interspersed at the centromere in flies and human (Blower et al., 2002; Sullivan and Karpen, 2004). However, local CENP-A clusters with higher occupancy have been observed as well (Blower et al., 2002). CENP-A clustering may be even more pronounced in the 3-D architecture of the centromere as indicated by high resolution microscopy (Andronov et al., 2019).

ChIP-Seq is a powerful tool to study epigenetic modification genome-wide but currently does not provide single cell resolution like the chromatin fiber technique. Moreover, studying CENP-A distribution at native centromeres using ChIP-Seq is limited by the repetitive nature of the centromere. Sequencing reads cannot unambiguously be mapped to centromeric DNA in the way it would be required to determine a CENP-A nucleosome sequence. However, ChIP-Seq on chromosomes that are not characterized by highly repetitive DNA sequences (such as the Z chromosome of chicken or human neocentromeres (**see section 2.4** and **2.6**)), where sequence mapping is feasible have been used to study CENP-A nucleosome occupancy at the centromere (Bodor et al., 2014; Shang et al., 2013, 2010). These studies revealed that, in agreement with the chromatin fiber analysis, CENP-A nucleosomes are indeed not found in one major cluster but smaller sets of local clusters with very high CENP-A occupancy seeming to exist within the centromere. Recent advances in the chromatin fiber technique and ChIP approaches describe the presence of two distinct CENP-A populations at the centromere. Only one population appears to be strongly associated with CENP-C and is likely directly involved in kinetochore formation (Kyriacou and Heun, 2018; Melters et al., 2019) while the second population might have a different function in CENP-A homeostasis (Melters et al., 2019). This is in agreement with the estimated average molecule number of CENP-C (215 molecules) and CENP-A (400 molecules) (Bodor et al., 2014; Suzuki et al., 2015).

Moreover, based on native ChIP-Seq data nucleosomes at the centromere were found to be phased on human centromeric DNA. Centromeric DNA is highly repetitive and one characteristic is the

recurrent presence of a 17-bp sequence motif called the CENP-B binding box, to which CENP-B protein sequence-specific binds (**see section 2.5**). Nucleosome phasing occurs independently of CENP-B along centromeric DNA but the presence of CENP-B seems to generally improve phasing of canonical and especially of CENP-A nucleosomes (Ando et al., 2002; Hasson et al., 2013; Nechemia-Arbely et al., 2017). CENP-B might additionally support partial unwrapping of CENP-A nucleosomes, a feature that could generally support a centromere/kinetochore formation (Hasson et al., 2013).

The presence of a CENP-T/-W-/S-/X nucleosome-like structure further contributes to the centromere architecture. In contrast to the CENP-A nucleosome, this structure induces not negative, but positive supercoils in the DNA (Takeuchi et al., 2013).

Electron microscopy of metaphase chromosomes revealed that CENP-A occupies a compact domain at inner centromere regions but is absent from the most inner regions where major fractions of CENP-B reside (Owen J Marshall et al., 2008). This indicates that many copies of CENP-A nucleosomes are clustered in 3-D space even stronger from what is known from the nucleosome sequence perhaps especially during mitosis. Interestingly, a recent study suggests a local orientation of CENP-A nucleosomes during mitosis mediated by asymmetric binding of CENP-C and CENP-N molecules. As CENP-A nucleosomes maybe interconnected this local asymmetry might be expanded on a global centromeric level (Allu et al., 2019).

The underlying centromeric DNA is further proposed to affect the architecture of the centromere. Intrinsic features of the highly repetitive DNA could promote structure formation (**see section 2.4**) (Aze et al., 2016; Kasinathan and Henikoff, 2018). In this regard, the binding of CENP-B to centromere DNA likely induces conformational changes of centromeric chromatin. The crystal structure of the CENP-B DNA binding domain bound to DNA revealed a ~60° bending of the DNA by CENP-B (Tanaka et al., 2001). Moreover, *in vitro* CENP-B was proposed to possess a bundling activity for two distant CENP-B boxes as the protein can dimerize which possibly could promote local heterochromatin formation (Tawaramoto et al., 2003; Yoda et al., 1998, 1992) (**see section 2.5**).

II.3 The Never-ending Story of Centromere Identity - CENP-A's Epigenetic Self-Assembly Loop

In human and most species studied so far, centromeres can propagate through an epigenetic mechanism based on the epigenetic mark CENP-A (Zasadzińska and Foltz, 2017). New CENP-A chromatin assembly is ultimately directed by the presence of preexisting CENP-A nucleosomes which generates an epigenetic self-assembly loop (**Figure 3**).

Centromeric CENP-A nucleosomes are further described to be remarkably stable (**see section 2.3d**) and CENP-A molecules are very efficiently recycled at the centromere when cells undergo S-phase (unlike other H3 histones) (Bodor et al., 2013; Shelby et al., 2000; Zasadzińska et al., 2018).

DNA replication however, causes dilution of CENP-A nucleosomes to both daughter strands, which commonly occurs in a random fashion (Bodor et al., 2013; Shelby et al., 2000) with most likely some interesting exceptions discovered in the context of stem cell differentiation of the *Drosophila* mid-gut (Arco et al., 2018). CENP-A nucleosome replication does not occur immediately in S-phase but CENP-A is replenished in many species studied so far in a different cell cycle phase (Bernad et al., 2011; Black et al., 2007; Erhardt et al., 2008; Takayama et al., 2008). This is however different in budding yeast where CENP-A^{Cse4} is replenished during S-phase (Pearson et al., 2004; Wisniewski et al., 2014) and in fission yeast that can replenish CENP-A^{Cnp1} in S-phase and G2-phase (Takayama et al., 2008).

In human, the timing of new CENP-A reloading was uncovered by taking advantage of the SNAP-tag technology (Jansen et al., 2007a). At a time point zero preexisting SNAP-tagged proteins in the cell can be chemically quenched using a non-fluorescent benzylguanine derivative which rapidly and irreversibly binds to the SNAP-tag. Following a chase period, newly synthesized SNAP-tagged proteins can be labeled using a different, this time fluorescent benzylguanine derivative allowing exclusive detection of newly synthesized proteins by fluorescence microscopy. Using this technique Jansen et al. revealed that CENP-A^{SNAP} is newly deposited at the centromere in late telo-/early G1 phase (Jansen et al., 2007a) (**Figure 3A**).

CENP-A dilution during DNA replication creates vacant nucleosome positions which are proposed to be filled with nucleosomes containing the histone variant H3.3 which consequently acts as a placeholder for CENP-A during G2/M (Dunleavy et al., 2011) (**Figure 3A**). Interestingly, during DNA replication cells evict CENP-A from non-centromeric regions and retain only centromeric CENP-A, which demonstrates that DNA replication can act as an error correction mechanism to

restrict CENP-A accumulation to the centromere (Nechemia-Arbely et al., 2019). This provides a theory to explain why new CENP-A deposition is uncoupled from DNA replication since concomitant new CENP-A assembly in S-phase could likely continuously create new mis-incorporations along chromosome arms. Centromeric CENP-A retention was proposed to be linked to the CCAN (Nechemia-Arbely et al., 2019) but the molecular details of this error correction pathway and what prevents the assembly of CCAN proteins involved in this error correction system at non-centromeric CENP-A prior to S-phase is still unclear.

The major players involved in new CENP-A deposition are: CENP-A's chaperone HJURP, the reloading licensing complex Mis18 and the CCAN as the reader of preexisting CENP-A (**Figure 3B**).

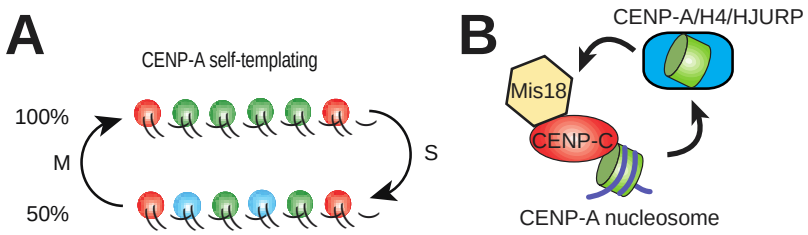


Figure 3: The CENP-A epigenetic self-assembly loop of the centromere.

*(**A**) Preexisting CENP-A nucleosomes (illustrated in green) are diluted equal during S-phase to both daughter strands resulting in half maximal centromeric CENP-A level after DNA replication. H3.3 nucleosomes (blue) are incorporated as CENP-A placeholder. After M-phase preexisting CENP-A molecules direct the deposition of new CENP-A molecules at the centromere to restore CENP-A level at the centromere.*

*(**B**) Major players of the CENP-A self-assembly loop. CENP-C acts as the epigenetic reader of preexisting CENP-A nucleosomes and directs the Mis18 complex to the centromere which in turn licenses new CENP-A incorporation mediated by the chaperone HJURP.*

a) HJURP – CENP-A's chaperone

Like other histones, CENP-A's deposition is mediated by a specific chaperone which in the case of CENP-A is the Holiday Junction Recognition Protein (HJURP) (Dunleavy et al., 2009; Foltz et al., 2009). At the native centromere, HJURP was proposed to have two functions during the course of the cell cycle. During telo-/early G1 phase HJURP mediates new CENP-A deposition at the

centromere and during S-phase HJURP was described to retain CENP-A at the centromere to avoid its loss specifically from the centromere (Nechemia-Arbely et al., 2019; Zasadzińska et al., 2018). In the latter, more recently discovered mechanism, the authors propose that HJURP, preexisting CENP-A and the replicative helicase complex MCM2-7 interact to ensure CENP-A retention through S-phase (Zasadzińska et al., 2018).

Pre-nucleosomal CENP-A/H4 dimers bind to an amino-terminal HJURP domain which is highly conserved across species including budding yeast where HJURP is called Scm3 (Suppressor of chromosome mis-segregation). Budding yeast Scm3 is much shorter than human HJURP and consists mainly of the CENP-A/H4 binding domain (Camahort et al., 2007; Mizuguchi et al., 2007; Pidoux et al., 2009; Stoler et al., 2007; Williams et al., 2009). For this reason, the domain is also called Scm3 domain in human. The Scm3 domain importantly recognizes the CATD domain of CENP-A (Shuaib et al., 2010) (**see section 2.2a**) (**Figure 5**). For CENP-A/H4 deposition the N-terminal tail of H4 (residues K5 and K12) must be acetylated, which is mediated by a complex of RbAp46/48 and histone acetyltransferase 1 (HAT1) (Shang et al., 2016). Binding of CENP-A/H4 to HJURP was further shown to increase the stability of soluble CENP-A (Bassett et al., 2012).

Targeting HJURP to an ectopic location is sufficient for CENP-A assembly at this locus in ~60% to ~90% of the cases (Barnhart et al., 2011; Shono et al., 2015) or to trigger centromere formation on human artificial chromosomes (Logsdon et al., 2019). Although the process of histone H3.3 eviction and CENP-A nucleosome assembly through HJURP is not well understood much progress has been made in understanding key steps of the mechanism and of its regulation.

Based on the crystal structure the association of HJURP with CENP-A/H4 was found to prevent tetramerization of CENP-A/H4 (Hu et al., 2011). Assembly of a homotypic CENP-A nucleosome somehow requires an interplay of two CENP-A/H4/HJRUP complexes. Dimerization of HJURP was proposed to be important in this context (Zasadzińska et al., 2013). This model however was more recently contradicted by including interaction studies of HJURP with the Mis18 complex, a protein complex upstream of HJURP and important for HJURP recruitment to the centromere at the end of mitosis (Barnhart et al., 2011; Pan et al., 2019). The authors of this study instead propose either a different coordination of two CENP-A/H4/HJURP recruited individually by two adjacent Mis18 complexes or a sequential deposition (Pan et al., 2019). In addition to the Scm3 domain, there are two HJURP Carboxy Terminal Domains (HCTD1 and HCTD2) which are sufficient for HJURP localization to the centromere in telo-/early G1 phase by interacting with the Mis18 complex (Foltz

et al., 2009; Nardi et al., 2016; Pan et al., 2019, 2017; Wang et al., 2014; Zasadzińska et al., 2013) and a Central Domain (CD) which binds to DNA, is important for CENP-A deposition (Müller et al., 2014) and interacts with cyclin A (Stankovic et al., 2017) (**Figure 5**). The latter has important regulatory function for CENP-A reloading (see below).

Once recruited, HJURP is proposed to be locally confined at the centromere through the formation of nuclear actin triggered by Diaphanous formin (mDia2). This HJURP confinement was proposed to be important for productive new CENP-A incorporation although the molecular details remain less clear (Liu et al., 2018; Liu and Mao, 2016). Moreover, condensin II is recruited in a HJURP dependent manner to the centromere which is reported to be important for centromere decondensation and new CENP-A deposition (Barnhart-Dailey et al., 2017).

CENP-A assembly is restricted to late telophase/early G1 phase which is mediated by the Cyclin Dependent Kinases 1 and 2 (CDK1/2). CDK1/2 are moderately active in S/G2 and are highly active during M-phase. Inhibition of CDK1/2 before M-phase was shown to be sufficient to promote precocious CENP-A reloading at the centromere in G2 phase (Silva et al., 2011). CDK1/2 have many substrates among them HJURP and the Mis18 complex (Müller et al., 2014; Stankovic et al., 2017) (**Figure 4**).

HJURP phosphorylation sequesters the protein away from centromeres and mutation of three phospho-residues (S210/S211/S412) in HJURP$^{\Delta HCTD2}$ to alanines was found to be partially sufficient to enable CENP-A reloading in G2 phase (Stankovic et al., 2017). Important for CDK1/2 phosphorylation seems to be the interaction of HJURP with cyclin A/CDK (high levels in S/G2) and cyclin B/CDK (high levels in M-phase) (**Figure 4**) (Stankovic et al., 2017). Recently, Pan et al. (2019) identified that the HCTD 1 and 2 domains of HJURP directly interact with the Mis18 complex. S412 is located inside HCTD1 which could prevent Mis18 complex interaction (Pan et al., 2019; Stankovic et al., 2017) (**Figure 5**). It is tempting to speculate that S595 (resides inside HCTD2 and is likely phosphorylated during S/G2 (Stankovic et al., 2017)) negatively impacts interaction with Mis18 in the HCTD2 in a similar way as the S412 site. However, mutations of S412A and S595A are insufficient to trigger precocious CENP-A reloading emphasizing the importance of S210/211 phospho-residues, which interestingly are in very close proximity of a HJURP-DNA interaction domain (285-321aa) (Müller et al., 2014).

HJURP was also suggested to be important for CENP-C assembly at the centromere (Tachiwana et al., 2015) and recently for CENP-T assembly (Ding et al., 2019). The molecular mechanism however of both events remain poorly understood. CENP-T/-W was additionally also described to interact

with the H2A-H2B chaperone FACT (Prendergast et al., 2016), suggesting that this chaperone may also be involved in the assembly.

b) Mis18 complex - the licensing factor

The Mis18 complex forms an octameric complex of two M18BP1 (Mis18 Binding Protein 1), four Mis18α and two Mis18β subunits (Fujita et al., 2007; Pan et al., 2017; Spiller et al., 2017). The Mis18 complex is required for HJURP reloading but Mis18 does not require HJURP for its recruitment demonstrating that Mis18 is upstream of HJURP for new CENP-A deposition (Barnhart et al., 2011). Similar to HJURP the Mis18 complex is not constitutively present at the centromere but only in anaphase until early G1 (Fujita et al., 2007; Hayashi et al., 2004) and phosphorylation not only regulates its centromere recruitment but also the assembly of the Mis18 subunits in human and fission yeast (McKinley and Cheeseman, 2014; Nardi et al., 2016; Pan et al., 2017; Stankovic et al., 2017; Subramanian et al., 2016). The Mis18 complex is highly conserved and essential for CENP-A assembly (Fujita et al., 2007; Hayashi et al., 2004; Maddox et al., 2007) but prominently not identified in *Drosophila*, suggesting the evolution of alternative CENP-A reloading pathways (Zasadzińska and Foltz, 2017).

The N-terminal region of M18BP1 (2-130 aa) interacts with Mis18α/β complex involving the so-called YIPPEE domain found in a middle region of Mis18α (78-191 aa) and β (73-189 aa) (Pan et al., 2017; Spiller et al., 2017). CDK1 phosphorylation of the residues T40 and S110 of M18BP1 reduces binding affinity of Mis18α/β and is described as an important negative regulator for Mis18 complex assembly (Pan et al., 2017; Spiller et al., 2017). Recruitment of the Mis18 complex to the centromere is controlled by a single phospho-residue on M18BP1 (T653) (Stankovic et al., 2017) (**Figure 4**). The HJURP HCTD domains interact with three helix bundles derived from the C-terminus of two Mis18α and one Mis18β units (Pan et al., 2019; Subramanian et al., 2016; Wang et al., 2014). An interaction between M18BP1 and HJURP is not described suggesting that M18BP1 perhaps connects Mis18α/β and in turn indirectly HJURP to the centromere. The first 140 amino acids of M18BP1 are sufficient for the interaction with Mis18α/β while most likely the C-terminus of M18BP1 interacts with centromere components (Dambacher et al., 2012; Pan et al., 2017). Mis18β is also reported to directly interact with the C-terminus of CENP-C$^{694\text{-}943}$. This interaction is either important for Mis18β recruitment (Stellfox et al., 2016) or to stabilize the Mis18 complex once recruited as an octameric complex can be reconstituted *in vitro* even without CENP-C (Pan et al., 2017) (**Figure 5**).

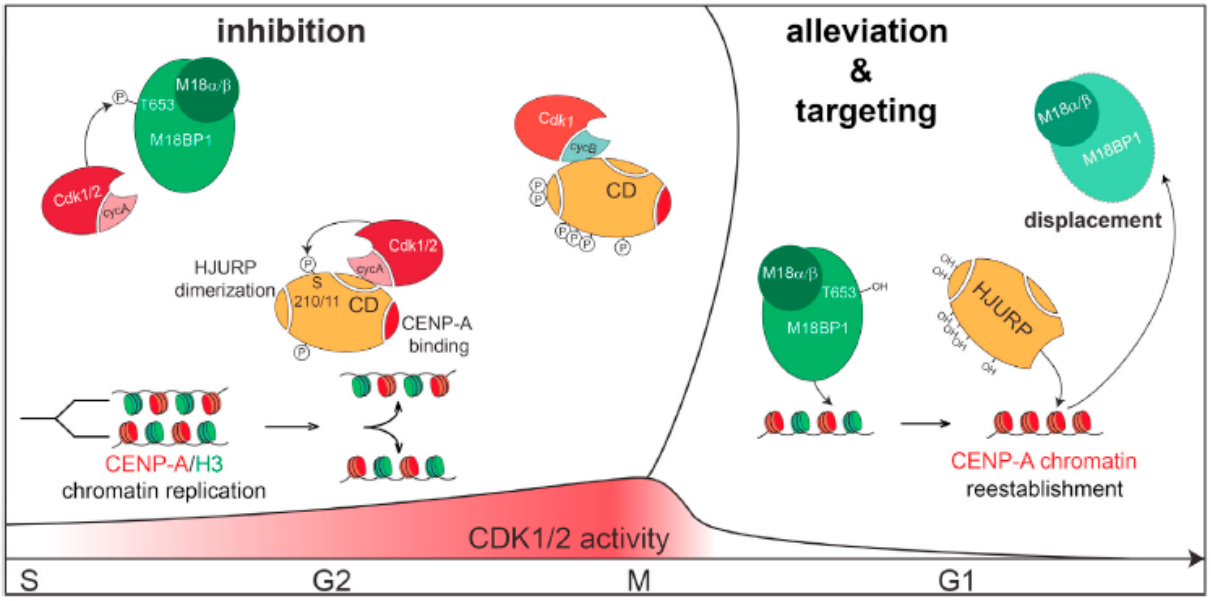


Figure 4: Model for CDK1/2 mediated cell cycle regulation of CENP-A reloading by Stankovic et al. 2017.

CDK1/2 prevent CENP-A reloading during S/G2 to M-phase by interacting and phosphorylating HJURP and M18BP1. After M-phase CDK2 level drop allowing new CENP-A deposition. Complete CENP-A reloading further requires Mis18 complex displacement from the centromere.

In certain aspects similarly to CENP-C (140 kD), M18BP1 is a big (130 kD) highly unstructured protein that is implicated to bind several proteins such as CENP-C (Dambacher et al., 2012; Moree et al., 2011; Shono et al., 2015), CENP-I (Shono et al., 2015), Polo-like kinase 1 (Plk1) (McKinley and Cheeseman, 2014), MgcRacGAP (Lagana et al., 2010) and KAT7 (Ohzeki et al., 2016) (**Figure 5**). For most if not all of these interactions the molecular details are only poorly understood.

Plk1 phosphorylation of M18BP1 was – in contrast to CDK phosphorylation described above – found to be required for M18BP1 centromere localization and is thus necessary for CENP-A deposition at the centromere (McKinley and Cheeseman, 2014). This study also proposed another functional control of M18BP1 by Plk1 phosphorylation in addition to the regulation of M18BP1 recruitment. Hence, Plk1 phosphorylation adds another layer to the tight regulation process of CENP-A reloading although CDK inhibition alone was sufficient to trigger precocious CENP-A reloading (Silva et al., 2011), thereby suggesting a more ancillary control by Plk1.

M18BP1 was also described to interact with the histone acetyltransferase KAT7 (Ohzeki et al., 2016). This was proposed to be important to maintain a transcriptionally permissive centrochromatin

(**see section 2.2c**) and thus prevent centromere inactivation (Ohzeki et al., 2016). Using the tetR/tetO system, the C-terminus of M18BP1 (851-1132) was described to be the main driver of KAT7 recruitment (Ohzeki et al., 2016). In HeLa cells KAT7 depletion and simultaneous overexpression of the histone methyltransferase Suv39h1 results in centromere dysfunction caused by spreading of heterochromatin in the centromere region (Ohzeki et al., 2016). KAT7 association with M18BP1 is surprising as this suggests that KAT7 recruitment occurs concomitant with new CENP-A assembly. However, it remains unclear if the timing of KAT7 recruitment during the cell cycles is important.

In a functional proteomics study, the small GTPase activating protein (GAP) MgcRacGAP was further identified to interact with M18BP1 in human (Lagana et al., 2010). Together with its antagonist the GTP exchange factor Ect2, this small GAP especially regulates the activity of the small GTPase Cdc42. The authors propose that Cdc42 activity modulates newly incorporated CENP-A but the mechanism as well as the binding site of MgcRacGAP on M18BP1 are not identified (Lagana et al., 2010). CENP-A stabilization especially at the centromere could be another important mechanism to restrict the centromere to its unique position and enable CCAN assembly only at the stable centromere location.

Finally, M18BP1 is proposed to interact with a C-terminal region of CENP-C although also this interaction remains poorly described (Dambacher et al., 2012; Hori et al., 2012; Moree et al., 2011; Shono et al., 2015).

Since T653 phosphorylation prevents M18BP1 recruitment to the centromere (Stankovic et al., 2017), this residue is likely involved in the CENP-C interaction and is located within a ~50 amino acid conserved Myb-like domain which was identified to be important for CENP-A loading in *C. elegans* (Maddox et al., 2007). Interestingly, a long N-terminal construct (1-490aa) lacking the T653 residue can be targeted to the centromere (McKinley and Cheeseman, 2014). Thus, it is tantalizing to speculate that M18BP1 may be able to connect to the CCAN through alternative interfaces. A possibility for this could be an indirect connection between CENP-C and M18BP1 complex through CENP-I (Shono et al., 2015) but future studies are required to validate the existence of such a connection and elucidate the molecular details.

In chicken and frogs, M18BP1 was found to directly connect to CENP-A nucleosomes, which therefore presents another alternative pathway for CENP-A reloading (French et al., 2017; Hori et al., 2017). However, the residues of M18BP1 involved in this interaction are not conserved in human suggesting that human M18BP1 cannot directly interact with CENP-A nucleosomes but likely require CENP-C or CENP-I.

Tethering M18BP1 to an ectopic site was described to be insufficient to trigger CENP-A assembly at this locus in human (Ohzeki et al., 2019; Shono et al., 2015). This could suggest that in human, in addition to the presence of the Mis18 complex, HJURP also requires components of the CCAN like CENP-C and CENP-I to successfully assemble CENP-A (Dambacher et al., 2012; Hori et al., 2012; Okada et al., 2006; Shono et al., 2015).

c) The CCAN – the reader of the epigenetic mark

Although in frogs and chicken the CCAN may not be required for HJURP recruitment and CENP-A assembly, in human, CENP-C (and maybe CENP-I) seems to be very important to function as a "reader" of CENP-A nucleosomes (Dambacher et al., 2012; Hori et al., 2012; Moree et al., 2011; Shono et al., 2015; Stellfox et al., 2016).

As described above CENP-C interacts via its C-terminal domain with M18BP1 and Mis18β (Dambacher et al., 2012; Hori et al., 2012; Moree et al., 2011; Shono et al., 2015; Stellfox et al., 2016).

When bound to an ectopic site CENP-C and CENP-I are indeed individually sufficient to recruit the Mis18 complex and trigger the incorporation of endogenous CENP-A at this locus even though with a rather low efficiency (~30%) (Hori et al., 2012; Shono et al., 2015).

In *Drosophila*, the Mis18 complex is not identified (see Zasadzinska and Foltz 2017 for a detailed view on the conservation of the CENP-A self-assembly loop). Using human cells as a heterologous system in combination with a LacO array, three *D. melanogaster* components (CENP-A^{CID}, CENP-C and $HJURP^{Cal1}$) have been found to be sufficient for centromere propagation (Roure et al., 2019). *Drosophila* CENP-C directly interacts with HJURP; however, the poor sequence conservation of the proteins does not allow any conclusions for the human case. Interestingly, CENP-C fragments but not the full-length protein were found to interact with HJURP *in vivo* in human (Tachiwana et al., 2015). Although the authors concluded that HJURP may be important for CENP-C recruitment it seems possible that a transient interaction involving the Mis18 complex plays a more important role for new CENP-A reloading which could potentially also explain why the Mis18 complex alone is insufficient for CENP-A reloading (Shono et al., 2015).

In contrast to HJURP and the Mis18 complex, CENP-C is clearly constitutively present at the centromere. Its replenishment follows rapidly after CENP-A reloading in G1/S phase but is not

observed in G2/M (Hoffmann et al., 2016; Hoffmann and Fachinetti, 2018). It is tantalizing to speculate that new CENP-C reloading to the centromere is similarly regulated by CDK activity as is CENP-A reloading but little is known about the regulation mechanism of CENP-C reloading.

While CENP-I is connected to CENP-C via the REST domain (**Figure 5**), CENP-C in turn as described previously, binds to CENP-A nucleosome and is hence the direct "reader" of the epigenetic mark (**Figure 5**). CENP-C possesses two nucleosome binding sites called central region and CENP-C-motif. *In vitro* studies suggest that both domains exclusively bind CENP-A nucleosomes and the central region has higher affinity for CENP-A (Ali-Ahmad et al., 2019). The central region (especially residue R522) *in vivo* was found to be most important for CENP-A stabilization while the CENP-C motif for this matter was described to be dispensable (Guo et al., 2017). CDK1 mediated phosphorylation of CENP-C was further described to strengthen CENP-A/C interaction presenting an additional regulation of CENP-C stabilization at the centromere (Watanabe et al., 2019).

CENP-C binds to CENP-A nucleosomes but also connects to chromatin in other ways: indirectly via CENP-I/H/K/M/L/N and CENP-T complex (**see section 2.1a**) and CENP-B and directly through an interaction with DNA (Fachinetti et al., 2015; Hoffmann et al., 2016; Politi et al., 2002; Suzuki et al., 2004) (**see section 2.5**). Whether these connections substantially contribute to centromere identity is not well understood. Tethering CENP-T and CENP-B to an ectopic site in some studies however, were found to be insufficient for ectopic CENP-A assembly (Gascoigne et al., 2011; Hori et al., 2012; Okada et al., 2007) (**see section 2.5**).

Moreover CENP-C has a dimerization domain at the C-terminus (Carroll et al., 2010; Cohen et al., 2008). The role of CENP-C dimerization for centromere organization and potential regulations (e.g. by CENP-B interaction, binding of other proteins or PTMs) is however only poorly understood.

d) CENP-A – the epigenetic mark

The presence of CENP-A is crucial for long term viability of proliferating cells and CENP-A is required for centromere propagation for indefinite cell cycles (Fachinetti et al., 2013; Howman et al., 2000; Régnier et al., 2005).

As mentioned earlier, the CATD domain of CENP-A is sufficient for CENP-A targeting to the centromere at telo-/early G1 phase in human (Black et al., 2007, 2004).

This provides the first step in a two-step cycle by which CENP-A acts as an epigenetic mark to self-direct its own chromatin assembly. Once assembled into chromatin, in the second step, CENP-A nucleates the assembly of the CCAN via its N- and C-terminal tails (Fachinetti et al., 2013; Kato et al., 2013; Logsdon et al., 2015).

CENP-A deposition at native centromeres occurs in nucleosomes that are in close proximity of preexisting CENP-A nucleosomes (Ross et al., 2016). The density of CENP-A may be important for *de novo* CENP-A deposition as CENP-A is highly enriched at native centromeres and even form clusters within the centromere (**see section 2.2e**) (Andronov et al., 2019; Bodor et al., 2014). An open question in this regard is whether the presence of preexisting CENP-A nucleosomes is at all required to template new CENP-A deposition (**see section 2.7**). As mentioned previously tethering experiments could demonstrate that CENP-A assembly at an ectopic locus does not require already existing CENP-A nucleosomes as long as specific components of the CENP-A assembly pathway were present (e.g. (Barnhart et al., 2011; Hori et al., 2012; Shono et al., 2015)). However, in these assays protein attachments and concentrations are tremendously different in comparison to native centromeres. This therefore makes any conclusion on *de novo* CENP-A deposition in absence of preexisting CENP-A at native centromeres uncertain.

Furthermore, it is unclear if CENP-A has any guiding function to target the replacement of specific neighboring H3 nucleosomes at native centromeres (Ross et al., 2016). A di-nucleosome model has been proposed where a CENP-A nucleosome is connected through CENP-C with a neighboring H3 nucleosome (Musacchio and Desai, 2017; Pan et al., 2019). To what extent the existence of a hetero-di-nucleosome is important for new CENP-A deposition and if the CENP-A deposition machinery perhaps preferentially replaces H3 nucleosomes that are associated with CENP-C and indirectly or directly with CENP-A is unclear. Overall, the contribution of preexisting (perhaps even post-translationally modified) CENP-A to new CENP-A reloading in particular needs to be tested.

While the core of the CENP-A epigenetic loop seems to mainly involve CENP-A, CENP-C, Mis18 and HJURP (**Figure 3 and 5**) other factors as indicated previously were described to be important regulators or stabilizers of CENP-C or other centromere components. Indeed, CENP-A incorporation was also proposed to require some level of transcription (Bobkov et al., 2018) (**section 2.2**).

Centrochromatin transcription also presents a risk for CENP-A loss since chromatin is disassembled similar to DNA replication. HAC formation for example was not detected when centromere DNA was flanked by actively transcribed regions (Nakashima et al., 2005). As mentioned earlier, HJURP was indicated to be involved in CENP-A recycling during S-phase (Zasadzińska et al., 2018). A recent study suggests that the histone chaperone and transcription elongation factor Spt6 provides a similar function for CENP-A retention during transcription (Bobkov et al., 2019). Spt6 interacts with both H3 and CENP-A nucleosomes but shows higher affinity for un-phosphorylated CENP-A *in vitro*, which could explain recycling of parental CENP-A and replacement of H3.3 in G1 (Bobkov et al., 2019). In addition to recycling, CENP-A was also proposed to be rejuvenated with a very slow turnover in temporary quiescent cells (Swartz et al., 2019). Fully differentiated, non-cycling cells like, for example, muscle cells however do not display any centromeric CENP-A accumulation any longer (Swartz et al., 2019).

In addition to CENP-A recycling mechanisms other factor contribute to the stability and longevity of CENP-A that is observed in different model systems (Arco et al., 2018; Bodor et al., 2013; Fachinetti et al., 2013; Smoak et al., 2016).

In human, a crucial stabilizer for both CENP-A and CENP-C is the protein CENP-B, which is the only centromere protein that binds to a specific DNA sequence motif of the centromere (Fachinetti et al., 2015; Fujita et al., 2015; Hoffmann et al., 2016; Masumoto et al., 1989; Suzuki et al., 2004) (**see section 2.5**).

Moreover, binding of CENP-C and CENP-N to CENP-A nucleosomes was reported to increase the stability of CENP-A nucleosomes (Cao et al., 2018; Falk et al., 2016, 2015; Guo et al., 2017). Hence factors that regulate CCAN stability in turn also may contribute to CENP-A stability. The sumo-protease SENP6 was recently identified as a major regulator of CCAN stability. Its depletion leads to hyper-sumoylation and subsequent degradation of the CCAN proteins CENP-C and CENP-I (Liebelt et al., 2019; Mitra et al., 2020). It is tempting to speculate that this regulation may further be involved in the prevention of neocentromere formation by restricting SENP6 to centromere regions.

Centromeric RNA transcripts were described to be important for CENP-A and CCAN stabilization in different species (Bergmann et al., 2011; Carone et al., 2013; McNulty et al., 2017; Rošić et al., 2014). A recent single cell study using single molecule FISH suggests that, *in vivo*, transcripts however do not co-localize with centromeres, further promoting debate around the function of the centromere transcripts (Bury et al., 2020).

CENP-A's stability is an important factor for its inheritance and function as the epigenetic mark of the centromere. In bull sperm, CENP-A even escapes protamine transition that removes all non-centromeric histones; further supporting the notion that CENP-A acts as a heritable mark (Palmer et al., 1990). During meiosis in mice, CENP-A was proposed to be loaded in G1 phase before meiosis I and last throughout both meiotic divisions with no substantial intermediate replenishment events. This is in particularly remarkable due to the long-lasting prophase I arrest in female meiosis (Smoak et al., 2016); although recently, slow gradual CENP-A renewal was also identified in prophase I-arrested starfish oocytes via a Plk1 independent, transcription-driven reloading mechanism (Swartz et al., 2019). A transgenerational role of centromere inheritance seems to be conserved and was further described in flies (Raychaudhuri et al., 2012).

Another question arises as to how CENP-A deposition is terminated in G1 phase? It has been proposed that the Mis18 complex or parts of it are actively dissociated from the centromere and Mis18 removal is even a necessity for productive CENP-A reloading (Nardi et al., 2016; Stankovic et al., 2017) (**Figure 4**). In fact, the Mis18β subunit was described to be ubiquitinated and degraded after completion of CENP-A reloading (Kim et al., 2014). An alternative model suggests that new CENP-A deposition is limited to the availability of paired CENP-A-H3 di-nucleosomes which are connected through CENP-C (and perhaps other CCAN proteins). Once all H3 nucleosomes of all di-nucleosomes are replaced, new CENP-A deposition is terminated (Musacchio and Desai, 2017). Clear evidence for active Mis18 complex disassembly or the di-nucleosome model however are missing. In chicken and frogs, CENP-C, which in the di-nucleosome model connects CENP-A nucleosomes and H3 nucleosomes, is not required for new CENP-A reloading (French et al., 2017; Hori et al., 2017); further supporting a Mis18 disassembly scenario.

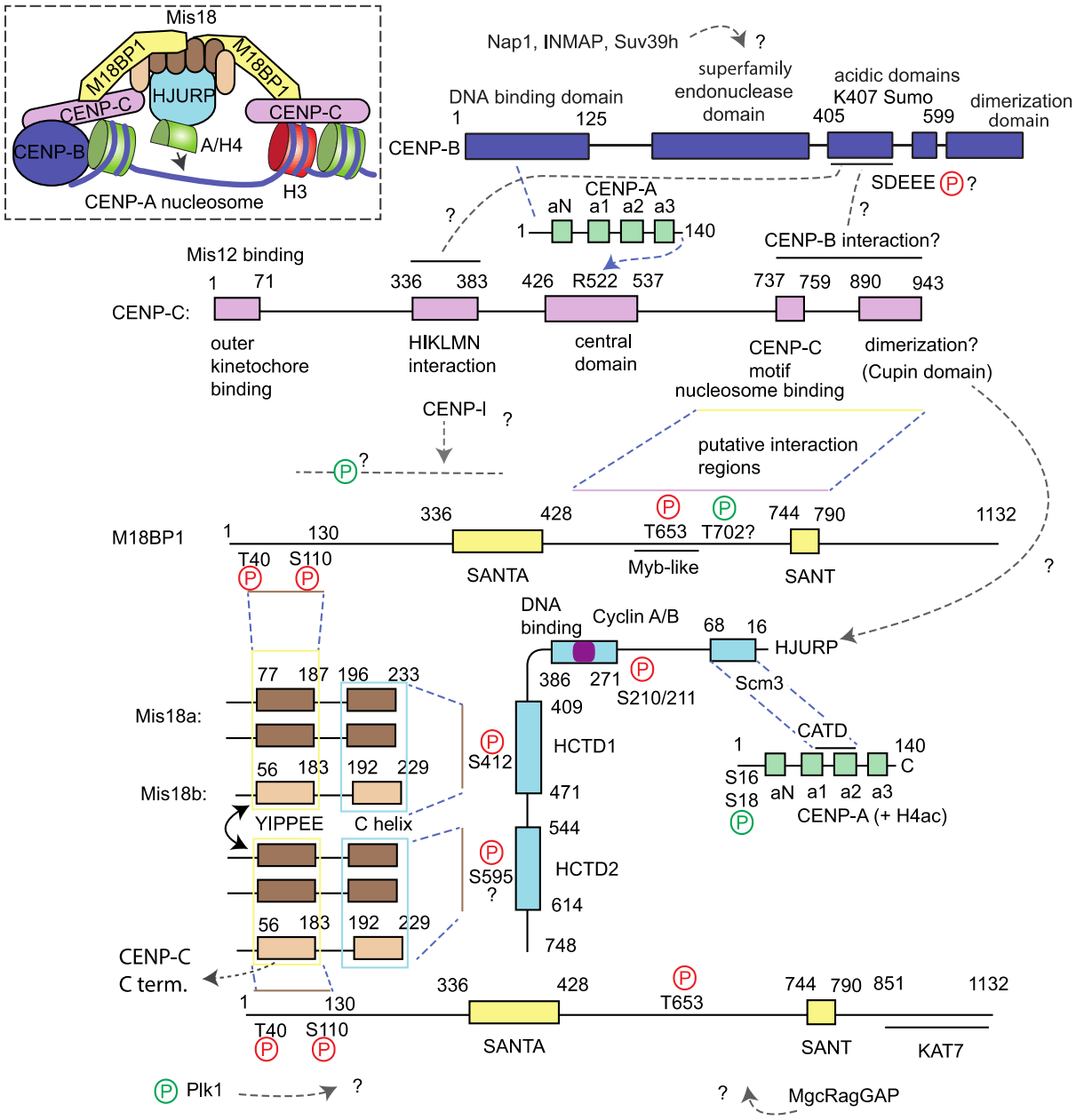


Figure 5: Protein interactions to direct new CENP-A assembly at the centromere.

Schematic map of proposed/reported protein interactions of HJURP, CENP-C, CENP-A, M18BP1, Mis18α and Mis18β as well as CENP-B. Top left inset: A model for new CENP-A deposition based on an interplay of two CENP-C molecules to assemble a new CENP-A nucleosome at the centromere during early G1 phase.

II.4 The Curious Case of Centromeric DNA (CenDNA)

Budding yeast possess so-called point centromeres that connect each sister chromatid to a single microtubule during mitosis (Pluta et al., 1995). The location of these point centromeres in the genome are specified by a ~125 bp sequence comprising three conserved DNA elements called Centromere DNA Element (CDE) I/II/III. Specific single point mutations in CDEIII lead to centromere dysfunction and chromosome segregation errors (McGrew et al., 1986; Ng and Carbon, 1987). However, the budding yeast case appears to be exceptional. In other monocentric species, the centromere spans along significantly longer DNA sequences (regional centromeres). The functional relevancies of the CenDNA in these cases are far less obvious and will be discussed in this section.

a) Centromere Evolution

CenDNA and CenDNA-binding proteins undergo rapid evolutionary changes which contradict the conserved function of the centromere and is referred to as "the centromere paradox" (Henikoff et al., 2001). An explanation of the centromere paradox may rely in the co-evolution of CenDNA with centromere components as the result of a meiotic drive. The meiotic drive theory suggests that during asymmetric female meiosis I bivalents undergo a selection process based on centromere features. Homologues with inferior centromere features are proposed to be more likely segregated into the polar body rather than into the egg and are thus less likely inherited in the next generation (Henikoff et al., 2001).

In mice, biased distribution of bivalents was observed in female oocytes obtained from hybrid mice strains with different centromere features (Akera et al., 2019, 2017; Chmátal et al., 2014; Iwata-Otsubo et al., 2017). These studies suggest that larger centromeres form more unstable spindle interactions, especially when oriented towards the cortical side of the spindle compared to their shorter homologue counterparts. This greater instability at the cortical side causes reorientation of the larger centromere towards the egg pole. The authors propose that spindle-kinetochore detachment and reorientation is a consequence of elevated Bub1 (Budding Uninhibited by Benzimidazoles) kinase activity found at larger centromeres. Bub1 activity can activate MCAK (Mitotic Centromere Associated Kinesin), a component of the error correction pathway (**see section 1.4**). MCAK preferentially destabilizes spindle fibers that are tyrosinated. This spindle fiber modification is more pronounced at the cortical side, which could explain the observed instability of the spindle to

kinetochore attachment during female meiosis I (Akera et al., 2019, 2017). Whether this proposed mechanism is the sole, universal centromere selection mechanism or whether there are other molecular ways of centromere selection during female meiosis I requires further investigation in different model systems.

In addition to centromere driven meiotic drive mechanism, non-centromeric "selfish" elements found on chromosome arms are found to foster biased bivalent segregation as well. In maize, a kinesin 14 motor protein can bind to heterochromatic knobs present on chromosome arms and impact chromosome segregation (Dawe et al., 2018).

A biased selection process during asymmetric female meiosis I may cause greater incompatibilities during symmetric male meiosis I, where for example, a strong maternal centromere needs to pair with a potentially drastically weaker, paternally-derived centromere. This in turn could lead to increased chances of segregation errors and hence infertility (Henikoff et al., 2001). Indications of other negative effects during mitosis or male meiosis as the result of a biased female meiosis I selection process have been reported in different species: in cells derived from the Muntjac deer for example, very long centromere sequences were found to encounter microtubule capture errors more frequently (Drpic et al., 2018). In monkeyflower hybrids, alleles with very clear selective advantage during female meiosis I were identified, but a plant homozygous for these alleles produced less pollen (Fishman and Kelly, 2015; Gamba and Fachinetti, 2020). Therefore, centromere proteins and CenDNA likely experience a selective pressure counteracting biased female meiosis I. This constant arms race for centromere selection during female meiosis I and centromere compatibility outside of meiosis I could contribute to the lack of a perfect centromere consensus sequence which emerged over evolution (Henikoff et al., 2001).

The repetitiveness of CenDNA likely also confers to more frequent unequal crossing over events during meiosis and more frequent sequence conversion events which could additionally promote rapid evolution of CenDNA sequences (Henikoff et al., 2001). Repetitive DNA with high AT content as found at the centromere of many species is proposed to be more vulnerable to DNA breaks which likely has implication for centromere evolution. In certain cancers such as colorectal cancer, breaks and rearrangement involving (peri-)centromere regions are very common (40-60%) [for a deeper review on this issue see (Barra and Fachinetti, 2018)]. However, centromere proteins normally contribute to the protection of CenDNA integrity (Giunta and Funabiki, 2017) (Guinta, Hervé et al. submitted).

Moreover, it has been suggested that evolutionary new centromeres (ENC) may be deprived of repeat arrays and acquire CenDNA after their formation over time (Rocchi et al., 2012). New centromere acquisition events during evolution likely further contribute to poor sequence conservation during evolution. Potential ENCs can be found in different species (e.g. chicken (Shang et al., 2010) and horse (Wade et al., 2009)) where the native centromeres of some chromosomes indeed lack any repetitive DNA arrays [reviewed in (Giulotto et al., 2017)]. The existence and stable inheritance of these centromeres represents further indication that the requirement of centromere DNA sequences for centromere function can be bypassed.

In human, all native centromeres possess repetitive DNA sequence arrays. However, in rare cases, the formation of neocentromeres (NeoCen) has been discovered (Earnshaw and Migeon, 1985; Owen J. Marshall et al., 2008; Voullaire et al., 1993). NeoCen can be stably inherited to the next generation, demonstrating that also in human, a specific DNA sequence may not be absolutely essential to maintain centromere identity (Amor et al., 2004; Voullaire et al., 1993). Even pseudo-dicentric chromosome have been discovered. In these cases, the native centromere has been inactivated and a centromere is formed on a different chromosome location (Amor et al., 2004; Owen J. Marshall et al., 2008). The presence of the centromeric DNA at the original centromere is not sufficient to reactivate the centromere (Amor et al., 2004) (**see section 2.6a**).

b) Centromere DNA sequence organization

Despite evolutionary differences, a reoccurring sequence pattern of regional centromeres in fungi, plants and animals is repeat arrays that span over several kilo-bases (kb) to mega-bases (Mb) of DNA suggesting that such arrangements could provide favorable conditions for centromere formation and/or function (Gamba and Fachinetti, 2020; Willard et al., 1989).

Repeat arrays found at all human centromeres are built on a series of primate-specific DNA units called alpha-satellites (Aldrup-Macdonald and Sullivan, 2014; Waye and Willard, 1987). An accumulative sequence of alpha-satellite DNA monomers (2-16 monomers) form a so-called Higher Order Repeat (HOR, 1 to 3 kb length). HORs in turn are repeated multiple times at the centromere producing highly repetitive HOR arrays of ~200 kb to up to 5000 kb in human (**Figure 6**) [reviewed in (Aldrup-Macdonald and Sullivan, 2014)].

Human alpha-satellite monomers have a length of 171-bp (Waye and Willard, 1987). Similar hierarchical organization of monomer sequence blocks can also be found in other species although

the monomer sequences are highly divergent (Alkan et al., 2011). Interestingly however, monomers from other species show similar length (e.g. in fish (*Sparus aurata*): 186 bp, in insects (*Chironomus pallidivittatus*): 155 bp, in plants (*Arabidopsis* and maize): 180 bp or (rice): 168 bp) or are around twice as long (e.g. in pigs 340-bp) (Henikoff et al., 2001). It has been suggested that either this length might favor local non-B form DNA folding (Garavís et al., 2015a, 2015b) (**see section 2.6c**) or may be important for nucleosome positioning as it is reminiscent to the DNA length of a single nucleosome (Henikoff et al., 2001).

Alpha-satellite monomers in human show relatively high divergence (50-80% similarity) but within an HOR array individual HORs are commonly nearly identical, which greatly contribute to the repetitiveness of centromere sequences [reviewed in (Miga, 2017)]. It is possible to distinguish between arrays of HORs present at different chromosomes since the composition of alpha-satellites within their HOR is variable among chromosomes. However, some chromosomes share the same HORs (Miga, 2017).

Moreover, some chromosomes have centromeres with two or more distinguishable HOR arrays showing different CENP-A occupancy levels. Generally, one HOR array is considerably longer than the other HOR arrays present at the same chromosome and shows higher CENP-A occupancy level. This dominating HOR array is considered the functionally active array for centromere assembly but exceptions may exist (Aldrup-Macdonald and Sullivan, 2014). In the majority of the cases the same HOR array of two homologue chromosomes functions as the active HOR array but individuals can also be heterozygous meaning that different HOR arrays on the homologues show different CENP-A occupancy at the centromere (Maloney et al., 2012). Such an occurrence has been found at the centromere of chromosome 17 where interestingly, shifting of the active HOR array seems to correlate with DNA sequence polymorphism. HORs of the longer, more commonly active array were found to be shortened. The authors speculate that a loss of CENP-B boxes (**see section 2.5**) may be a determining factor for a switch of CENP-A occupancy but the causality of sequence polymorphism and array switch remain unclear (Maloney et al., 2012).

HOR arrays are usually separated from each other by highly divergent alpha-satellite monomers (Miga, 2017). These assemblies of transition monomers which are also found in the pericentromere, can span hundreds of kb adjacent to the HOR arrays, but are not observed to be involved in binding of centromere proteins and were found to be insufficient for HAC formation (in contrast to inactive HOR arrays) (Hayden et al., 2013; Maloney et al., 2012). Although Logsdon et al. (2019) observed centromere formation on a HAC deprived of any kind of centromeric DNA, a

certain higher rate of HAC formation with specific centromere DNA sequences seems still likely (**see section 2.6b**). Interestingly even in this study, when HACs were formed on non-centromeric DNA without initial centromere seeding, HACs had somehow acquired centromeric DNA sequences (Logsdon et al., 2019).

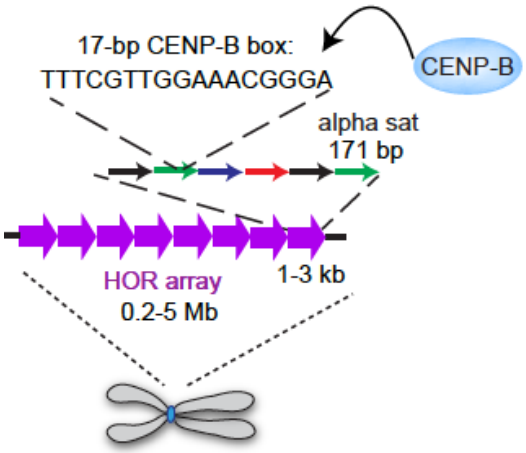


Figure 6: Schematic illustration of centromere sequence organization found at human centromeres. *Centromeric DNA is hierarchically organized based on tandem alpha-satellite monomers which can contain CENP-B boxes necessary for the binding of the only sequence specific DNA binding protein identified at human centromeres, CENP-B.*

c) Contribution of DNA sequence to the centromere architecture

The potential drive to acquire centromere repeat arrays in ENC (or even in HAC experiments) and especially their occurrence at all human centromeres supports a functional role of these DNA sequences.

Since centromeric DNA is commonly AT-rich and highly repetitive it has been suggested that these features could promote the formation of non-B-form DNA. Rather than a precise DNA sequence, the ability of any given sequence to induce a certain DNA topology favorable for centromere protein binding maybe the key selection factor for CenDNA (Dumont and Fachinetti, 2017).

An *in silico* approach further identified the reoccurrence of different dyad symmetries in small (<10bp) centromeric DNA sequence units (Kasinathan and Henikoff, 2018). These symmetries are not ubiquitously found in all species. Computational modelling suggest that sequences in which the dyad symmetry was present are more likely to establish an energetically stable non-B-form DNA structure (Kasinathan and Henikoff, 2018). Interestingly, species lacking the dyad-symmetry in their CenDNA (e.g. human with the exception of the CenDNA at the Y chromosome) are known to possess a sequence specific DNA binding protein like CENP-B (**section 2.5**). The authors suggest that in these cases CENP-B-like proteins might induce structure formation substituting the intrinsic propensity of CenDNA with dyad symmetry. Moreover, they hypothesize that HJURP might

preferentially bind to these non-B-form DNA configurations, further implicating a potential relevance for centromere identity (Kasinathan and Henikoff, 2018).

Experimentally it was shown that alpha-satellite monomers *in vitro* can form small non-B form conformations (e.g. so called i-motifs) using NMR (Gallego et al., 1997; Garavís et al., 2015a, 2015b). Whether this indeed occurs in longer DNA sequences and *in vivo* where additionally histones and other proteins are present, remains uncertain.

Larger topological structures of pieces of human centromeric DNA were observed by electron microscopy in a heterologous *Xenopus* egg extract system (Aze et al., 2016).

The presence of overwound DNA is dependent on topoisomerase I activity and condensins. In their model the authors suggest that smaller structures like hairpins might form during DNA replication as centromeric DNA is highly repetitive. The larger loops of dsDNA are proposed to prevent activation of DNA damage response in these difficult to replicate regions by sterically preventing hyper-loading of proteins of the DNA damage pathway (especially RPA, Replication Protein A). This, however, rather describes a mechanism of how cells control highly repetitive centromeric DNA during DNA replication, but does not necessarily allow any conclusion on the importance of such repetitive elements for centromere function per se.

A real functional implication of double stranded DNA loops at the centromere has been proposed during mitosis. These topological arrangements might act as molecular springs to absorb the microtubule pulling force exerted on centromeric chromatin (Bloom, 2014; Rosandić et al., 2008). However, experimental evidence for this is missing and how and if repetitive centromeric DNA engages topoisomerase I for loop formation requires further analysis.

Another intriguing finding in the study of Aze and colleagues was the assembly of CENP-A on the introduced centromeric DNA but not on non-centromeric control DNA in the egg extracts. This indicates an importance of centromeric DNA for CENP-A assembly. If structure formation or other features such as the high AT content, periodic presence of tandem repeat monomers, an accumulative presence of deoxyuridine bases (Shu et al., 2018) or other CenDNA features are important for this remains unclear. In *S. Pombe* centromeric DNA was also shown to destabilize H3 nucleosomes to favor CENP-A deposition, presenting another alternative mechanism for how centromere sequence could favor centromere formation/propagation (Shukla et al., 2018a).

Yet another theory to explain the functional importance of CenDNA involves transcription. The role of transcription for centromere function has been discussed previously (**see section 2.2d**). DNA structure formation or non B form DNA could regulate centromere transcription and slow down the process with potential implication for new CENP-A deposition at sites of RNA Pol II stalling for example (Kouzine et al., 2017). Moreover, transcripts of CenDNA themselves might show a special topology that could be favorable for its proposed functions, but neither experimental nor computational evidence for this exist to my knowledge.

The centromere transcript was also found to remain associated with DNA and form so called R-loops (Kabeche et al., 2018). These structures were proposed to be involved in the chromosome congression process by interacting with components important for Aurora B activity (Kabeche et al., 2018). Whether DNA sequence can promote R-loop formation remains unclear.

As indicated before, a very interesting feature of centromeric DNA in many species is the presence of a protein binding motif. In human, this binding motif is called CENP-B box and allows the binding of the only sequence specific DNA binding protein identified so far, named CENP-B (**Figure 6**) (Earnshaw et al., 1987). CENP-B may further contribute to the architecture of the centromere as discussed in the next chapter.

II.5 CENP-B

CENP-B is a 2 helix-turn-helix DNA binding protein that binds to a 17-bp DNA sequence motif called CENP-B box (Muro et al., 1992). CENP-B shows sequence similarity with class II DNA transposase of the pogo-like family (Gamba and Fachinetti, 2020). This led to the hypothesis that CENP-B could be a domesticated transposase. Indeed, residues that are crucial for transposase activity are mutated and sequence conservation is especially high in the DNA binding region (Kipling and Warburton, 1997).

CENP-B seems highly conserved in mammals suggesting that the exaptation event of CENP-B occurred in a common ancestor but CENP-B was not identified in most other vertebrates (Gamba and Fachinetti, 2020). Convergent domestications of transposases presumably also took place in other species such as fungi (Casola et al. 2008) and insects leading to CENP-B-like proteins (Gamba and Fachinetti, 2020).

Potential precursor CENP-B box sequences were found in evolutionary distant species (plants, insects, nematodes), which suggest that a similar sequence motif could have contributed to the domestication of different transposases in evolution (Gamba and Fachinetti, 2020).

Studies in closely related new world monkey species showed a tremendous variation of CENP-B box abundance and organization among these species (Thongchum et al., 2020). The authors concluded that CENP-B boxes are – similar to alpha-satellite sequences in general - acquired gradually over time at ENCs for example and not present at early events of centromere formation (Thongchum et al., 2020).

In human, CENP-B boxes are found within type I, but not type II alpha-satellite monomers (Miga, 2017). The active HOR array of a chromosome typically shows the strongest abundance of CENP-B boxes. Different chromosomes display variations of type I or II monomer abundance with CENP-B boxes being present in centromeres of all chromosomes except for the male sex chromosome, which possesses its own type of alpha-satellite sequences (Miga, 2017). As Y chromosomes do not undergo the selection process of asymmetric female meiosis I it has been speculated that this might account for the absence of CENP-B boxes at the Y centromere.

Precisely how CENP-B contributes to centromere strength in female meiosis I remains to a certain extent untested. As discussed below, CENP-B contributes to centromere stability, which could potentially also promote higher Bub1 kinase level and thus advantage in egg pole re-orientation during female meiosis I. Another theory suggests that CENP-B could have promoted centromere expansion by frequently inducing breakage-repair mechanism in centromeric regions. This is based on a transposase-derived nickase domain found in CENP-B that potentially is still active (experimentally untested) (Gamba and Fachinetti, 2020). A larger centromere region as discussed previously (**see section 2.4a**) may be an advantageous feature for selection in female meiosis I (Akera et al., 2019).

a) CENP-B's functions

Selfish gene behavior and acquisition of functional relevancies outside of female meiosis I do not necessarily exclude each other; but in the case of CENP-B could present two cooperative driving forces for positive selection. Several functional roles of CENP-B at human centromeres have been reported.

CENP-B presence at the centromere is proposed to affect centromere organization by optimizing nucleosome positioning along centromere DNA (Hasson et al., 2013). The positioning effect was also reconstituted *in vitro* using CENP-B, alpha-satellite DNA and core histone complexes, which were assembled with the aid of NAP-1 (Yoda et al., 1998). In addition, as mentioned in **section 2.2f** CENP-B is also predicted to affect the 3-D organization of centromeric DNA. Its presence may be important at chromosomes that have lost intrinsic sequence-specific features to promote structure formation at the centromere, which could be important for centromere function or identity, as suggested by an *in-silico* approach (Kasinathan and Henikoff, 2018).

CENP-B dimerization was proposed to be important to physically bring together two CENP-B boxes and promote heterochromatin formation or control transcription in certain regions of the centromere (**Figure 7**) (Tawaramoto et al., 2003; Yoda et al., 1998, 1992).

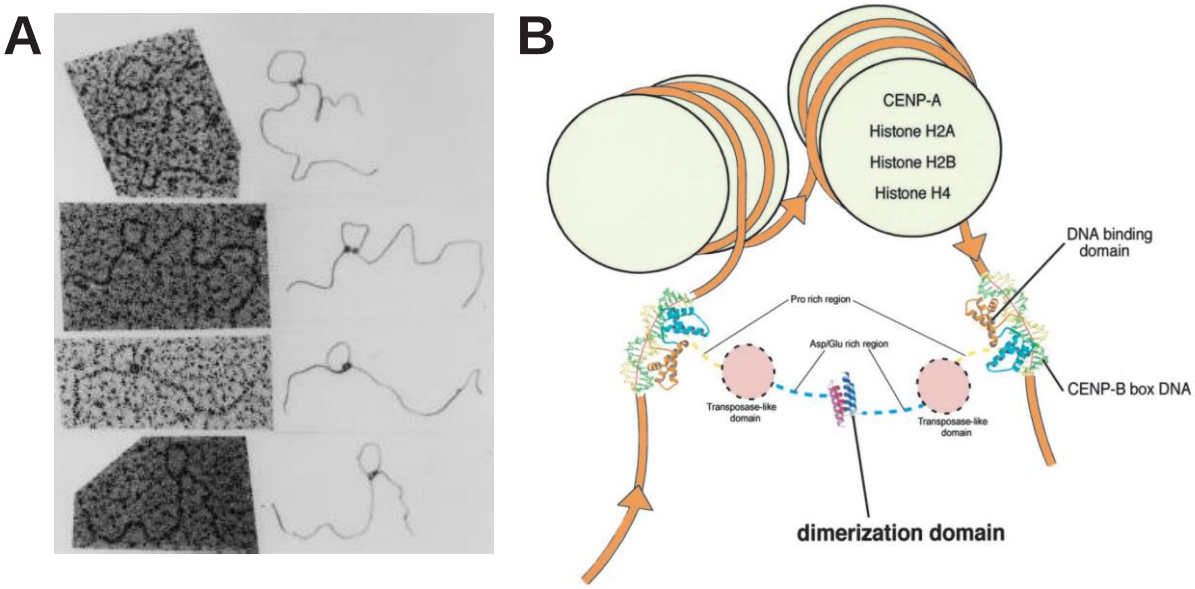


Figure 7: CENP-B dimerization could cause higher order structure formation.

*(**A**) Electron microscopy of CENP-B bound to DNA possessing two CENP-B binding sites assembled in vitro (Yoda et al. 1998).*

*(**B**) Model of nucleosome bundling by CENP-B from (Tawaramoto et al., 2003).*

In addition to CENP-B's potential effect on the architecture of the centromere, CENP-B was found to influence centrochromatin in different ways:

Two out of three potential CpG dinucleotides (common target sites for methylation) of type I alpha-satellite monomers reside within the CENP-B box motif. On one hand, CENP-B binding likely prevents methylation of alpha-satellite DNA but on the other hand methylation of these sequences were also shown to reduce CENP-B binding affinity to centromeric DNA, suggesting an interplay between DNA methylation and CENP-B binding regulation (Mitchell et al., 1996; Scelfo and Fachinetti, 2019; Tanaka et al., 2005). The regulation and physiological relevance of this for the centrochromatin state requires further exploration (Scelfo and Fachinetti, 2019). Moreover, expression of human CENP-B in CENP-B knock-out mouse embryonic fibroblast (MEF) cells increases H3K9me3 level at the mouse centromere, presumably through the interaction of CENP-B with a histone methyltransferase (Suv39h) (Okada et al., 2007). Future studies are however required to better elucidate this interaction and its regulation at native human centromeres.

Moreover, CENP-B depletion was found to regulate DAXX and H3.3 (deposited by DAXX) levels at centromeres (Morozov et al., 2017). The recruitment of DAXX was reported to be regulated by sumoylation and sumoylation-specific proteases (SENP). This study further reports changes in H3K9me3, ATRX and HP1α levels and hence also suggest an overall contribution of CENP-B to heterochromatin formation at the centromere (Morozov et al., 2017).

In conclusion, there are indications that CENP-B as a genetic feature could have an impact on the epigenetic centromere landscape. The functional relevance and mechanism behind this however remains only partly uncovered.

As indicated earlier, overwound CenDNA structures are suggested to be important for DNA replication (Aze et al., 2016). CENP-B was implicated to be involved in controlling DNA replication possibly through its chromatin regulatory functions (Erliandri et al., 2014); but how CENP-B and regulation of structure formation or resolution are related to DNA replication need further clarification. Also a CENP-B homologue in fission yeast was described to counteract replication-fork blocks which can be a result of Sap1 (Switch-activating protein 1) recruitment to long terminal repeat sequences present at fission yeast centromeres (Zaratiegui et al. 2011).

In addition to its implications in centrochromatin remodeling or centromere architecture, the abundance of CENP-B at the centromere also correlates with centromere strength in human during mitosis (Dumont et al., 2020; Fachinetti et al., 2015) (**Figure 8C**). The Y chromosome or chromosomes with neocentromeres for instance, have a higher probability to mis-segregate than other CENP-B possessing centromeres (Fachinetti et al., 2015) (**Figure 8B**). The Y chromosome is

also found to be frequently lost in hematopoietic cells in an age-related manner (Pierre and Hoagland, 1972). In 1 out of 5 men over 80, mosaic loss of chromosome Y in white blood cells has been identified, which in turn also reflects the ability of these cells to survive even without the Y chromosome (Forsberg et al., 2014). This further suggests an additional explanation for the lack of CENP-B boxes at the Y chromosome since its stable continuous transmission at least in certain contexts might be less essential as it is for other chromosomes. Although CENP-B might be involved in the regulation of the centrochromatin state, its positive impact on segregation fidelity most likely is conferred predominantly by its direct interaction with and stabilization of centromere components. In an evolutionary context, centromere stabilization by CENP-B could facilitate CENP-A/DNA evolution by permitting weaker interfaces.

Amino acids 10-25 were found to be essential for CENP-B's ability to bind to CENP-B boxes (Earnshaw et al., 1987; Okada et al., 2007; Yoda et al., 1992). Moreover, the N-terminal domain (1-44) was also proposed to interact with the N-terminal tail of CENP-A which substantially stabilizes CENP-A nucleosomes *in vitro* (Fujita et al., 2015). Furthermore, full-length and also N-terminal CENP-B constructs (1-159) were sufficient to trigger CENP-A incorporation when a bacterial artificial chromosome (BAC) comprising alphoid DNA with CENP-B boxes was transfected into mouse embryonic fibroblasts (MEF) (Okada et al., 2007). Altogether, this therefore leads to the conclusion that the N-terminus has a dual function for a) CENP-B to DNA binding and; b) CENP-A binding. Both functions together could indicate that CENP-B has a functional relevance to act as an identification mark of the centromere position perhaps upstream of the epigenetic mark CENP-A (Okada et al., 2007). Indeed, centromere formation was observed in various studies using HACs with integrated alphoid DNA which spanned a minimal length of 30 kb in MEF and human cell lines (Basu et al., 2005; Harrington et al., 1997; Nakashima et al., 2005; Okamoto et al., 2007; Tachiwana et al., 2013). Although this seems to support the idea that CENP-B facilitates centromere formation, alphoid DNA lacking CENP-B boxes were also effective at generating HACs (Harrington et al., 1997), but with strikingly reduced efficiency (Basu et al., 2005; Harrington et al., 1997). Moreover, to some extent puzzlingly, CENP-B was described to promote heterochromatin formation suggesting it to promote chromatin inaccessibility for centromere formation when alphoid DNA was integrated on a chromosome (Okada et al., 2007; Tachiwana et al., 2013). Another study also found HAC integration in the host genome. On these CENP-B box containing sequences, centromeres had formed occasionally (Nakashima et al., 2005); but it remains unclear if they formed before or after HAC sequence integration into the host genome. Recently, the team of Hiroshi Masumoto reported that CENP-B can recruit CENP-C to an ectopic centromere region and also CENP-A recruitment is

observed at least when overexpressed (Ohzeki et al., 2020; Otake et al., 2020). The authors propose that CENP-B has the ability to recruit chromatin remodelers with counteracting functionality but the regulation of this remains uncertain. CENP-B can recruit ASH1L, a H3K36 methylase, therefore promoting the formation of a permissive chromatin environment favorable for CENP-A incorporation but also heterochromatin components (Suv39h1 and HP1) (Otake et al., 2020).

Earlier studies also initially suggested that integration of human alphoid DNA containing CENP-B boxes into the genome of African Green Monkey (AGM) cells caused difficulties in chromosome segregation (Haaf et al., 1992). However, in hamster cells integration of human alphoid DNA also caused segregation defects; but no additional kinetochore formation was observed. Instead, integration of additional repeated DNA arrays (not necessarily alphoid DNA) caused DNA sister chromatid disjunction problems providing an explanation for the defects seen in the AGM cells (Warburton et al., 1997).

CENP-B interacts not only with CENP-A but also with CENP-C (Fachinetti et al., 2015; Hoffmann et al., 2016; Suzuki et al., 2004). Yeast-two-hybrid (Y2H) studies identified CENP-B's acidic domain (404-470) as a possible interaction region (**Figure 5**). Where exactly CENP-B binds on CENP-C is less clear as the Y2H identified three domains (336-383; 737-759; 890-943). Overexpression of a N-terminal CENP-B constructs lead to aberrant CENP-C organization (CENP-C no longer formed uniform round foci) especially during G2-phase and increased mitotic timing suggesting that CENP-B may be important for CENP-C organization at the centromere although a direct interaction of CENP-C with the N-terminus of CENP-B is not reported. A different pattern of aberrant CENP-C organization in G2 was also observed when a CENP-B construct including the acidic domain (1-561) was overexpressed. However, it remains poorly understood how CENP-B modulate the organization of CENP-C at the centromere. CENP-B depletion reduces centromeric CENP-C level *in vivo* by two-fold demonstrating that CENP-B critically contributes to CENP-C stability in addition to CENP-A (Fachinetti et al., 2015).

Our lab further found that CENP-B was neither required nor sufficient to recruit CENP-C to the centromere although we followed CENP-C loading in absence of CENP-A only for a short time. It remains unclear if preexisting CENP-C could have blocked free binding sites on CENP-B (Fachinetti et al., 2015; Hoffmann et al., 2016).

Once the centromere is assembled, the presence of endogenous CENP-B is sufficient for faithful chromosome segregation even when CENP-A is depleted (Hoffmann et al. 2016). We

concluded that the interaction between CENP-B and CENP-C is in this case sufficient to maintain the integrity of the centromere and allow proper kinetochore formation and function during mitosis (Hoffmann et al., 2016). Indeed, rapid removal of both CENP-A and CENP-B leads to a rapid loss of CENP-C (Hoffmann et al. 2016) (**Figure 8A**). In agreement with this explanation, Y chromosome missegregation can be induced in cells expressing a CENP-A chimera containing the C-terminal tail of histone H3 (Ly et al., 2017). Altogether, CENP-B and CENP-A seem to cooperatively bind to CENP-C to cooperatively support CENP-C stability.

The CENP-B/C interaction could also be at the core of non-random mis-segregation rate of different chromosomes. Active centromere length correlates with CENP-B protein level and CENP-C level and anti-correlates with centromere mis-segregation rate. Especially large chromosomes with small centromeres were found to mis-segregate more frequently when mitotic chromosome segregation was challenged (Dumont et al., 2020) (**Figure 8C**).

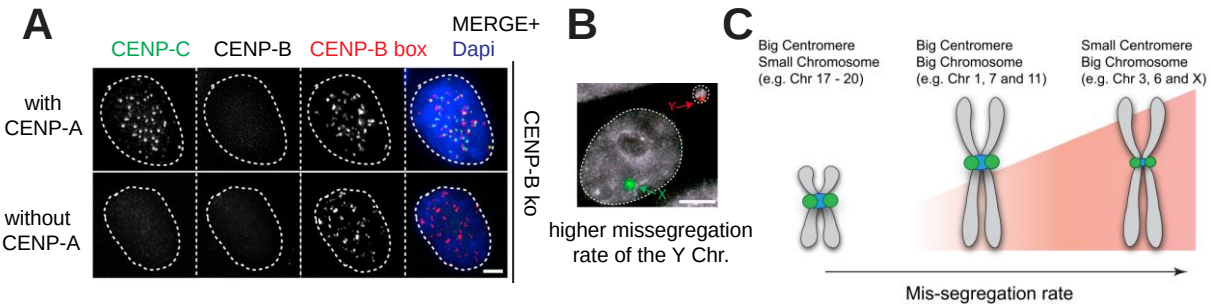


Figure 8: CENP-B and centromere strength.

*(**A**) Immunofluorescence coupled with fluorescence in situ hybridization (IF-FISH) in a CENP-B knock-out RPE-1 cell line with CENP-A or after rapid removal of CENP-A (from Hoffmann et al. 2016). CENP-B is missing to stabilize CENP-C following rapid CENP-A removal.*

*(**B**) FISH image of a human DLD-1 cell possessing a micronucleus containing Y centromeric DNA (from Fachinetti et al. 2015). Centromere strength of the Y centromere is weaker resulting in higher mis-segregation rate of the Y chromosome.*

*(**C**) Model of centromere strength from Dumont et al. 2020. Centromeric CENP-B level crucially impacts centromere strength. CENP-B abundance is especially important for larger chromosomes.*

While CENP-B, CENP-A and CENP-C are interacting partners, a great number of CENP-B molecules likely do not engage in these interactions. IP experiments suggest that ~20-25% of total endogenous CENP-B is in the same chromatin fraction like CENP-C (Suzuki et al., 2004); and around half of the amount of CENP-B interacts with canonical H3 nucleosomes rather than CENP-A nucleosomes (Ando et al., 2002). ChIP experiments in addition report that CENP-C and CENP-B bind to similar alpha-satellite sequences but at different regions at the centromere (Politi et al., 2002), however the repetitive nature of CenDNA likely prevents a definitive answer. This is further supported by fluorescence microscopic imaging and electron microscopy images that reveal partial co-localization of CENP-A/C with CENP-B (Cooke et al., 1990; Owen J Marshall et al., 2008; Saitoh et al., 1992; Sugimoto et al., 1999) (**Figure 9**). What regulates CENP-B interaction with centromere proteins remains unclear. In addition to heterochromatin formation mechanisms, CENP-B is posttranslationally modified which could be another important component in the regulation of CENP-B interaction with centromere proteins. Interestingly, CENP-B can be phosphorylated *in vitro* within the acidic domain at a conserved SDEEE motif likely via the casein kinase 2 and sumoylated at K407 residue (Maalouf et al., 2018; Sugimoto and Himeno, 1992). It remains untested if these PTMs regulate CENP-C interaction, for instance. Further, there is little evidence for casein kinase 2-mediated SDEEE phosphorylation *in vivo*.

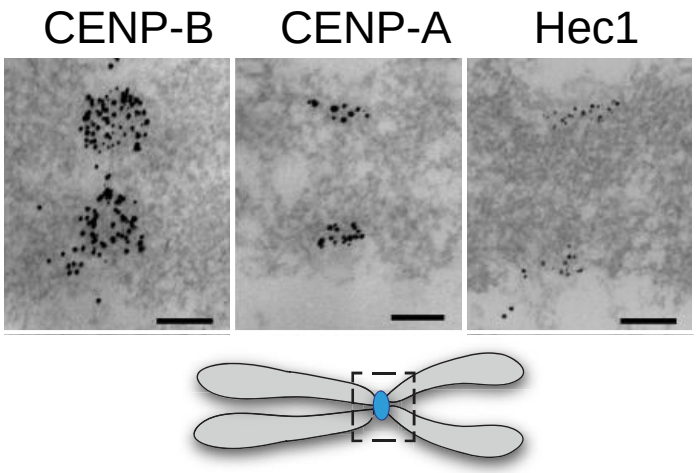


Figure 9: Electron microscopy images of CENP-B, CENP-A and Hec1 adapted from Mashall et al. 2008. *In contrast to CENP-A, CENP-B is also found in most inner centromeric regions. Scale bars 200 nm.*

Finally, in the context of CENP-B-mediated centromere stabilization, FRET (Förster resonance energy transfer) analysis also suggested that CENP-B could interact with CENP-T (Hellwig et al., 2008). However, CENP-A's N-terminal tail likely interacts with both CENP-B and CENP-T

(Fachinetti et al., 2015; Hoffmann et al., 2016; Logsdon et al., 2015; Suzuki et al., 2004) which could attribute to spatial proximity of CENP-B and -T. Further evidence that demonstrating a possible CENP-B/CENP-T interaction is missing so far.

Despite the functional implications described above, in mice, CENP-B knock-out did not cause obvious strong phenotypic changes in the first generation (Hudson et al., 1998; Kapoor et al., 1998; Perez-Castro et al., 1998). These results are in agreement with the observations that centromeres can be to some extent functional and propagated even in the absence of CENP-B. However, mice can tolerate genomic instability to some extent, which could explain the lack of phenotypic abnormalities in absence of CENP-B. Aneuploid cells were found to be eliminated by the immune system or are eliminated through p53 activation (Chunduri and Storchová, 2019). Another indication that this may be an important aspect to consider comes from the study of Hudson et al. (1998) in which weight and sperm count of CENP-B knock-out mice were indeed significantly decreased suggesting more frequent cell death. Furthermore, follow-up investigations also identified age-dependent uterine dysfunction, loss of fertility in female mice and reduced reproductive fitness in CENP-B knock-out mice (Fowler et al., 2000).

In conclusion, centromere function and propagation can occur in absence of CENP-B but the protein contributes to centromere strength. On one hand, CENP-B seems to promote heterochromatin formation and is insufficient to promote centromere formation; but on the other hand, CENP-B interacts and stabilizes CENP-A and CENP-C. CENP-B's relevance for centromere identity remains thus unclear and requires deeper characterization.

b) Regulation of CENP-B binding at the centromere

CENP-B was reported to show high dynamic turnover rates in G1/S phase and almost no turnover in G2/M (Hemmerich et al., 2008). A variety of indications and hypotheses have been put forward on how CENP-B might be regulated but a clear picture is still missing.

As mentioned earlier, one key regulatory feature of CENP-B binding to CENP-B boxes could be DNA methylation (Scelfo and Fachinetti, 2019). In addition, CENP-B itself was described to be tri-methylated at the N-terminus. This methylation was proposed to increase CENP-B's affinity to bind to CENP-B boxes perhaps as a response to cellular stress (Dai et al., 2013).

Another interesting protein for CENP-B regulation is NAP1, which was reported to have different cellular functions, most prominently as a histone chaperone for H2A and H2B (**see section 1.2**). NAP1 was found *in vitro* to interact with the DNA binding domain of CENP-B and with full-length CENP-B *in vivo*. NAP1 facilitates CENP-B binding to CENP-B boxes on one hand, while on the other hand seems to prevent association of CENP-B with CENP-A in absence of CENP-B boxes association (Tachiwana et al., 2013).

Another study proposed that INMAP (Interphase Nucleus and Mitotic Apparatus Associated Protein), a protein also associated with centrosome function, interacts with CENP-B and somehow promotes the cleavage of the N-terminal domain of CENP-B during interphase and thus regulates CENP-B centromere localization (Tan et al., 2014). CENP-B was recently also described to be controlled by yet another regulation mechanism involving sumoylation and ubiquitination/proteasomal-dependent degradation (Maalouf et al., 2018).

Finally, CENP-B centromere stability is also influenced by centromere proteins like CENP-A *in vivo*. Depletion of CENP-A or its N-terminal tail leads to a slight decrease of centromeric CENP-B level (Fachinetti et al., 2015, 2013).

In conclusion, although several mechanisms of CENP-B regulation have been proposed, the interplay of these regulation mechanisms and the details of CENP-B's cell cycle-dependent stability at the centromere require deeper investigation.

II.6 De Novo Centromere Formation

The preservation of a unique centromere per chromosome is a crucial challenge for monocentric species. As discussed in **section 1.5**, gain or loss of centromeres can have severe consequences for chromosome segregation. Centromeres are maintained by an epigenetic mechanism (**section 2.3**). DNA sequence and in particular CENP-B seem to be genetic factors that support the stability of the centromere but their contribution to centromere identity and potential for centromere formation independent of the epigenetic mark remains a subject of debate (**section 2.4** and **2.5**). Especially, the discoveries of neocentromeres strongly support the view that DNA sequence can neither be sufficient nor be required for centromere propagation and function (Amor et al., 2004; Voullaire et al., 1993).

a) Human Neocentromeres and inactivated centromeres

In human, over 100 neocentromere cases have been discovered in patients (Owen J. Marshall et al., 2008; Murillo-Pineda and Jansen, 2020). Human neocentromere form on different ectopic chromosome locations and are deprived of CENP-B boxes and CENP-B. Neocentromeres are characterized by the presence of CENP-A, all CCAN proteins and their ability to form kinetochores/spindle connections (Alonso et al., 2007; Owen J. Marshall et al., 2008; Saffery et al., 2000). Moreover, neocentromeres are stably inherited in cycling cells and can be inherited over generations (Amor et al., 2004; Bodor et al., 2014; Murillo-Pineda and Jansen, 2020). The epigenetic environment found at neocentromeres is highly similar to native centromeres further supporting the epigenetic nature of centromere function and identity (**see section 2.2**) (Bodor et al., 2014; Murillo-Pineda and Jansen, 2020).

An intriguing case of a neocentromere is a patient-derived PseudoDicentric NeoCentromere found on the q-arm of chromosome 4 (PD-NC4) (Amor et al., 2004). While neocentromeres are usually associated with genomic rearrangements, often related to loss or breakage of the native centromere (Hasson et al., 2011), in this case chromosome 4 as well as the rest of the genome appeared normal (46, XX) with the exception of the centromere position on one of the chromosome 4 homologues. The neocentromere was discovered in a patient with mild cognitive impairment but father and brother of this patient carried the same neocentromere and were healthy. The pseudodicentric chromosome was traced back to the grandfather, however, because the grandfather was not available for this study, it remains unclear if the generation of the neocentromere and inactivation occurred in the germline or was already present before (Amor et al., 2004). Analysis of the inactive centromere showed that the alpha-satellite arrays were slightly reduced compared to a normal chromosome 4 but more or less in the range of individual variation. What caused the inactivation of the original centromere in this case and which are the order of events that lead to *de novo* centromere formation and centromere inactivation remain elusive. There is some indication that centromere inactivation could be a triggering event for neocentromere formation (**see section 2.6b**). However, it is also possible that ectopic centromere formation is a stochastic event with extremely low probability and that inactivation of the centromere occurred after neocentromere formation. Artificial generation of neocentromeres in *Drosophila* was recently found to promote the inactivation of the original centromere and chromosome breaks at the centromere appeared to be a prerequisite for this (Palladino et al., 2020).

Analysis of the epigenetic environment of inactive regional centromeres in yeast, plants and on inactive HOR arrays in human revealed that repeat arrays at these inactive centromere sites are more compact, enriched for heterochromatin marks and deprived of acetylated histones (Han et al., 2009; Maloney et al., 2012; Sato et al., 2012; Zhang et al., 2010). In an artificially generated cell line with a X chromosome comprising two active centromeres, centromere inactivation of one of the two centromeres was occasionally observed upon treatment with a histone deacetylase inhibitor (TSA, Trichostatin A) (Higgins et al., 2005). In contrast, TSA treatment was also proposed to trigger the opposite effect: The ectopic integration of artificially generated centromere sequence arrays into a host chromosome usually do not evoke the generation of an active centromere. However, upon TSA treatment centromere activation in this case has been observed (Nakano et al., 2003). This study hence further indicates that centromere DNA sequence in a certain epigenetic environment could provide centromere formation potential even when integrated in a chromosome. However, TSA treatment in the PD-NC4 case was insufficient to trigger activation of the inactive original centromere (Amor et al., 2004). It is possible that fine-tuning of the chromatin state for centromere activation may be challenging and centromere activation could be a rare event easy to miss. Hence, it remains unclear to what extent epigenetic silencing can overrule a putative intrinsic capability of centromeric DNA sequence to initiate centromere formation (**see section 2.6c**).

Interestingly, CENP-A level at the PD-NC4 neocentromere were only slightly lower compared to a normal centromere of chromosome 4; but CENP-C level and centromere strength of the neocentromere was later found to be weaker due to the absence of CENP-B (Amor et al., 2004; Fachinetti et al., 2015). In addition, the mitotic error correction mechanism (**see section 1.4**) is impaired at neocentromeres due to altered Aurora B location suggesting perturbations in the inner centromere that are either related to missing centromeric DNA sequences or other factors such as a non-optimal chromatin environment (Bassett et al., 2010).

b) Artificial Generation of Neocentromeres

Since neocentromeres are potentially associated with karyotype speciation and disease development, understanding how they might form is of great interest. Neocentromere formation was successfully induced following different approaches on chromosomes in different species (**Figure 9**).

Ectopic centromere formation can be achieved as mentioned previously by targeting certain centromere/kinetochore proteins or HJURP to ectopic chromosome sites (Barnhart et al., 2011;

Gascoigne et al., 2011; Mendiburo, 2011; Palladino et al., 2020; Shono et al., 2015) (**Figure 9B**). While this shows that tethering centromeric proteins to ectopic region can be sufficient to trigger centromere formation, strong accumulation and solid binding of centromere proteins at non-centromeric regions is highly artificial and give little insight into how neocentromeres could form naturally.

In fungi (*S. Pombe* and *C. Albicans*), chicken, and human, neocentromeres were occasionally generated following the excision or replacement of the endogenous centromere (Ishii et al., 2008; Shang et al., 2013) (Murillo-Pineada et al. (Jansen lab)) (**Figure 9C**). In *Drosophila*, centromere formation was also found as a consequence of gamma-irradiation and chromosome breaks (Williams et al., 1998). Similar results were obtained in plants (Nasuda et al., 2005). It is thus tantalizing to speculate that neocentromeres naturally occur when the native centromere is inactivated in some way or that DNA breaks require neocentromere formation to prevent loss of genetic material during mitosis. However, this may not be the case for the PD-NC4 observed in human (Amor et al., 2004) (**Figure 9D**). In *Drosophila*, neocentromere formation was also induced by CENP-A overexpression (Heun et al., 2006; Olszak et al., 2011) (**Figure 9A**). In contrast to tethering assays, in both cases, neocentromeres formed at non-artificially designated chromosome sites raising the question if there are preferential sites within the genome for neocentromere formation.

Interestingly, hotspots for neocentromere formation were sub-telomeric regions or regions with boundaries between heterochromatin and euchromatin in flies. As described previously, native centromeres are weakly transcriptionally-permissive but surrounded by heterochromatin regions (**see section 2.2c**). Hence these boundary regions seem to provide favorable conditions for centromere establishment. Also, in *S. pombe,* heterochromatin was found to be important as the depletion of the methyltransferase Clr4 (mediates H3K9 methylation) reduces the chance for neocentromere formation (Ishii et al., 2008). In contrast however, in chicken and in several human patient-derived neocentromeres, hotspots for neocentromere formation are not identified (Alonso et al., 2010; Murillo-Pineda and Jansen, 2020; Shang et al., 2013).

In conclusion, it still remains unclear especially in vertebrates what triggers neocentromere formation and if certain chromatin sites could function as potential hotspots to overcome a potential natural barrier of chromatin to form a centromere *de novo*.

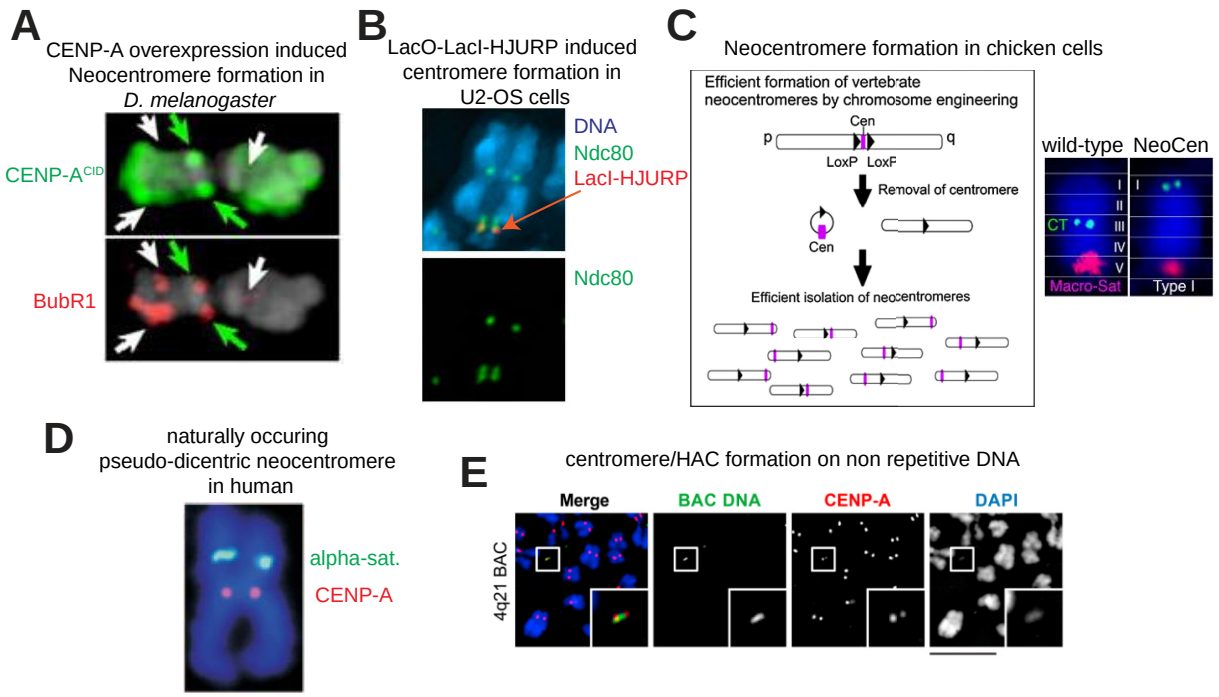


Figure 10: Selection of Neocentromeres naturally (D) or artificially (A-C and E) formed in different systems.

*(**A**) CENP-A induced neocentromere formation by CENP-A overexpression in flies (adapted from Heun et al. 2006).*

*(**B**) Artificial tethering of HJURP can induce centromere formation at an ectopic chromosome location in human (adapted from Barnhart et al. 2011). Generation of a neocentromere on the Z chromosome in chicken cells after excision of the original centromere (CT=CENP-T) (adapted from Shang et al. 2013).*

*(**D**) Naturally occurring pseudo-dicentric neocentromere chromosome 4 (PD-NC4) discovered in human.*

*(**E**) Centromere formation without the aid of artificial tethering strategy on a bacterial artificial chromosome deprived of repetitive DNA sequence in human cells. HAC formation however was accompanied with acquisition of genomic sequences in the BAC sequence. (adapted from Logsdon et al. 2019).*

In contrast to flies, CENP-A overexpression in human is insufficient to trigger neocentromere formation, even though a strong increase of CENP-A and CENP-C on chromosome arms are observed (Lacoste et al., 2014; Shrestha et al., 2017). The absence of the Mis18 complex suggests

that the regulation of centromere formation is less complex in flies (**see section 2.3b** for the role of Mis18). It is thus tantalizing to speculate that the lack of the Mis18 complex could explain neocentromere formation after CENP-A overexpression in flies while this is not observed in human. Also, ectopic CENP-A was described to be assembled mostly in CENP-A-H3.1 heterotypic particles unlike homotypic nucleosomes found at the centromere (Lacoste et al., 2014). Hence it is possible that CENP-A nucleosomes need to be homotypic and highly accumulated in close proximity in order to trigger centromere formation in human (Bodor et al., 2014; Lacoste et al., 2014), which even after CENP-A overexpression is unlikely to occur at ectopic locations. Moreover, overexpressed CENP-A is mis-incorporated through different assembly mechanisms in human and *Drosophila*. In human, the H3.3 chaperone DAXX is responsible for CENP-A mis-incorporation resulting mainly in CENP-A mis-incorporation in euchromatic regions (Lacoste et al., 2014). In *Drosophila*, the DAXX homologue is not involved; instead a chromatin remodeling complex (NuRD) was reported to be important for CENP-A mis-incorporation (Demirdizen et al., 2019), prompting speculations to what extent the assembly machinery could enable (*Drosophila*) or suppress (human) neocentromere formation. Despite the lack of neocentromere formation upon CENP-A overexpression in human, CENP-A overexpression is commonly found in cancers which may be related to gene expression alterations caused by CENP-A (Hu et al., 2010; Lacoste et al., 2014; Tomonaga et al., 2003). However, future studies will be required to improve our understanding of the causality.

Another way to study centromere formation are artificial chromosome formation assays. Especially, bottom-up assembly strategies in which naked DNA containing centromere DNA sequences is introduced into cells demonstrated that certain centromeric DNA sequences and especially the presence of CENP-B boxes greatly facilitate centromere formation on HACs (Harrington et al., 1997; Ikeno et al., 1998; Ohzeki et al., 2002; Okada et al., 2007).

However, as mentioned before, centromere formation was described to be distinct on inserted potentially naked DNA (BACs) compared to ectopic chromosome sites where the presence of centromeric DNA sequences were found to be insufficient for centromere formation and CENP-B presence instead promote heterochromatin formation (Okada et al., 2007) (**section 2.5a**). Logsdon et al. 2019 recently reported for the first time that HACs can be generated even in complete absence of centromeric DNA sequences or centromere seeding using LacO/I or tetO/R (**section 2.4**). However, in these cases genomic sequences were acquired and in half of those, even alpha-satellite DNA sequences.

HACs have been used to better understand the effect of chromatin environment on centromere formation (**section 2.2c, section 2.5a**). Moreover, HAC formation is of greater general interest due to its potential biotechnological application especially for gene therapy or as a gene delivery system. HACs indeed can deliver multiple genes, avoid genomic integration but remain stably transmitted and can be inactivated in patients or livestock if desired (Cardinale et al., 2009; Kononenko et al., 2015). HACs with reporter genes were also used as sensors to screen compounds that induce chromosome instability (Kim et al., 2016; Lee et al., 2013). Moreover, artificial chromosome formation is at the core of ambitious attempts to artificially create entire genomes from scratch with a great variety of possible biotechnological applications (Boeke et al., 2016).

III.1 Summary and aims of this thesis

Centromeres fulfil an essential and highly conserved role in proliferating cells (section 1.4/5 and 2.1). Centromeres possess a unique epigenetic landscape (section 2.2), which is sustained for indefinite cell cycles by an epigenetic mechanism (section 2.3). Despite this epigenetic nature, centromeres are also commonly assembled on special DNA sequences (section 2.4). The centromere protein CENP-B binds specifically to a sequence motif found at the centromere (section 2.5). DNA sequence and CENP-B contribute to centromere strength but centromeres can be functional in their absence. Great uncertainty surrounds the role of centromere DNA and especially CENP-B bound to this DNA for centromere identity.

During my PhD, I aimed to investigate the interplay between genetic and epigenetic centromeric features to better understand centromere specification.

First, I scrutinized the requirement of preexisting CENP-A molecules for centromere identification. Using a CENP-A$^{OFF/ON}$ system, I aimed to understand several outstanding questions regarding the role of preexisting centromeric CENP-A molecules for a) *de novo* CENP-A deposition at the native centromere; b) an epigenetic memory of CENP-A level in the next cell generation(s) and; c) guiding new CENP-A assembly to designated reloading sites.

These questions were addressed by taking advantage of the rapidness and reversibility of the auxin inducible degron (AID) system (Hoffmann et al., 2016; Hoffmann and Fachinetti, 2018; Nishimura et al., 2009).

Furthermore, I aimed to understand if and how CENP-B plays a role in centromere specification at a CENP-A-deprived native centromere.

III.2 Résumé et objectifs de cette thèse

Les centromères remplissent un rôle essentiel, fortement conservé, dans la prolifération des cellules (sections 1.4/5 et 2.1). Les centromères possèdent un paysage épigénétique unique (section 2.2) qui est maintenu par un mécanisme épigénétique au cours des cycles cellulaires illimités (section 2.3). Malgré cette nature épigénétique, les centromères sont aussi couramment assemblés sur des séquences d'ADN particulières (section 2.4). La protéine CENP-B du centromère se lie spécifiquement à un motif de séquence localisé au centromère (section 2.5). La séquence d'ADN et CENP-B renforcent le centromère, mais celui-ci maintient sa fonction en leur absence. Le rôle de l'ADN du centromère et surtout de la protéine CENP-B liée à cet ADN dans l'identité du centromère reste très incertain. Au cours de mon doctorat, j'ai cherché à étudier l'interaction entre les caractéristiques génétiques et épigénétiques des centromères afin de mieux comprendre la spécification des centromères.

Les centromères remplissent un rôle essentiel, fortement conservé, dans la prolifération des cellules (sections 1.4/5 et 2.1). Les centromères possèdent un paysage épigénétique unique (section 2.2) qui est maintenu par un mécanisme épigénétique au cours des cycles cellulaires illimités (section 2.3). Malgré cette nature épigénétique, les centromères sont aussi couramment assemblés sur des séquences d'ADN particulières (section 2.4). La protéine CENP-B du centromère se lie spécifiquement à un motif de séquence localisé au centromère (section 2.5). La séquence d'ADN et CENP-B renforcent le centromère, mais celui-ci maintient sa fonction en leur absence. Le rôle de l'ADN du centromère et surtout de la protéine CENP-B liée à cet ADN dans l'identité du centromère reste très incertain. Au cours de mon doctorat, j'ai cherché à étudier l'interaction entre les caractéristiques génétiques et épigénétiques des centromères afin de mieux comprendre la spécification des centromères.

The following sections (results, discussion and methods) are published in the EMBO Journal (Hoffmann et al. 2020).

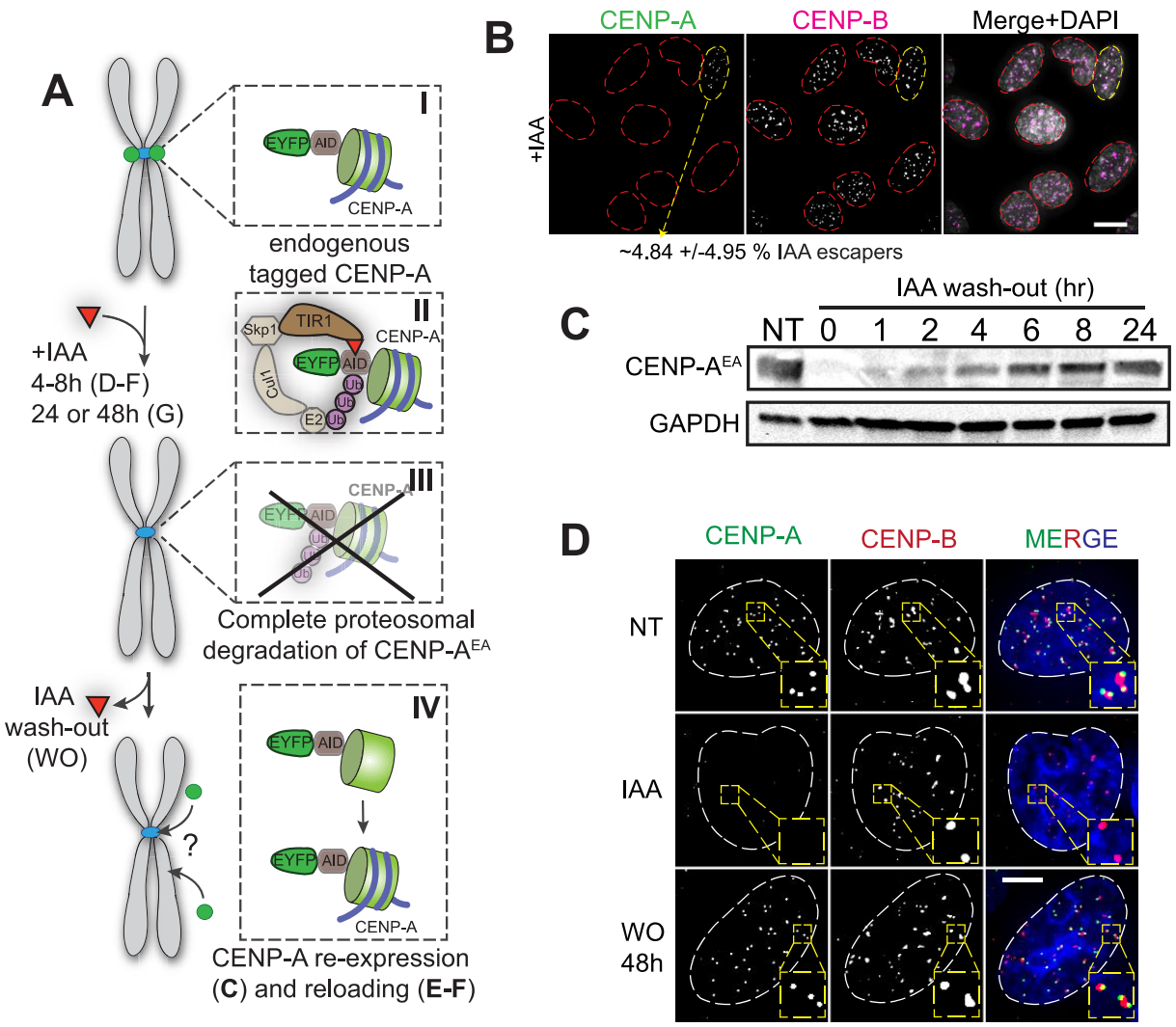
5. RESULTS

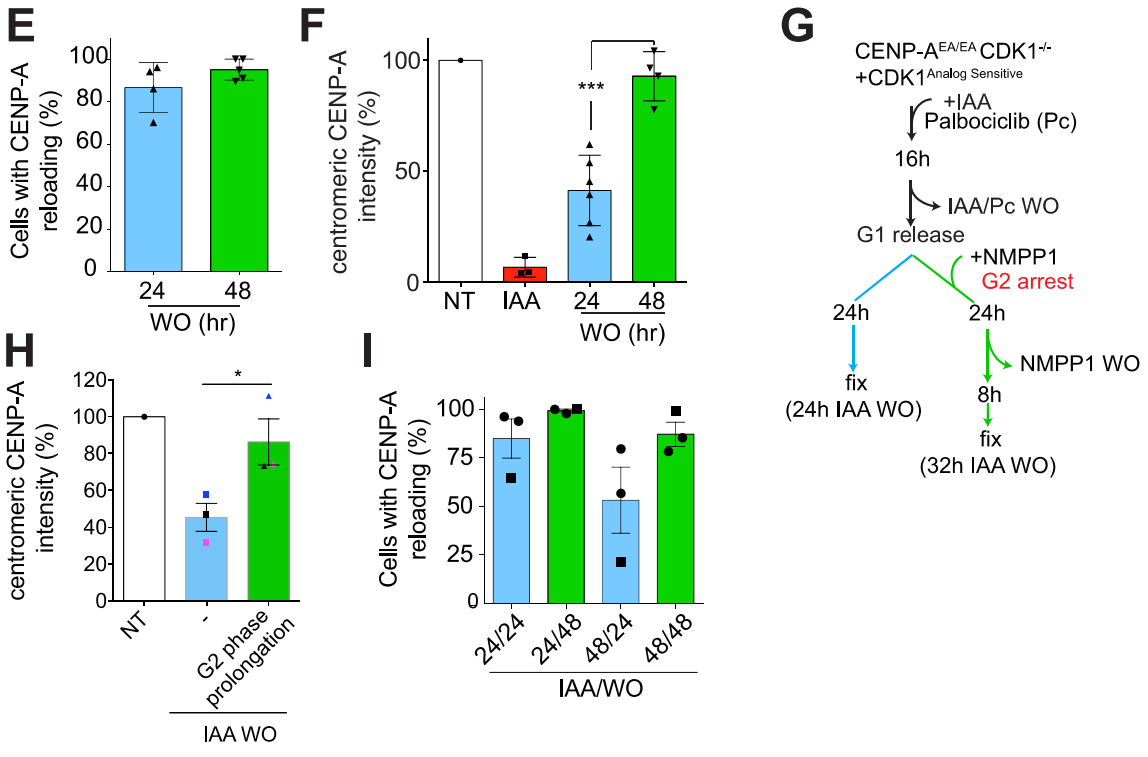
Previously deposited CENP-A is not essential for new CENP-A deposition at endogenous centromeres

CENP-A is well known to maintain centromere position via an epigenetic self-assembly loop (McKinley and Cheeseman, 2016). This suggests that at least a pool of CENP-A must always be maintained at the centromere to mediate new CENP-A deposition. Here, we sought to challenge this concept and test if previously deposited centromeric CENP-A is required to license new CENP-A deposition at the native centromere position. To this aim, we used a two-step assay (hereafter referred to as CENP-A$^{OFF/ON}$ system) that allows us, in a first step, to deplete endogenous CENP-A and, subsequently, to re-express it (**Figure 11A**). To generate this unique tool/model, we took advantage of the reversibility of the Auxin Inducible Degron (AID) system that allows rapid protein depletion and re-accumulation following synthetic Auxin (indol-3-acetic acid, IAA) treatment and wash-out (WO), respectively (Hoffmann and Fachinetti, 2018; Holland et al., 2012; Nishimura et al., 2009). By combining genome editing with the AID tagging system, we previously showed our ability to rapidly and completely remove endogenous CENP-A from human centromeres with a half-life of about 15 min (Hoffmann et al., 2016), with only a small percentage of cells that remain unaffected by the IAA treatment (**Figure 11B)**. Following IAA WO, endogenous CENP-A$^{EYFP\text{-}AID}$ (hereafter referred to as CENP-A^{EA}) is rapidly (within 1-2 hours) re-expressed at detectable level as observed by immuno-blot (**Figure 11C**). The rapid CENP-A re-accumulation could be explained by the continuous presence of mRNA CENP-A transcripts despite the immediate protein degradation in presence of IAA. Hence, the generated CENP-A$^{OFF/ON}$ system provides a powerful tool to test CENP-A reloading in the absence of previously deposited CENP-A.

We next tested if newly expressed CENP-A is reloaded back at the native centromere position by immuno-fluorescence (IF). Following CENP-A depletion for 4-8 hours (hr) and IAA WO for 24 hr or 48 hr, we found that in most (~90%) of the cells, newly-expressed CENP-A re-localizes with centromeric regions marked by CENP-B, which remains tightly bound to the CENP-B boxes (**Figure 11D-F**). Interestingly, centromeric CENP-A levels recovered to only ~50% of untreated

levels after one cell cycle (24 hr WO), but fully recovered to untreated levels after IAA 48h WO (**Figure 11E**). This was due to incomplete recovery of total CENP-A levels rather than absence of pre-existing CENP-A molecules or preexisting centromeric factors that mediate CENP-A assembly. Indeed, centromeric CENP-A level recovered to untreated level within one cell cycle when we prolonged G2 phase – time where most CENP-A is transcribed (Shelby et al, 1997) – after IAA WO using a CDK1 analog sensitive inhibition system (Hochegger et al., 2007; Saldivar et al., 2018) (**Figure 11G-H**). Following short-term CENP-A depletion, many CCAN components partially remain at centromeric regions (Hoffmann et al., 2016), potentially promoting CENP-A deposition (Hori et al., 2008a; McKinley et al., 2015; Okada et al., 2006). We thus depleted CENP-A for longer durations (24-48 hr) as, in these conditions, most CCAN proteins are lost from centromeres (Hoffmann et al., 2016). We used p53-deficient DLD-1 cells and chromosomally unstable U-2OS cells to bypass cell cycle block due to events of chromosome mis-segregation following CENP-A depletion. Even in these conditions, newly expressed CENP-A molecules were reloaded at CENP-A-depleted centromeres (**Figure 11I**).





(figure legend on next page)

Figure 11. The CENP-A epigenetic self-assembly mechanism is not required for de novo CENP-A deposition at native human centromere.

*(**A**) Schematic illustration of the two-step CENP-A$^{OFF/ON}$ assay using the Auxin (IAA) inducible degradation system.*

*(**B**) Image of IAA treated cells. IAA escaper is highlighted with a dashed yellow circle, CENP-A depleted cells are contoured with red dashed lines. Scale bar, 10 µm.*

*(**C**) Immunoblot showing CENP-A^{EA} protein level at the indicated time in RPE-1 cells.*

*(**D**) Representative immunofluorescence images showing CENP-A reloading at CENP-B-marked centromeres in DLD-1 cells. White dashed circles contour nuclei. Scale bar, 5 µm.*

*(**E**) Quantification of the percentage of cells showing centromeric CENP-A 24 hr or 48 hr after IAA WO. Each dot represents one experiment (~30-50 cells per condition per experiment), error bars represent standard deviation (SD) of 5 independent experiments.*

*(**F**) Quantification of centromeric CENP-A levels normalized to non-treated level. Each dot represents one experiment, error bars represent SD. Unpaired t test: ***p= 0.0005.*

*(**G**) Schematic for the experiments shown in H.*

*(**H**) Quantification of centromeric CENP-A levels normalized to non-treated level. Each dot represents one experiment, error bars represent SD. Unpaired t test: *p= 0.0493.*

*(**I**) Quantification of the relative number of DLD-1 (square) and U-2OS (circle) cells with centromeric CENP-A at the indicated timing of IAA treatment and recovery. Each dot represents one experiment with at least 20 cells per condition. Error bars represent standard deviation (SD) from 3 independent experiments.*

We then followed *de novo* CENP-A reloading using live cell imaging taking advantage of the EYFP tag on the endogenous CENP-A (Hoffmann et al., 2016). To mark centromere position in live cells, we endogenously tagged CENP-B with mCherry using CRISPR/Cas9 in RPE-1 cells. Following the induction of a CENP-A$^{OFF/ON}$ cycle we observed a burst of reloading of CENP-A shortly after mitotic exit (approximately 30 min after anaphase onset) (**Figure 12**), in agreement with the previously described timing of CENP-A reloading (Jansen et al., 2007b). This experiment showed that CENP-A reloading in the absence of any previously deposited centromeric CENP-A is

still tightly restricted to a narrow cell cycle window. So, we further examined if *de novo* CENP-A reloading relies on the same key regulation mechanisms as canonical CENP-A deposition (McKinley and Cheeseman, 2016).

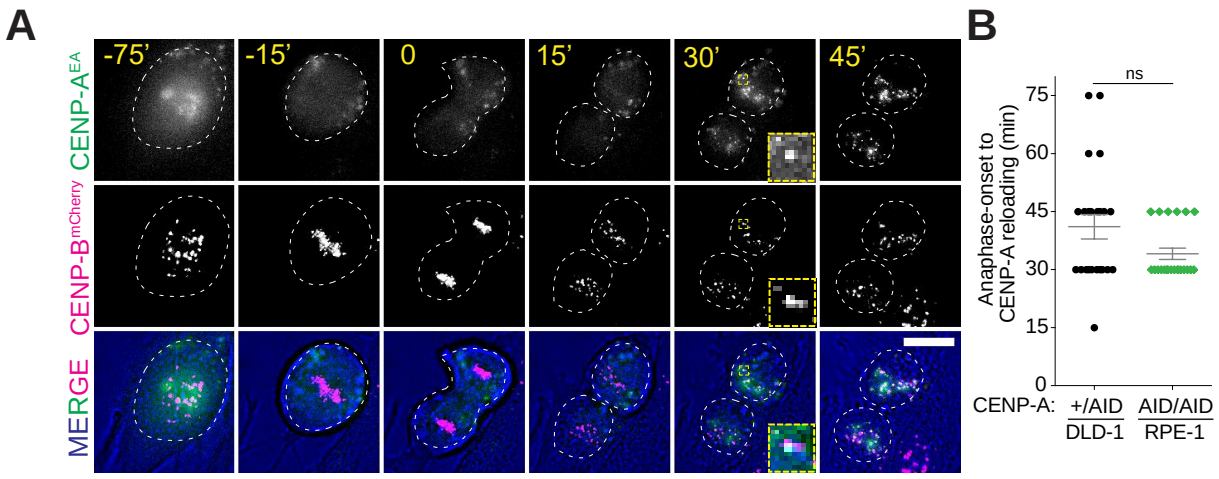


Figure 12: De novo CENP-A deposition occurs shortly after mitotic exit.

*(**A**) Stills of live cell imaging to follow de novo CENP-A^{EA} reloading in RPE-1 cells harboring endogenously tagged CENP-B^{mcherry}. Images were taken every 15 min. White dashed circles contour nuclei prior/after mitosis and cells during mitosis based on bright field images. Scale bar, 10 µm.*

*(**B**) Dot plot showing the timing of CENP-A^{EA} reloading after anaphase onset in the indicated cell lines. Each dot represents one cell, error bars represent standard deviation. Unpaired t test, ns.*

We confirmed that *de novo* CENP-A deposition is dependent on M18BP1 and HJURP, but not DAXX, a histone chaperone that was shown to be involved in non-centromeric CENP-A deposition (Lacoste et al., 2014) (**Figure 13**). However, as HJURP depletion strongly impacts the stability of soluble CENP-A, re-expression of CENP-A was also strongly affected by HJURP depletion making a direct conclusion on HJURP dependency uncertain (**Figure 13F**).

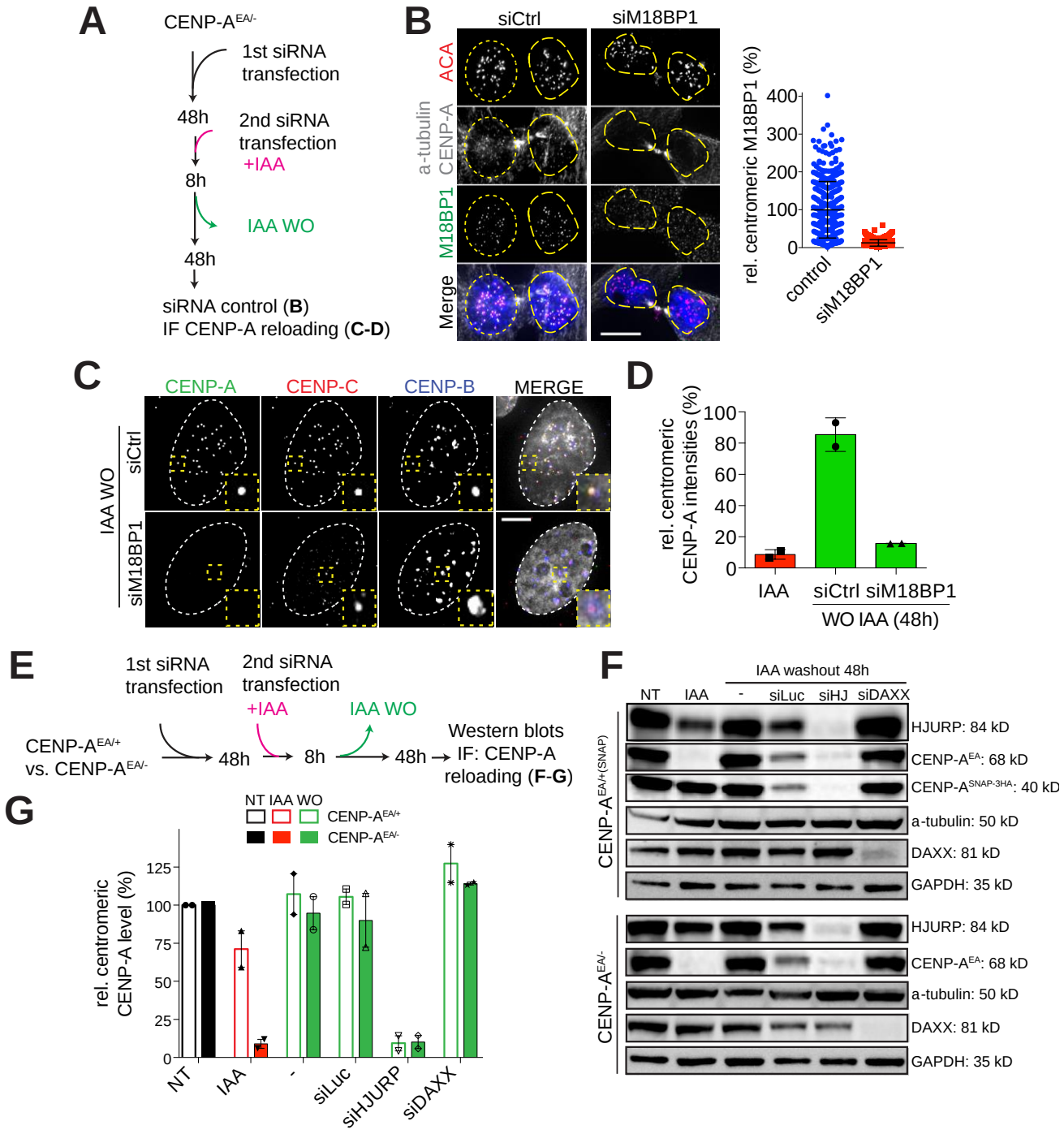


Figure 13: De novo CENP-A deposition follows the canonical CENP-A deposition pathway.

*(**A**) Schematic for the experiments shown in C, D.*

*(**B**) Left panel: representative images to confirm M18BP1 knock-down in late M-phase cells. Scale bar, 10 µm. Yellow dashed lines highlight nuclei of daughter cells. Right panel: relative M18BP1 levels in late M/early G1 phase after siRNA knock-down using M18BP1 antibody. Each dot represents one centromere, error bars represent standard deviation.*

*(**C**) Representative images of de novo CENP-A reloading upon M18BP1 knock-down. Nuclei are highlighted with white dashed lines. Scale bar, 5 µm.*

*(**D**) Quantification of centromeric CENP-A intensities in the indicated conditions (relative intensities normalized to CENP-A level in untreated cells). Each dot represents one experiment (>30 cells per condition per experiment), error bars represent SD of 2 independent experiments.*

*(**E**) Schematic representation for the experiments shown in F-G.*

*(**F**) Immunoblot of total protein levels in the indicated cell lines and conditions*

*(**G**) Bar graphs showing quantification of centromeric CENP-A intensities following the indicated treatment. Each dot represents one experiment with at least 30 cells. Error bars represent SD of 2 independent experiments.*

Altogether, our data indicate that centromere position is preserved even in the absence of CENP-A. Also, it demonstrates that previously assembled CENP-A is not essential for *de novo* deposition of CENP-A nor to control its abundance, as levels of new CENP-A rise very fast at CENP-A-depleted centromeres, a result in disagreement with the template model. Further, like canonical CENP-A reloading in the presence of previously deposited CENP-A, *de novo* CENP-A reloading is cell cycle regulated and occurred exclusively after mitotic exit.

Our results rely on complete depletion of CENP-A following IAA addition, as we have previously shown by IF, immuno-blot and immuno-precipitation (Hoffmann et al., 2016). To further prove the efficiency of the Auxin system, we challenged it by inducible, doxycycline-mediated Over-Expression (OE) of CENP-A tagged with EYFP and AID (**Figure 14A**). CENP-A OE leads to elevated CENP-A incorporation at the centromere and also outside the centromere region (Lacoste et al., 2014; Nechemia-Arbely et al., 2017) . Using this system, we obtained ~2000 fold higher nuclear CENP-A^{EA} protein level as compared to endogenous CENP-A levels at a single centromere (**Figure 14B-C**). Despite the vast excess of CENP-A in these cells, no CENP-A was detectable upon IAA addition by IF (**Figure 14B-C**). We concluded that the AID system remains - by far - unsaturated under endogenous CENP-A expression level conditions, as we are able to deplete higher CENP-A^{EA} levels to non-detectable level.

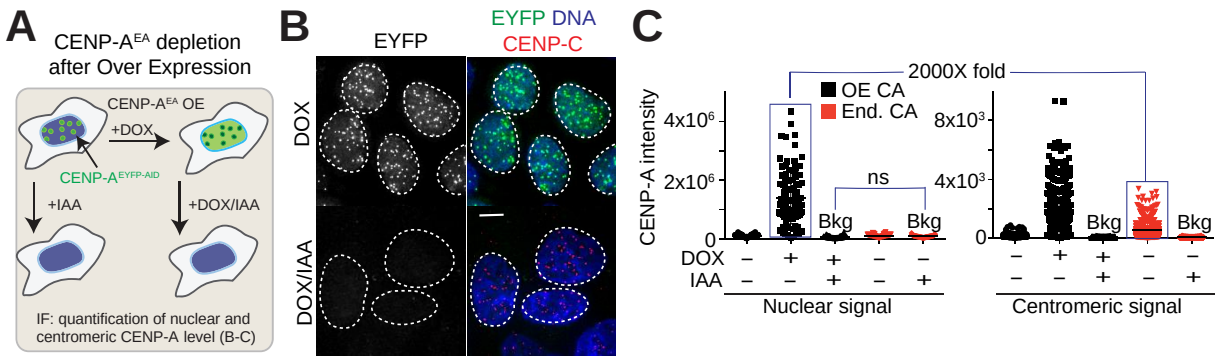


Figure 14: Depletion of overexpressed CENP-A$^{EYFP\text{-}AID}$ to undetectable level using the AID system.

*(**A**) Schematic illustration of experiment shown in B, C.*

*(**B**) Representative images showing complete depletion of CENP-A^{EA} despite doxycycline (DOX) induced overexpression in DLD-1 cells. Nuclei are contoured with white dashed lines. Scale bar, 5 µm.*

*(**C**) Quantification of CENP-A intensities in in the nucleus and at the centromere in presence or absence of IAA/DOX. Endogenous (End.) CENP-A$^{EYFP\text{-}AID}$ (CA) or overexpressed (OE) CA is depleted to non-detectable background-level in presence of IAA. Each dot represents a cell. Mean CA intensities are indicated by a black line.*

We then used live cell Single Molecule Microscopy (SMM) to assess the presence of single CENP-A molecules tagged with EYFP after IAA addition. We first confirmed the ability to detect a single EYFP molecule by transiently expressing EYFP-tagged GlucocorticoidReceptor (GR^{EYFP}) as it was previously used to study dynamics of single molecules in human cells (Harkes et al., 2015) (**Figure 15A**). Here, we observed a clear diffraction limited spot with a Gaussian profile that bleached in a single bleach step as expected when observing a single molecule (**Figure 15B-D**). In contrast, we were unable to detect such profile for CENP-A^{EA} at CENP-B^{mCherry}-positive centromeres following IAA treatment. Quantification of EYFP fluorescent intensities of IAA-treated CENP-A^{EA} cells at centromeres was significantly lower than the fluorescent signal of a GR^{EYFP} single molecule and comparable to the background signal obtained in EYFP-free RPE-1 CENP-C$^{mCherry\text{-}AID}$ cells (**Figure 15B-D**).

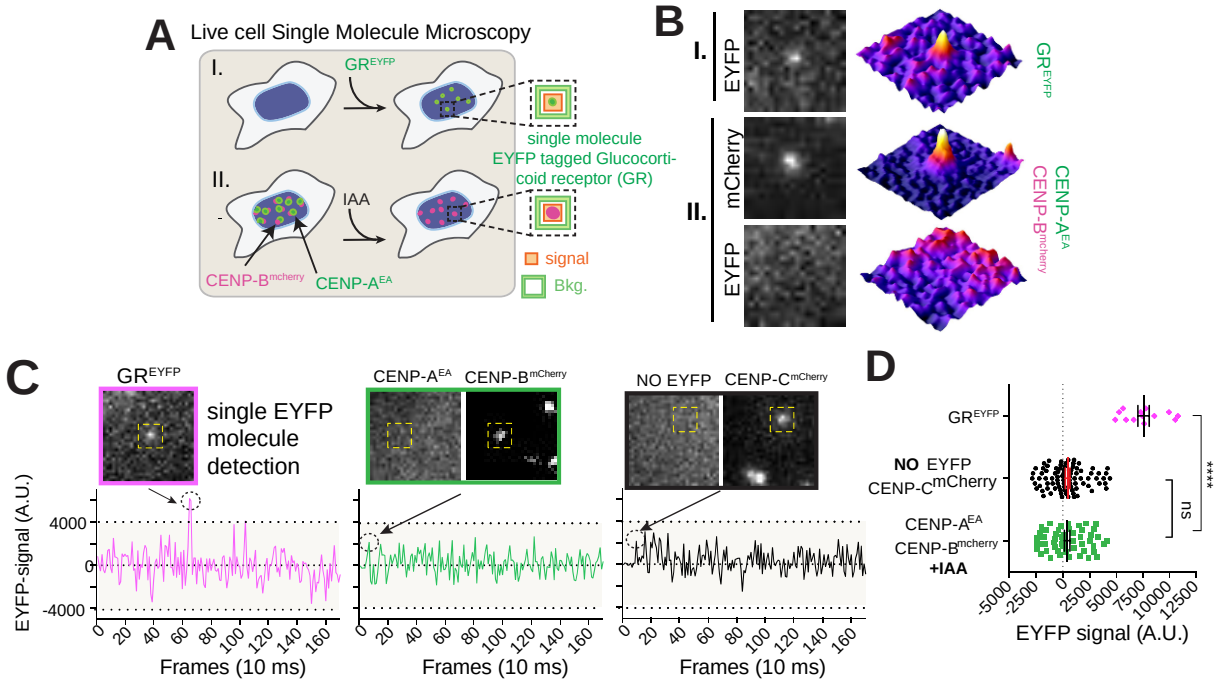


Figure 15: After depletion CENP-A$^{EYFP\text{-}AID}$ cannot be detected by Single Molecule Microscopy (SMM)

*(**A**) Schematic illustration of the single molecule microscopy (SMM) experiments shown in B, D.*

*(**B**) Representative microscopy images from live cell imaging and corresponding 3D surface plots showing single molecule GR^{EYFP} detection in I and following IAA treatment CENP-A^{EA} signal absence at CENP-B^{mcherry} marked centromeres in II, using SMM acquisition settings.*

*(**C**) Examples of background-corrected EYFP signal intensities quantified over time (as shown in D) for single GR^{EYFP} molecules (in magenta), centromeric EYFP signals in IAA treated CENP-A$^{EA/EA}$ cells (in green) and in absence of EYFP molecules (in black).*

*(**D**) Signal quantification as shown in D in the indicated conditions. Unpaired t test, ns (p=0.88), ****p < 0,0001, error bars represent standard deviation. Each dot represents the quantification of one GR^{EYFP} signal (GR^{EYFP}, n=13) or one centromere, respectively (No EYFP CENP-C^{mcherry}, n=85 and CENP-B^{mcherry} CENP-A^{EA}, n= 52).*

In summary, we concluded that CENP-A re-loading following CENP-A$^{OFF/ON}$ is unlikely to be due to any remaining CENP-A molecules. In addition, as CENP-A loading occurs only at mitotic exit, the dynamic equilibrium between its IAA-mediated degradation and re-expression does not impact on our results.

De novo CENP-A localization remains unaltered in absence of old centromeric CENP-A

Human centromeres display a hierarchical organization ultimately structured as higher order repeat arrays (HOR) (Miga, 2017). Some chromosomes have centromeres with multiple HOR arrays which display different abundance of CENP-B boxes and CENP-A occupancy (Sullivan et al., 2011). In some cases (e.g. chromosomes 7 or 17), the two homologue chromosomes display equal CENP-A occupancy but at different HOR arrays (epiallelic status) (Maloney et al., 2012). Also, it has been demonstrated that inactive HOR arrays with low CENP-A occupancy have the capacity to trigger HAC formation (Maloney et al., 2012). The CENP-A self-assembly mechanism model implies that previous incorporated CENP-A is required to avoid the sliding of the centromere to a different chromosomal position.

Using the CENP-A$^{OFF/ON}$ system, we tested if *de novo* CENP-A deposition was slightly displaced ultimately leading to a different distribution at HOR arrays within the same centromeric locus (**Figure 16A**). CENP-A – HOR array occupancy was determined by CUT&RUN (Cleavage Under Targets and Release Using Nuclease) combined with high-throughput sequencing (Skene and Henikoff, 2017). Sequencing reads were mapped to the latest HORs reference assembly for centromeric sequences, as described previously (Dumont et al., 2020). In agreement with previous observations of CENP-A occupancy, in untreated cells, CENP-A localized mostly to a defined HOR, but it could also be found on different HORs of the same chromosome (Nechemia-Arbely et al., 2019; Sullivan et al., 2011). Following CENP-A depletion and reactivation, we found that the distribution of *de novo* CENP-A along the different HORs was maintained, as we did not observe any remarkable differences of CENP-A HOR array occupancy compared to the untreated condition (**Figure 16B-E**). These results were confirmed by line scan and by chromatin fiber techniques coupled with FISH (**Figure 16F-G**). Here, we determined CENP-A occupancy along the HOR array of chromosome X (DXZ1) and found that CENP-A re-occupies around 35% of the DXZ1 array within 48 hr (**Figure 16F**), similar to the control and in agreement with previous results (Maloney et al., 2012).

Altogether, we concluded that CENP-A is reloaded at the same HOR array even in the absence of previously deposited CENP-A and that (epi-)genetic mechanism(s) other than CENP-A are involved in maintaining centromere position.

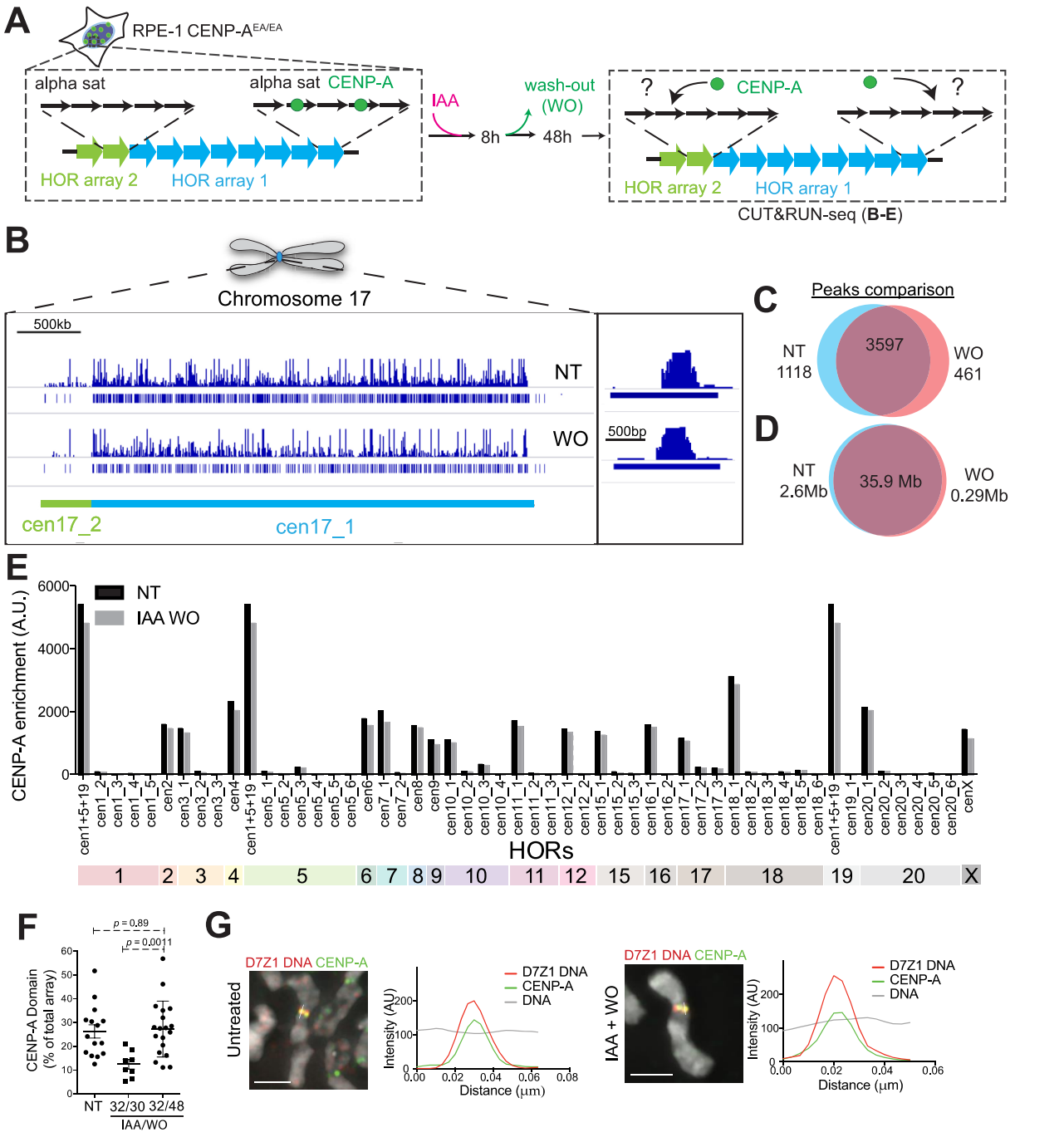


Figure 16. Previously deposited CENP-A is dispensable for precise de novo CENP-A incorporation. *(**A**) Schematic of DNA sequence organization at centromeres with more than one higher order repeat array (HOR) and experimental set-up of the CENP-A$^{OFF/ON}$ cycle performed in experiments B-D.*

*(**B**) Full coverage plot of chromosome 17 centromeric array (chr17:22,500,000-26,900,000 of the hg38 assembly) showing enrichment of CENP-A by CUT&RUN-seq in the untreated (NT) and IAA wash out (WO) sample. CENP-A is reloaded to cen17_1 (D17Z1). A shift of CENP-A occupancy to cen17_2 (D17Z1B) is not detected.*

*(**C**) Venn diagram showing the number of CUT&RUN-seq peaks that are NT specific (left, blue), WO specific (right, red) or overlapping (center).*

*(**D**): Venn diagram showing the total length in Mb of CUT&RUN-seq peaks that are NT specific (left, blue), WO specific (right, red) or shared between NT and WO (center).*

*(**E**) CENP-A levels at the indicated HOR arrays quantified by CUT&RUN-sequencing. CENP-A levels 48 h after IAA wash-out recover at the original HOR.*

*(**F**) Quantifications of CENP-A occupancy at the DXZ1 HOR array after one CENP-A$^{OFF/ON}$ cycle in DLD-1 cells using IF-FISH on chromatin fibers. Each dot represents a single chromatin fiber. Error bars show standard deviation. P-values from unpaired t test.*

*(**G**) Line scan analysis at the D7Z1 array on chromosome spreads in the indicated treatment. Scale bar, 2 µm.*

To further test the importance of the original centromere location, we then generated an *in vivo* competition assay between the native centromeres lacking endogenous CENP-A and ectopic site(s) enriched with exogenous CENP-A. To do so, we used an inducible CENP-A OE system with a binary (on/off) controllable activity in the CENP-A$^{OFF/ON}$ background. This binary control is achieved via a doxycycline-inducible expression of CENP-A tagged with a destabilization domain *E.coli*-derived *DiHydroFolate Reductase* (DHFR) protein (**Figure 17A**). Addition of a small ligand named TriMethoPrim (TMP) is required for protein stabilization (Iwamoto et al., 2010). We then assessed if *de novo* endogenous CENP-A reload at non-centromeric sites following mis-incorporation of exogenous CENP-A – which can be indistinguishable from centromeric CENP-A levels **(Figure 17F) -** and if this was sufficient to trigger ectopic centromere formation (**Figure 17B, C**). In most analyzed cases, centromere function (defined by the presence of CENP-C) was occurring at CENP-B marked centromeres and never observed at ectopic loci despite the initial presence of ectopic CENP-A along the chromosome arms (**Figure 17D, E**).

This data demonstrates the importance of alpha satellite DNAs and their embedded features in marking centromere position.

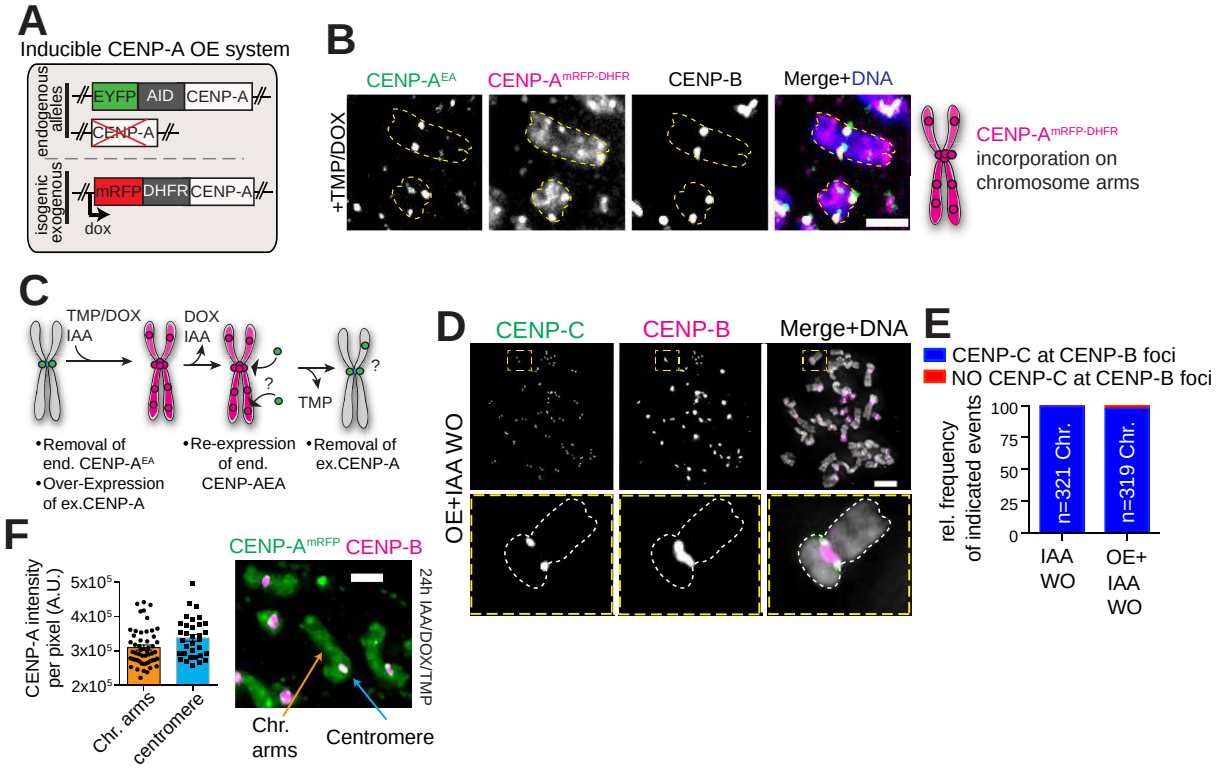


Figure 17: Re-expression of CENP-A into a temporary CENP-A overexpression background is insufficient to promote neocentromere formation.

*(**A**) Illustration of genomic make-up of DLD-1 with an exogenous CENP-A$^{mRFP\text{-}DHFR}$ overexpression system.*

*(**B**) Representative images showing ectopic CENP-A on a chromosome spread after induction of CENP-A overexpression in DLD-1 cells. Centromere position is marked using immunofluorescence staining for CENP-B. Yellow dashed lines contour representative chromosomes. Scale bar, 2µm. Schematic to the right illustrate the observed CENP-A$^{mRFP\text{-}DHFR}$ localization pattern on a chromosome.*

*(**C**) Schematic illustration for the experiments shown in D, E.*

*(**D**) Representative images of chromosome spreads after re-expression of endogenous CENP-A^{EA} into a CENP-A$^{mRFP\text{-}DHFR}$ overexpression background and subsequent removal of the overexpression as illustrated in (C). Antibody against CENP-C was used to score for the presence of the functional centromere. Scale bar, 5µm.*

*(**E**) Quantification of the indicated events after one CENP-A$^{OFF/ON}$ cycle compared to treatment shown in (C).*

*(**F**) Left panel: Comparison of CENP-A^{DHFR} level per pixel at the centromere and on chromosome arms following IAA and DOX/TMP treatment. Each dot represents one centromere or a region on the chromosome arms. Right panel: Immunofluorescence image of a chromosome spread of a cell overexpressing CENP-A^{DHFR} (green). CENP-B staining is shown in magenta. Scale bar 2µm.*

De novo CENP-A deposition is impaired in CENP-B deficient cells

We next asked what maintains centromere position in the absence of previously deposited CENP-A. We have already demonstrated that CENP-B plays a major role in stabilizing centromere proteins, including a fraction of CENP-C, on CENP-A-depleted centromeres (Hoffmann et al., 2016). We therefore hypothesized that CENP-B may be important for *de novo* CENP-A deposition. To test this, we assessed CENP-A *de novo* deposition in *CENP-B* KO RPE-1 cells harboring the CENP-A$^{OFF/ON}$ system (Hoffmann et al., 2016). Following CENP-A$^{OFF/ON}$ we found that *de novo* CENP-A reloading was strongly impaired in *CENP-B*$^{-/-}$ cells compared to *CENP-B*$^{+/+}$ cells, with less than 20% of the cells that reloaded CENP-A upon re-expression (**Figure 18A-C**).

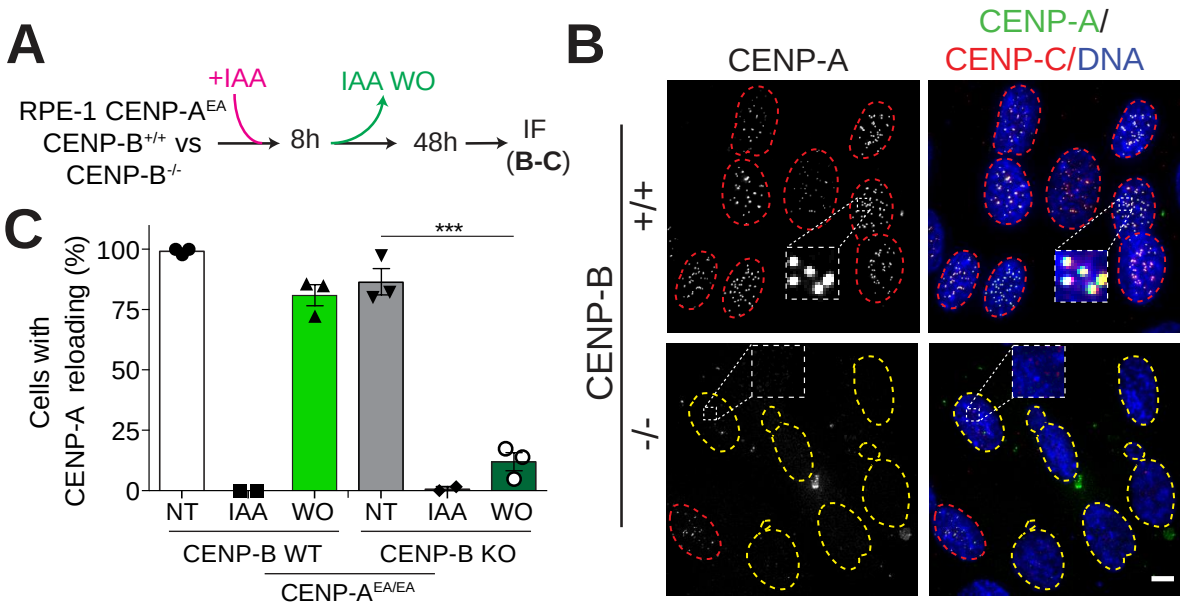


Figure 18. De novo CENP-A deposition is impaired in the absence of CENP-B.

*(**A**) Schematic illustration of the CENP-A$^{OFF/ON}$ cycle performed in the experiments shown in B, C.*

*(**B**) Representative images of de novo CENP-A reloading in CENP-B wild-type (+/+) and CENP-B knock-out (-/-) cells. Cells with centromeric CENP-A are marked with a red dashed contour line while a yellow contour lines mark cell without centromeric CENP-A. Scale bar, 5 µm.*

*(**C**) Quantification of relative number of RPE-1 cells with centromeric CENP-A in the indicated conditions. Each dot represents one experiment with at least 20 cells per condition. Error bars represent standard error of the mean (SEM) from 3 independent experiments. Unpaired t test, ***p=0.0003.*

Co-depletion of CENP-A and CENP-B leads to immediate mitotic defects (Hoffmann et al., 2016) that result in a p53-mediated cell cycle arrest. So, we first tested if massive chromosome mis-segregation *per se* has an effect on *de novo* CENP-A reloading by chemical inhibition of the spindle checkpoint kinase Mps1 (Santaguida and Musacchio, 2009). Despite massive chromosome mis-segregation in this condition, *de novo* CENP-A reloading was not affected (**Figure 19**).

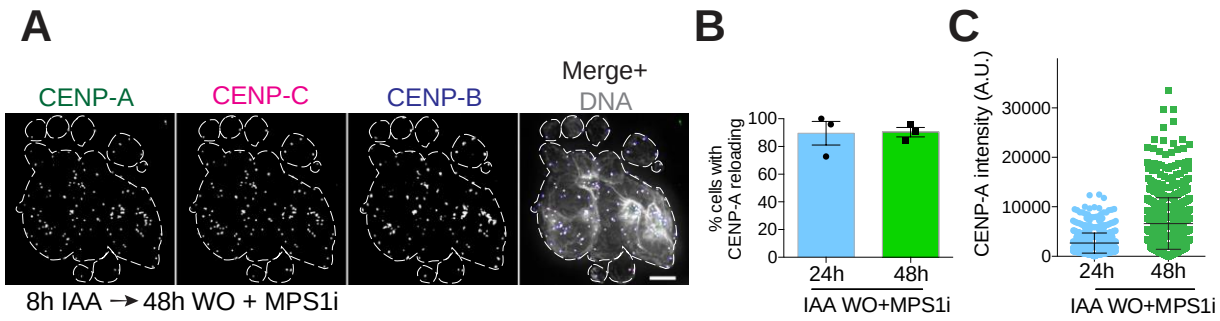


Figure 19: Reversine induced missegregation does not impair de novo CENP-A reloading

*(**A**) Representative immunofluorescence showing de novo CENP-A deposition in CENP-B (+/+) DLD-1 cells after reversine induced chromosome mis-segregation. DAPI staining is contoured by a white dashed line. Scale bar, 5 µm.*

*(**B-C**) Bar graphs showing the relative number of CENP-A positive cells (**B**) and the level of centromeric CENP-A level (**C**) in presence of reversine. Each dot represents one experiment with more than 20 cells per condition. Error bars represent SD of 3 independent experiments.*

To rule out that failure of *de novo* CENP-A deposition in *CENP-B* KO cells is not simply a consequence of cell cycle arrest, we depleted *CENP-B* using CRISPR in a p53-deficient DLD-1 cell line (**Figure 20A**). We then synchronized cells to specifically assess *de novo* CENP-A reloading uniquely in cells that underwent mitosis following the CENP-A$^{OFF/ON}$ pulse (**Figure 20B**). Using α-tubulin staining to visualize early G1 cells we found that ~90% of *CENP-B*$^{-/-}$ cells showed impaired *de novo* CENP-A reloading while almost all *CENP-B*$^{+/+}$ cells successfully reloaded CENP-A (**Figure**

20C, D), similar to control RPE-1 cells. Interestingly, ~ 50% of the *CENP-B*$^{-/-}$ cells failed to reload CENP-A completely, while around 40% of cells showed only partial reloading of CENP-A with less than 10 centromeric CENP-A dots per daughter cell (**Figure 20C-E**). We then tested if re-expression of CENP-B could rescue CENP-A *de novo* deposition in CENP-B KO cells. To this aim we integrated CENP-B into an isogenic FRT locus that can be induced by doxycycline addition. Following a CENP-A$^{OFF/ON}$ pulse, most cells rescued by CENP-B successfully reloaded CENP-A to a similar extent as the CENP-B wildtype cell line (**Figure 20F, G**).

Overall, our data demonstrate that CENP-B is required for efficient re-loading of *de novo* CENP-A in the absence of any previously deposited centromeric CENP-A.

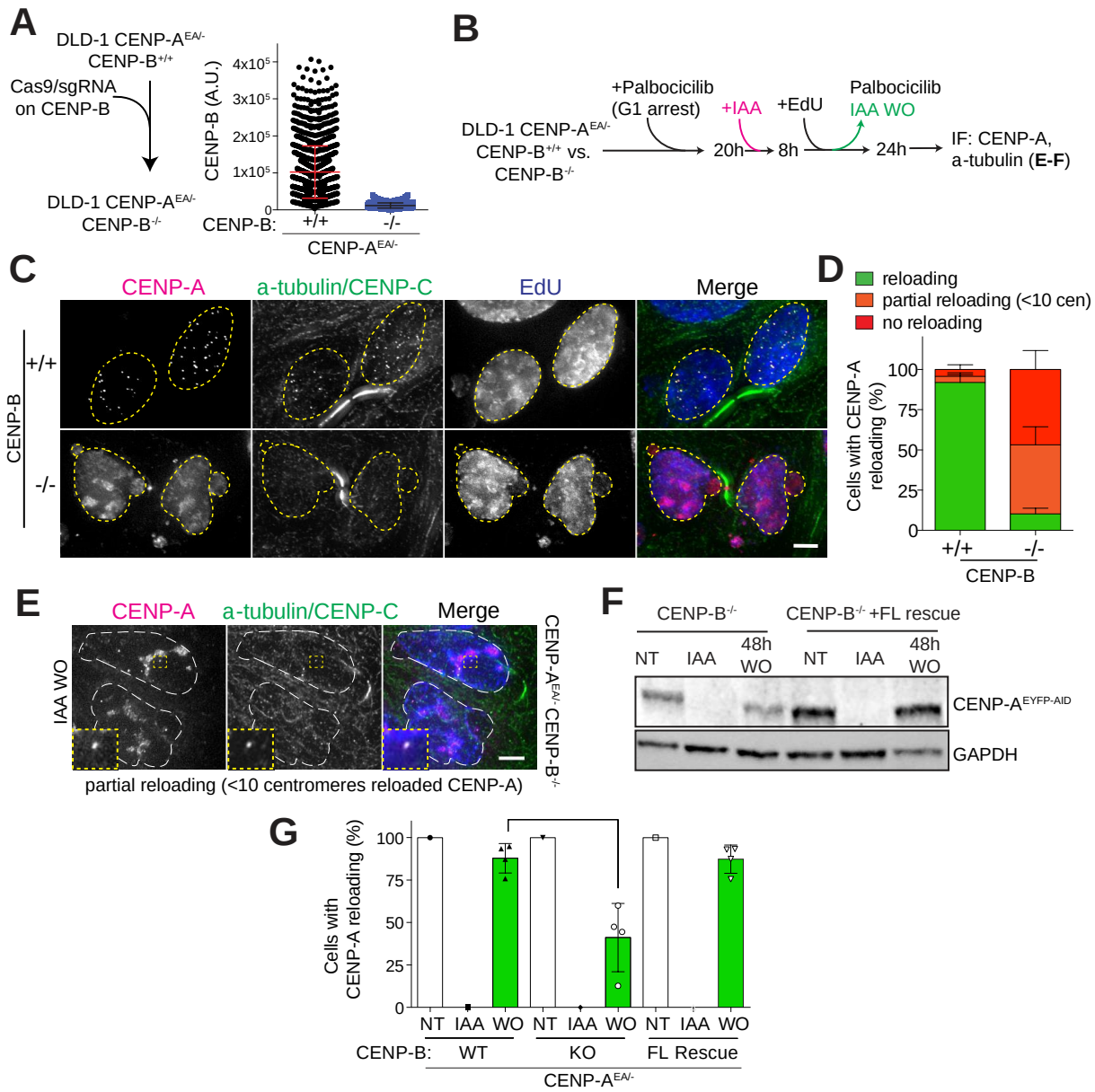


Figure 20: CENP-B is a key factor for efficient de novo CENP-A reloading in p53 deficient DLD-1 cells.

*(**A**) Left: schematic of the CRISPR/Cas9 strategy to deplete CENP-B in DLD-1 cells as measured in the dot plot on the right. Each dot represents one centromere, error bars show standard deviation.*

*(**B**) Schematic representation of experiments shown in C, D.*

*(**C**) Representative images of DLD-1 CENP-B (+/+) or (-/-) cells in late M phase following one CENP-A$^{OFF/ON}$ cycle. EdU staining was used to confirm successful wash-out of Palbociclib and cell progressing through S-phase. Yellow dashed lines contour nucleus of cells in late M phase. Scale bar 5 µm.*

*(**D**) Quantification of indicated events observed in late M phase cells in the indicated cell lines. Error bars show SEM from 4 independent experiments.*

*(**E**) Immunofluorescence images showing partial de novo CENP-A^{EA} reloading (<10 centromeres) in CENP-B$^{-/-}$ DLD-1 cells in late M phase. Nuclei of daughter cells are highlighted with white dashed lines. Scale bar, 5 µm.*

*(**F**) Immunoblot of total protein levels in the indicated conditions in the indicated cell line. FL = full length.*

*(**G**) Bar graphs showing the relative number of DLD-1 cells with centromeric CENP-A in indicated conditions in the indicated cell line. Each dot represents one single experiment, > 30 cells per condition. Unpaired t-test, **p=0.0053, error bars represent SD of 4 independent experiments.*

Neocentromere formation occurs on the Y chromosome under selective pressure

The centromere of the Y chromosome contains repetitive alpha satellite DNA sequences but lacks CENP-B binding sites. Hence, CENP-B is absent from the Y centromere (Earnshaw et al., 1989; Miga et al., 2014). To assess *de novo* CENP-A reloading at the Y centromere we used Fluorescence *In Situ* Hybridization (FISH) probes against the Y centromere or the X centromere, as a control, in combination with IF in interphase cells and on metaphase spreads in DLD-1 cells (**Figure 21A, B**). Following a CENP-A$^{OFF/ON}$ cycle, *de novo* CENP-A or CENP-C colocalized with the X centromere in both interphase and metaphase spreads, respectively, with no significant changes in abundance when compared to non-treated cells or to single allele AID-tagged CENP-A$^{AID/+}$cells (**Figure 21C-F**). Conversely, only ~25% of cells showed CENP-A/CENP-C at the Y centromere in

both interphase (**Figure 21E, F**) or mitotic (**Figure 21C, D**) cells. So, as observed in *CENP-B* KO cells, *de novo* CENP-A reloading is impaired at CENP-B lacking centromeres.

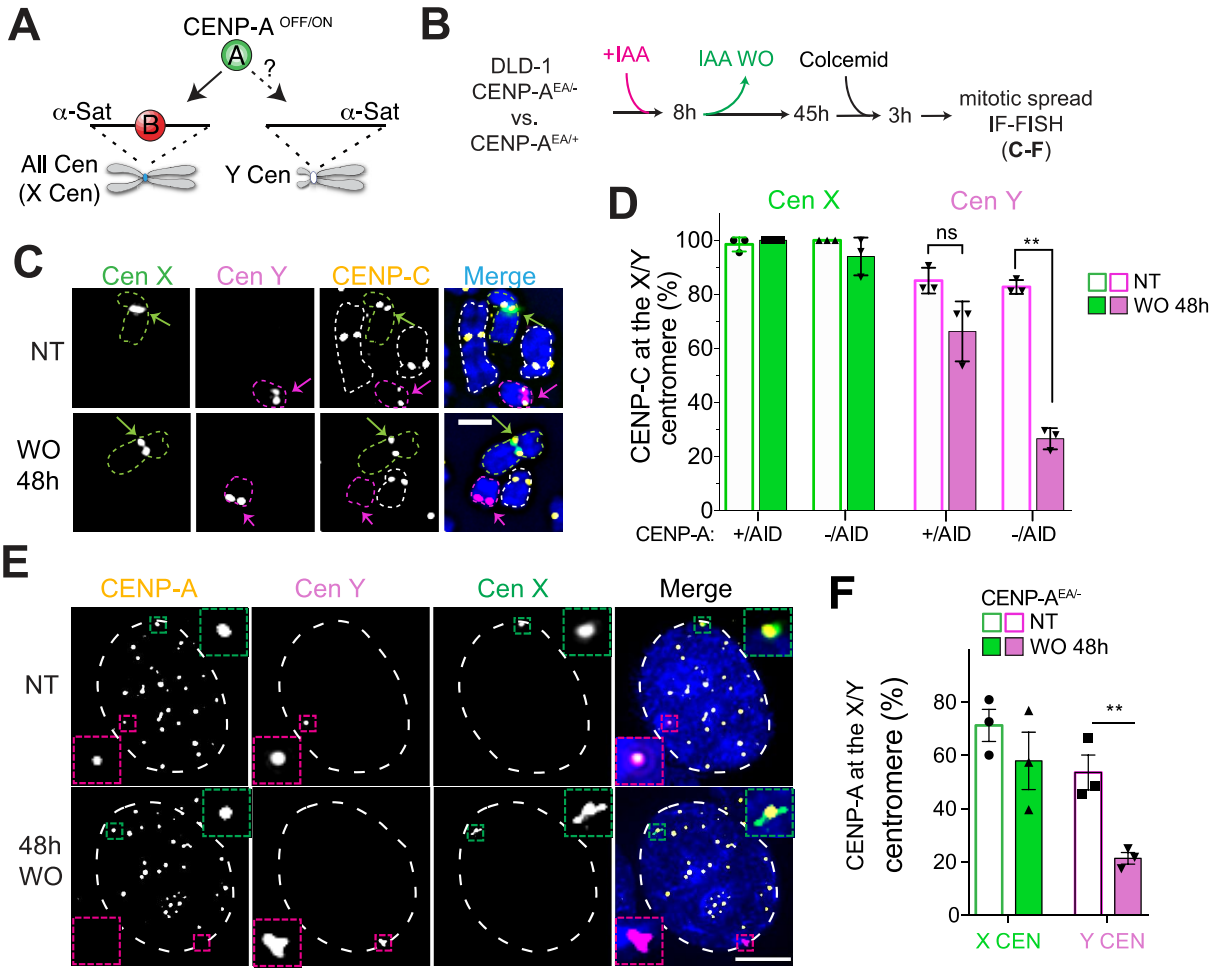


Figure 21: Impaired centromere formation at the Y chromosome after a CENP-A$^{OFF/ON}$ cycle.

*(**A**) Concept of experiments shown in B-D.*

*(**B**) Timeline of experiment performed in C, D.*

*(**C**) Representative IF-FISH images of mitotic spreads in non-treated cell (NT) and following a CENP-A$^{OFF/ON}$ cycle (IAA WO). Chromosomes are contoured by a green (X), magenta (Y) or white (other autosomes) dashed line. The centromeres of chromosome X/Y are highlighted by a green/magenta arrow. Scale bar, 2 µm.*

*(**D**) Quantification of CENP-C (as read-out of CENP-A) presence at the X (green) or Y (magenta) centromere in the indicated cell lines. Each dot represents one experiment. Error bars represent SD from 3 independent experiments. Unpaired t test, **p=0.0042.*

*(**E**) Representative IF-FISH images on interphase cells (nucleus highlighted by a white dashed line) showing de novo CENP-A deposition at the X but not at the Y centromere after a CENP-A$^{OFF/ON}$ cycle (48 h after IAA wash-out (WO). Scale bar, 5 µm.*

*(**F**) Bar plot showing frequency of CENP-A presence at the X (green) or Y (magenta) centromere in non-treated cells or following one CENP-A$^{OFF/ON}$ cycle. Unpaired t test, **p=0.0096. Error bars represent SEM of 3 independent experiments.*

We then studied whether CENP-A deposition can occur at a different, non-alpha-satellite location after CENP-A re-expression. However, the Y chromosome is not an essential chromosome in the *in vitro* cell culture model, and since neocentromere formation happens at very low frequency (Shang et al., 2013), the impaired CENP-A reloading is expected to lead to the loss of the Y chromosome in the majority of cells (Ly et al., 2017). To prevent its loss, we selected for the retention of the Y chromosome by inserting a selectable cassette (Neomycin) (**Figure 22A-C**). Following CENP-A depletion and re-activation, we subjected cells to G418 treatment to select clones that maintained the Y chromosome (Neomycin positive), as observed by colony formation assay (**Figure 22D-E)**. Following a CENP-A$^{OFF/ON}$ cycle, we observed a strong impact of G418 selection on cell viability that likely reflects failure of CENP-A deposition on the Y chromosome and subsequent loss of the chromosome bearing the selection marker. IF-FISH on the pool of surviving clones revealed three outcomes: i) in ~70% of the cells CENP-C was found to be at the original Y centromere location indicating that other centromere features different from CENP-A and CENP-B favor centromere formation at native centromere position (**Figure 22D, F-G)**; ii) In about 15% of the cells, the entire or portions of the Y chromosome (likely containing the Neomycin cassette) was fused to other chromosomes, an event that could be observed also in untreated condition, although at very low frequency (**Figure 22F-G**); iii) Intriguingly, in the about 15% of remaining cases CENP-C staining was observed elsewhere on the Y chromosome, and not coinciding with the Y centromeric probe, indicating the presence of a neocentromere (**Figure 22F-G**). Importantly, neocentromere-like elements were exclusively found after CENP-A depletion and re-activation by IAA WO, but never under untreated conditions (**Figure 22G**).

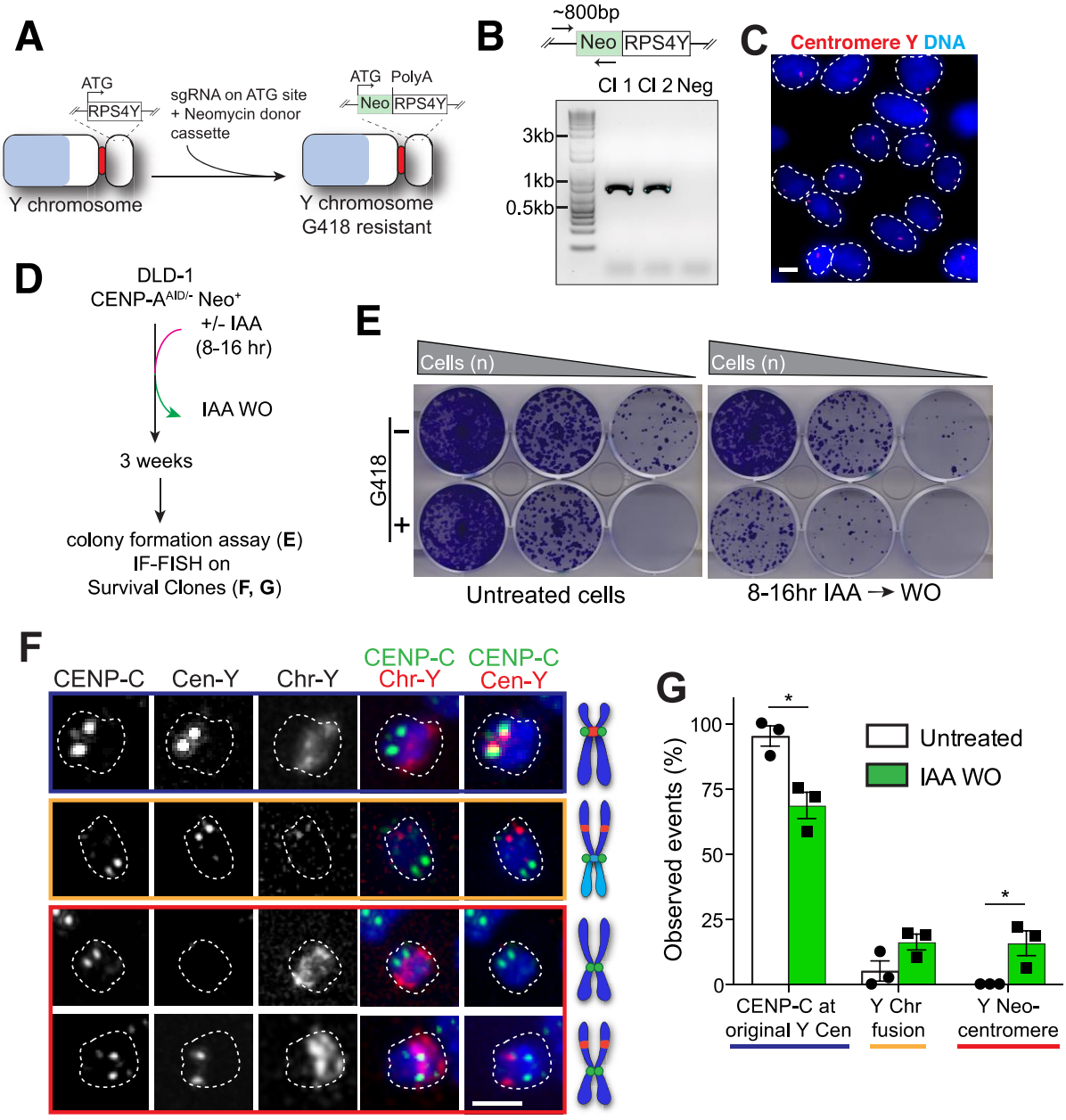


Figure 22. Neocentromere formation arises at the CENP-B negative Y-chromosome.

*(**A**) Schematic illustration of CRISPR/Cas9-mediated genome editing of the non-essential RPS4Y gene on the Y chromosome with the insertion of a neomycin resistance cassette.*

*(**B**) Agarose gel of a PCR run confirming successful integration of the Neomycin resistance cassette at the Y chromosome of DLD-1 cells.*

*(**C**) Representative FISH using a Y centromere probe in Neomycin selected DLD-1 cells (nuclei are highlighted by white dashed contour lines). Scale bar, 5 µm.*

*(**D**) Experimental set-up for E-G.*

*(**E**) Colony formation assay under selective pressure (Neomycin, G418) in untreated conditions or following a CENP-A$^{OFF/ON}$ cycle.*

*(**F**) Representative IF-FISH images showing CENP-C localization on the Y chromosome in cell growing under selective pressure (G418) following a CENP-A$^{OFF/ON}$ cycle. Chromosomes are highlighted by a white dashed contour line. CENP-C was found at the native centromere position (top panel, blue), the Y chromosome fused to another chromosome indicated by the increased size of the chromosome and the absence of Chr-Y painting probe staining (second panel, orange) or CENP-C was found on a different location on the Y chromosome and did not overlap with Cen-Y DNA (lower panels, red). Scale bar, 2 µm.*

*(**G**) Quantification of indicated events depicted in G. Each dot represents one experiment (>20 spreads for experiment). Error bars represent SEM from 3 independent experiments. Unpaired t-test, *p=0.0137 and 0.039.*

These data indicate that neocentromere-like elements can form on a CENP-B deficient chromosome following rapid CENP-A re-expression, but only when CENP-A-mediated centromere identity was transiently perturbed. Nevertheless, native centromere location continues to be the preferential site for centromere re-formation.

CENP-B bound to an ectopic location promotes CENP-C recruitment independently of CENP-A

We next investigated how CENP-B maintains centromere position at native centromeres. Previous studies showed that CENP-C interacts with CENP-B directly (Fachinetti et al., 2015; Suzuki et al., 2004), in addition to its well-known interaction with CENP-A (Fachinetti et al., 2013; Guse et al., 2011; Kato et al., 2013). Here, we noticed that exogenous CENP-B expression not only rescued *de novo* CENP-A reloading in CENP-B$^{-/-}$ cells (**Figure 23**), but also lead to increased CENP-C levels at the centromeres that were mostly maintained even after CENP-A depletion (**Figure 23**). These observations indicate that CENP-B not only maintains and stabilizes CENP-C (Hoffmann et al., 2016), but might also recruit CENP-C *de novo* to the centromere.

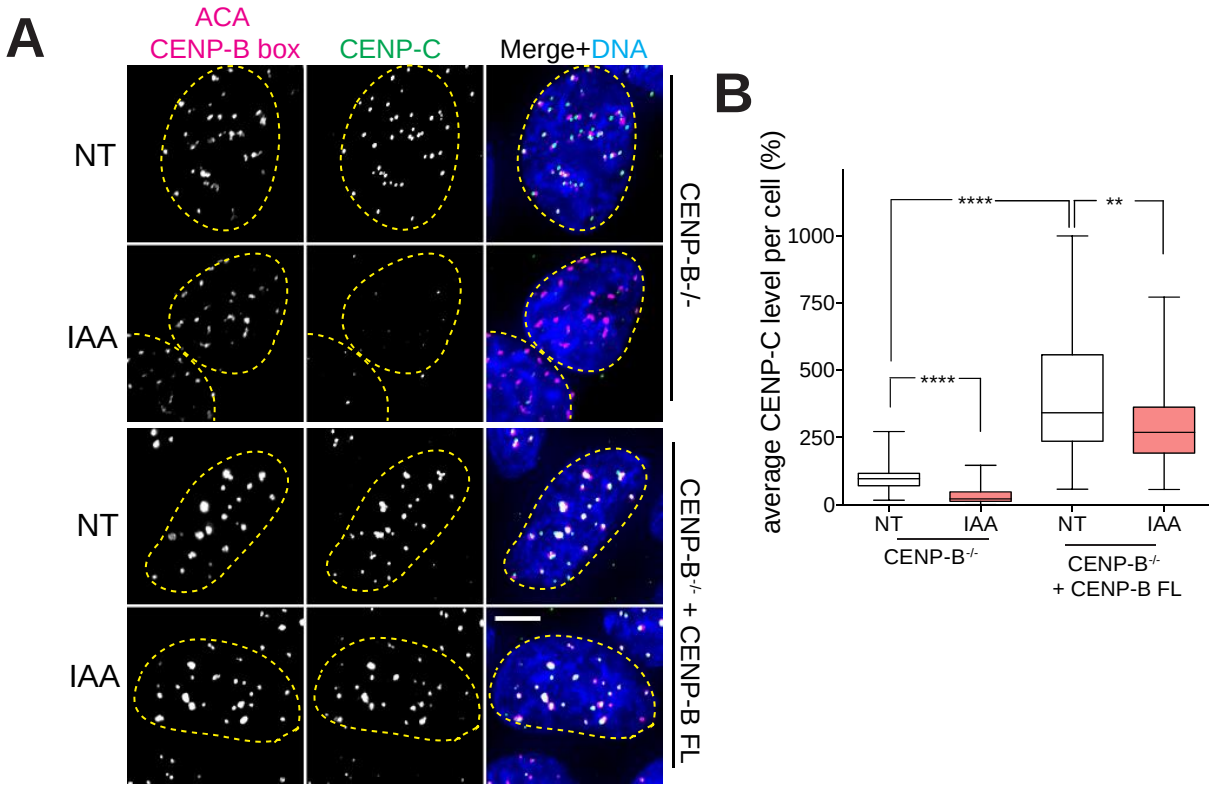


Figure 23: CENP-B overexpression leads to elevated CENP-C level at the centromere.

*(**A**) Representative IF-FISH images of non-treated and IAA treated (48 h) CENP-B$^{-/-}$ +/- CENP-B^{FL} rescue DLD-1 cells (nuclei are contoured by yellow dashed lines). Scale bar, 5 µm.*

*(**B**) Box plots of centromeric CENP-C levels in the indicated conditions (>70 centromeres per conditions). Box plot shows median and 25th and 75th percentiles, whiskers show minimum and maximum values. Unpaired t-test, ****p<0.0001, **p=0.0049.*

To test this hypothesis, we used the LacO-LacI system in a previously described U2-OS cell line (Janicki et al., 2004) in which we further integrated the CENP-A$^{OFF/ON}$ system at the CENP-A endogenous locus. This system allows us to test whether CENP-B can recruit CENP-C to an ectopic LacO locus independently of CENP-A (**Figure 24A**). Following transfection of CENP-B-LacI-mCherry, CENP-C recruitment to the ectopic big LacI/CENP-B sites was observed in the majority of the cells, even upon IAA treatment to deplete CENP-A, in contrast to the LacI/mCherry control (**Figure 24B-D**). Interestingly, CENP-C molecules at the LacO array were organized in patches rather than covering LacO sites homogenously. Based on previous studies (Hori et al., 2012; Shono et al., 2015), we predicted that once CENP-C is recruited to the LacO array it should recruit CENP-A. Indeed, following IAA WO, CENP-A was recruited at most CENP-B/CENP-C positive LacO

arrays (**Figure 24C-E**). CENP-A recruitment via CENP-C further stabilizes CENP-C, as CENP-C levels at the LacO array were higher in cells that had CENP-A (non-treated cells and in the IAA WO) compared to IAA treated cells (**Figure 24B-D**).

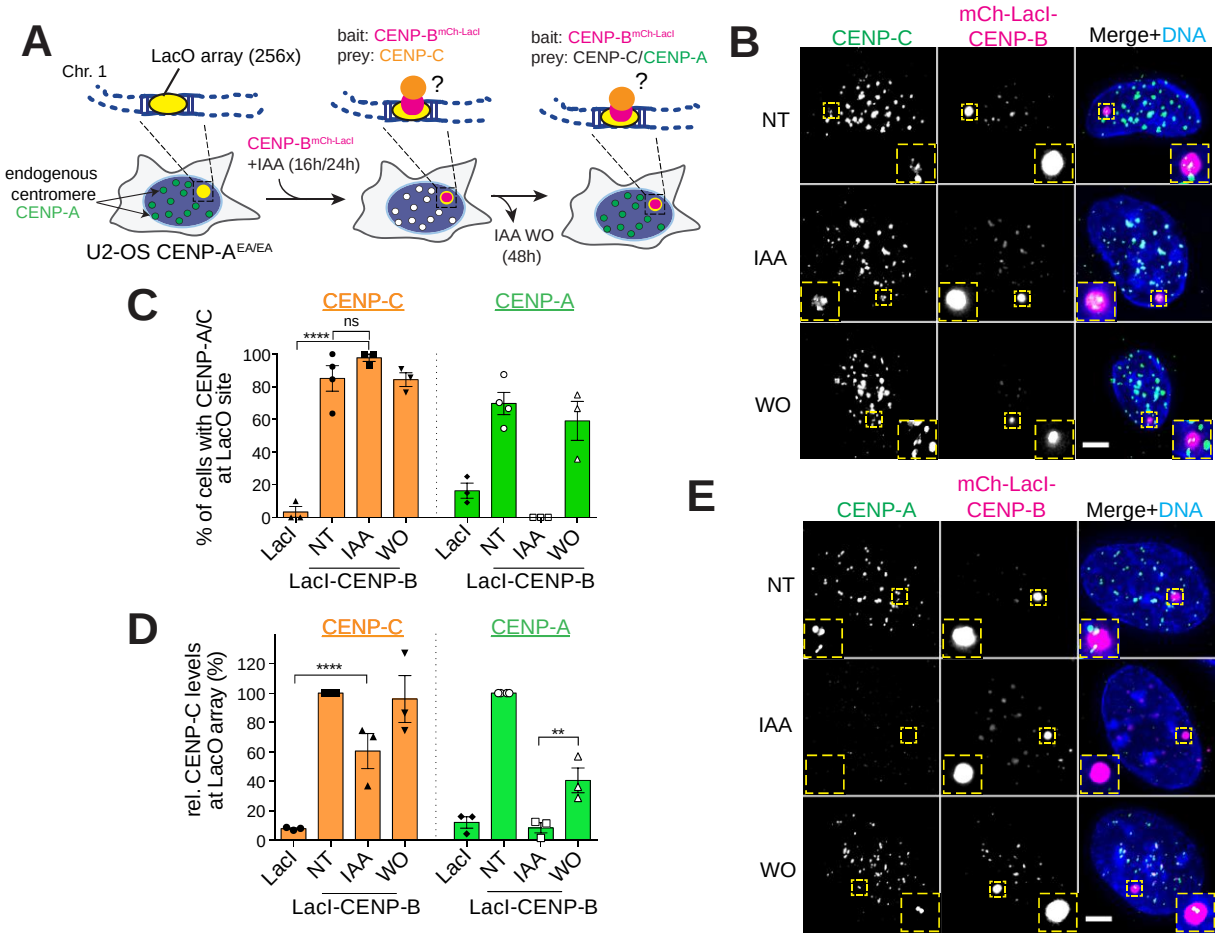


Figure 24. Ectopic CENP-B is capable to recruit CENP-C independently of CENP-A at LacO arrays.

*(**A**) Experimental schemes for B-G. LacI-CENP-B or LacI control (baits) were expressed into CENP-A$^{EA/-}$ lacO-tetO U-2OS cells. Immunostaining against CENP-C or CENP-A (preys) was performed following LacI-CENP-B transfection in untreated, IAA treated and IAA wash-out (WO) conditions.*

*(**B**) Representative immunofluorescence images showing CENP-C recruitment at the LacO array in the indicated conditions. LacO array is displayed in the insets. Scale bar, 5 µm.*

*(**C**) Bar plot showing frequency of CENP-C (orange) or CENP-A (green) recruitment to LacO array in the indicated conditions. Each dot represents one independent experiment (>20 cell analyzed for experiment). Error bars represent SEM from 3 independent experiments. Unpaired t test, ****p<0.0001.*

*(**D**) Quantification of CENP-C (orange) or CENP-A (green) protein levels at the LacO array in the indicated conditions normalized to protein intensities in non-treated conditions at the mCherry-LacI-CENP-B-LacO array using CENP-A or CENP-C antibody. Each dot represents one independent experiment (>20 cell analyzed for experiment). Error bars represent SEM from 3 independent experiments. Mann-Whitney test was performed on pooled single cell data of three independent experiments, ****p<0.0001.*

*(**E**) Representative immunofluorescence images showing CENP-A recruitment to ectopic mCherry-LacI-CENP-B in the indicated conditions. LacO array is shown in the inset. Scale bar, 5 µm.*

Efficient recruitment of CENP-A was dependent on CENP-C, as pre-removal of CENP-C by siRNA largely abolished CENP-A recruitment at CENP-B positive LacO arrays (**Figure 25**). Residual CENP-A recruitment in this condition could be due to incomplete CENP-C depletion by siRNA or to a weak ability of CENP-B to directly recruit CENP-A (Fujita et al., 2015).

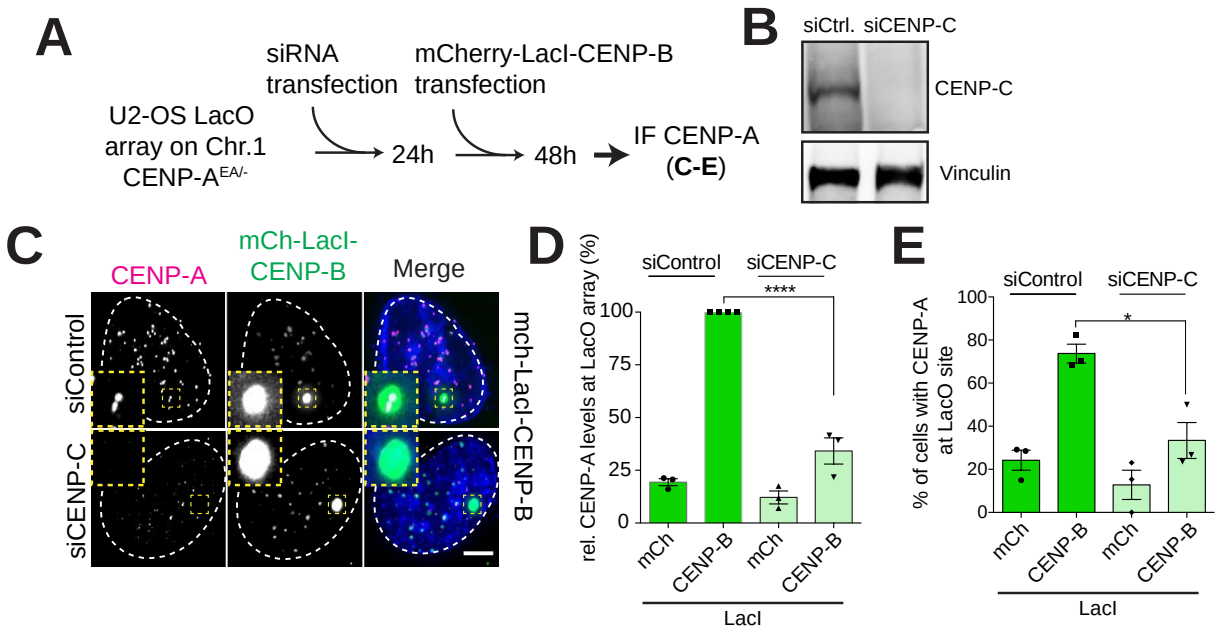


Figure 25: CENP-C is required for the recruitment of CENP-A at ectopically tethered CENP-B loci

*(**A**) Timeline of experiments shown in D-G.*

*(**B**) Immunoblot showing CENP-C knock-down after siRNA treatment.*

*(**C**) Representative immunofluorescence images showing CENP-B (bait) mediated CENP-A (prey) recruitment in the indicated conditions. Nuclei are marked by white dashed contour lines. Inset shows LacO array. Scale bar, 5µm.*

*(**D**) Bar graph showing CENP-A intensities at the LacO array in the indicated conditions normalized to CENP-A intensities in siRNA control conditions at the mCherry-LacI-CENP-B-LacO array. Each dot represents one independent experiment (>20 cell analyzed for experiment). Unpaired t test, ****p<0.0001, error bars show SD of 3 independent experiments.*

*(**E**) Frequency of CENP-A recruitment to the LacO array in the indicated conditions. Each dot represents one independent experiment. Unpaired t test, *p=0.0128, error bars show SD of 3 independent experiments.*

We next tested which CENP-B domains were involved in CENP-C recruitment. Previous yeast two-hybrid analysis suggested that CENP-B's acidic domain interacts with CENP-C (Suzuki et al., 2004). We transiently expressed several CENP-B constructs that lack either the DNA Binding Domain (DBD), the acidic domain or both. As positive and negative controls, we used CENP-B Full Length (FL) and H3.1 or CENP-T$^{\Delta C}$ (Gascoigne et al., 2011), respectively. As the DBD of CENP-B was shown to interact with CENP-A (Fujita et al., 2015; Suzuki et al., 2004), we performed these experiments in the absence of CENP-A (+IAA) to avoid any interference of CENP-A in CENP-C recruitment. In agreement with previous studies (Gascoigne et al., 2011; Hori et al., 2012), we could not observe CENP-C recruitment by H3.1 or ectopic CENP-T. Surprisingly, CENP-C was recruited at the CENP-B$^{\Delta acidic}$-LacI in a similar manner to that of FL or ΔDBD (**Figure 26A**). However, double deletion of both the DBD and the acidic domain almost completely prevented CENP-C recruitment to a level similar to that of CENP-T or the negative control H3.1 (**Figure 26A**). This remaining small fraction of CENP-C recruited to the LacO could be due to the presence of endogenous CENP-B that dimerize with the CENP-B variant. We confirmed these results using pull-down assays with purified recombinant proteins (**Figure 26B-C**). Here we observed direct interaction between GST-CENP-C (1-727) and $^{\Delta DBD}$CENP-B (**Figure 26D**). However, removal of the acidic domain in $^{\Delta DBD}$CENP-B abolished the interaction with CENP-C (**Figure 26D**). In summary, we concluded that both the acidic domain and the DBD of CENP-B are involved in CENP-A-independent recruitment and maintenance of CENP-C that, in principle, is capable to initiate the epigenetic centromere assembly loop mediated by CENP-A.

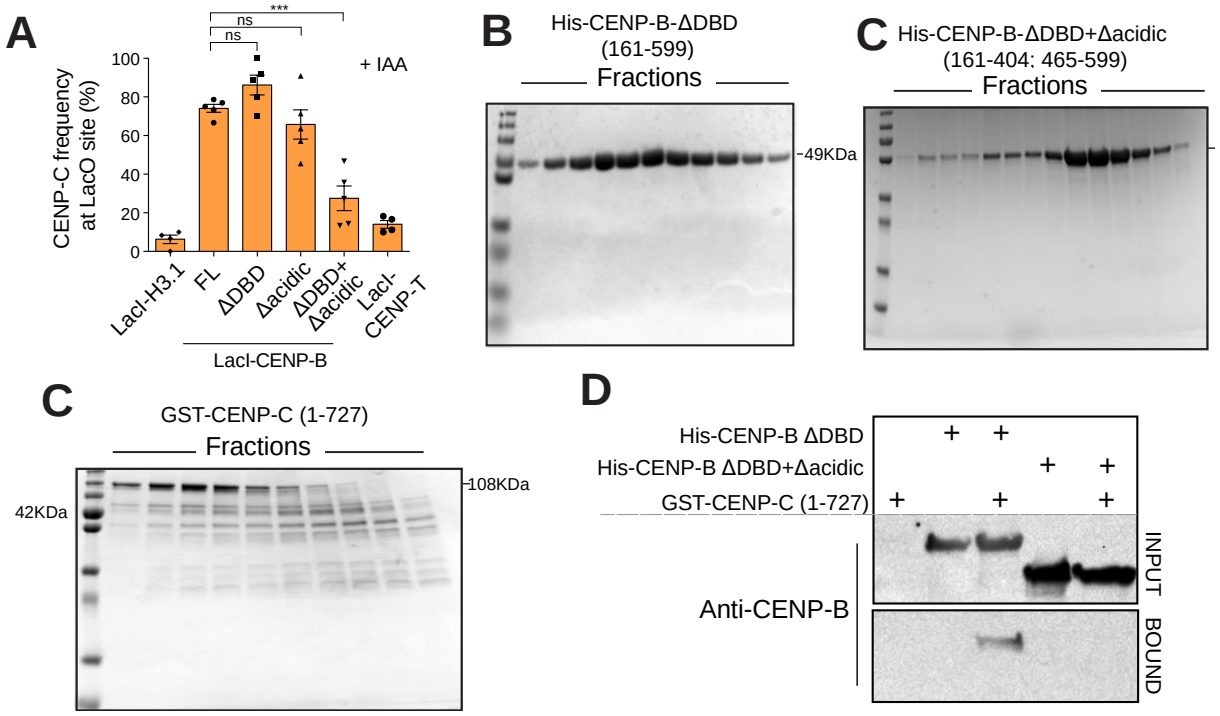


Figure 26: CENP-B's acidic and DNA binding domain (DBD) are involved in CENP-C binding.

*(**A**) Quantification of the frequency of CENP-C recruitment to LacO arrays by different LacI constructs. Error bars represent SEM from 5 independent experiments. Unpaired t test, ***p=0.0001.*

*(**B-C**) SDS PAGE analysis of fractions with an unique peak containing the protein of interest after a final purification step by size exclusion chromatography of the indicated constructs.*

*(**D**) CENP-B immunoblot following GST pulldown experiments using GST tagged CENP-C (1-727) as bait and with the indicated proteins as preys.*

CENP-B marks native centromere position to promote *de novo* CENP-A/C reloading

Given the strong enrichment of CENP-B-LacI at the LacO array, it was unclear if the interaction with CENP-C is relevant for new CENP-A deposition at endogenous centromeres. In addition, CENP-B-LacI at the LacO array could potentially cluster with endogenous centromeres and, to a certain extent, influence our observation of CENP-C recruitment. Therefore, we tested if CENP-B could promote centromere formation also at the native centromere position, where CENP-B levels are far lower compared to the LacI system. To test this, we generated a DLD-1 cell line in which

both endogenous CENP-A and CENP-C can be rapidly depleted (and re-expressed) simultaneously using the auxin inducible protein degradation system (CENP-A/C$^{OFF/ON}$; **Figure 27**).

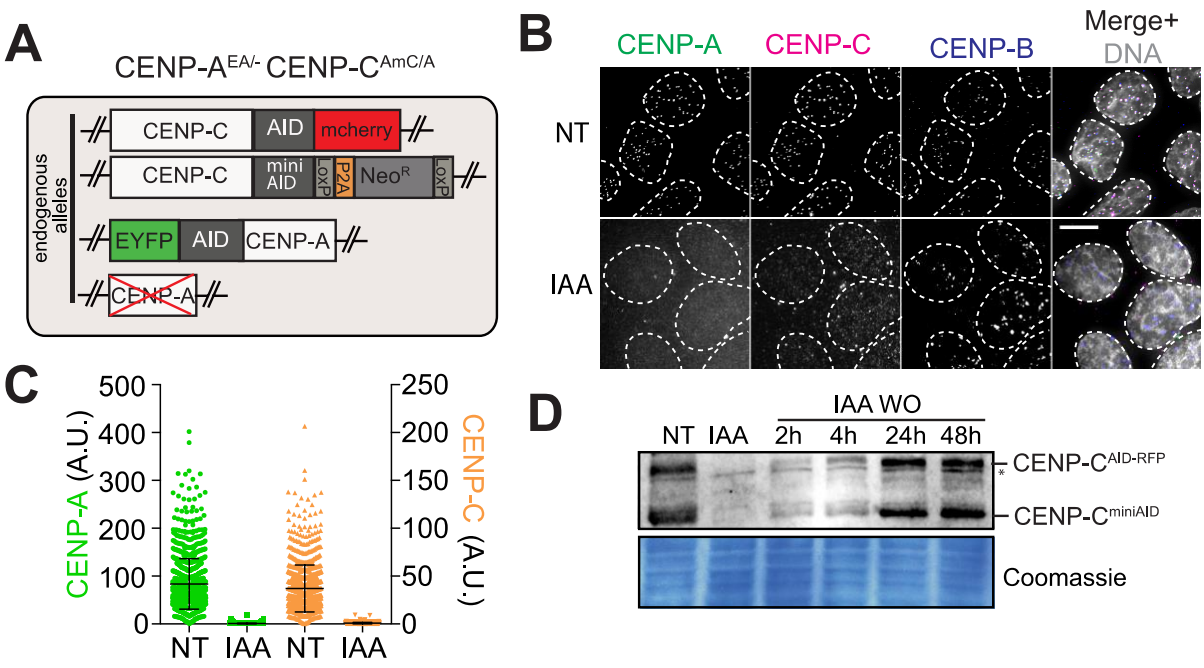


Figure 27: Development of a CENP-A/C$^{OFF/ON}$ AID system.

*(**A**) Schematic representation of the genomic make-up of DLD-1 cells used for CENP-A/C$^{OFF/ON}$ assays.*

*(**B**) Representative immunofluorescence images showing CENP-A and CENP-C depletion after IAA treatment. Nuclei are contoured by white dashed lines. Scale bar, 5 µm.*

*(**C**) Quantification of centromeric CENP-A (green) and CENP-C (orange) level in the indicated conditions in cells with genomic make-up shown in A. Each dot represents one centromere. Error bars show SD.*

*(**D**) Immunoblot following CENP-C depletion and re-expression in DLD-1 cells using the AID system. Asterisk marks an unspecific band.*

After CENP-A/C depletion and re-activation, we tested if both proteins were reloaded *de novo* to the centromere and if it was CENP-B-dependent (**Figure 28A, E**). Following CENP-A/C$^{OFF/ON}$ we observed reloading of both proteins at some CENP-B positive centromeres in ~30% of the cells (**Figure 28G, H**). CENP-A reloading under this condition was also observed by live cell imaging (**Figure 28B**) and CUT&RUN followed by qPCR (**Figure 28I**). We noticed that only a

fraction of centromeres per cell show efficient reloading, as also demonstrated by the low CENP-A/C centromeric protein levels upon reactivation (**Figure 28C**). When CENP-A/C reloading did not occur at all centromeres, there was a detrimental impact on cell viability, differently to what we observed in cells in which only CENP-A is removed and re-activated (**Figure 28D**).

Despite only a fraction of centromeres show CENP-A/C reloading, our data strongly indicate that the endogenous centromere position remains the hotspot for new centromere formation even in absence of CENP-A and CENP-C. We next tested if partial *de novo* CENP-A/C reloading was CENP-B dependent. Upon downregulation of CENP-B by siRNA (**Figure 28F**), *de novo* CENP-A/C reloading was almost completely lost, as observed by IF and CUT&RUN–qPCR (**Figure 28G-I**), reduced to a similar level as that obtained by downregulating M18BP1 expression via siRNA (**Figure 28G-H**).

Altogether, our results show that, occasionally, CENP-B can initiate CENP-A/C deposition to maintain centromere position along with the canonical CENP-A deposition machinery (at least M18BP1-mediated) in cells in which CENP-A/C were simultaneously co-depleted.

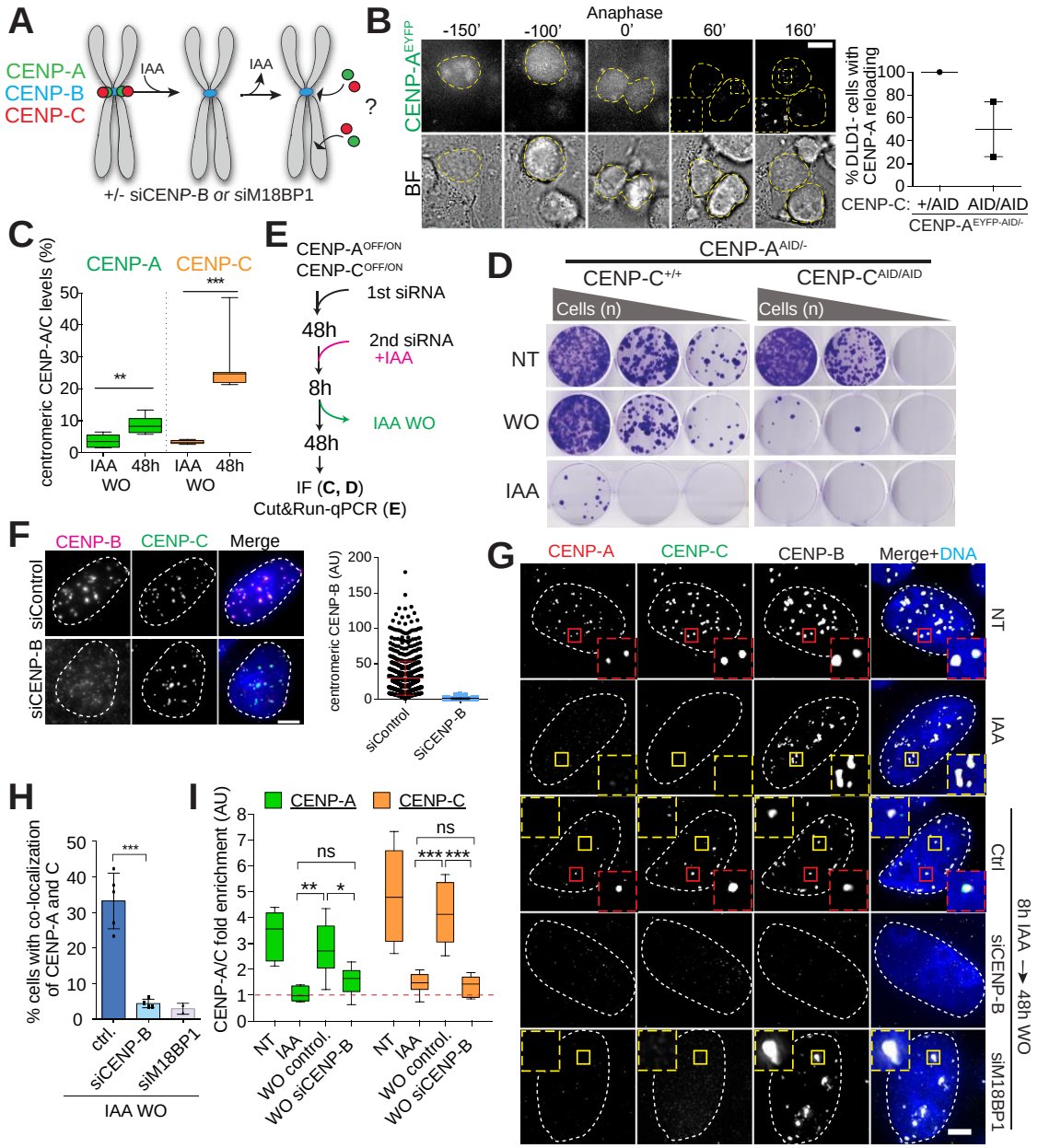


Figure 28. Partial de novo CENP-A and CENP-C reloading at CENP-A/C-depleted centromeres is dependent on CENP-B.

*(**A**) Schematic representation of experiments in B-E.*

***(B)** Left: stills of live cell imaging to monitor de novo CENP-A^{EA} reloading in DLD-1 cells after one CENP-A/C$^{OFF/ON}$ cycle. Images were taken every 10 min. Nucleus/mitotic cell is highlighted by yellow dashed contour line based on bright field (BF) images. Scale bar, 10 µm. Right: Plot to show the frequency of CENP-A reloading observed in cells that underwent mitosis. Each dot represents one experiment with at least 10 dividing cells. Error bars show SD of 2 independent experiments.*

*(**C**) Quantification of centromeric CENP-A (green) and CENP-C (orange) levels normalized to non-treated level using CENP-A/EYFP or CENP-C antibody. Box plot shows median and 25th and 75th percentiles, whiskers show minimum and maximum values. Average centromeric CENP-A/C intensities from five independent experiments (with at least 30 cells per condition) were used to generate the box plots. Unpaired t test: **p=0.0048, ***p=0.0003.*

*(**D**) Colony formation assay to assess cell viability in the indicated conditions and cell lines.*

*(**E**) Time line of experiment in G-I.*

*(**F**) Left: representative immunofluorescence image showing CENP-B knock-down during a CENP-A/C$^{OFF/ON}$ assay. White dashed lines highlight nuclei. Scale bar, 5 µm. Right: quantification of centromeric CENP-B level after RNAi treatment. Each dot represents one centromere (n>700 centromeres). Error bars show standard deviation.*

*(**G**) Representative immunofluorescence images showing CENP-A and CENP-C reloading in the indicated conditions. Cells with centromeric CENP-A/CENP-C are marked with a red dashed contour line while a yellow contour lines mark cell without centromeric CENP-A. The centromere position is marked using immunofluorescence staining for CENP-B. White dashed lines contour DAPI stained nucleus of representative cells. Scale bar, 5 µm.*

*(**H**) Bar plot showing the percentage of cells with CENP-A and CENP-C co-localization after a CENP-A/C$^{OFF/ON}$ cycle in the indicated conditions. Error bars show SD from 5 independent experiments. Unpaired t test, ***p=0.0002. Each dot represents one experiment (>30 cells analyzed for experiment).*

*(**I**) Box plot of CUT&RUN qPCR for CENP-A (green) or CENP-C (orange). Box plot shows median, 25th and 75th percentiles, whiskers show minimum and maximum values. Results from 3 independent experiments including combined data from qPCR reactions using alpha-satellite primers and primers binding at the centromere of chromosome 4. Enrichment is measured relative to the IgG control and normalized to Alu repeats. Mann-Whitney test, CENP-A: **p= 0.0087, *pp= 0.0411, CENP-C: **p= 0.0022.*

CENP-B initiates the CENP-A self-assembly epigenetic loop by recruiting CENP-C at native human centromeres

We next wanted to identify the mechanism by which CENP-B promotes *de novo* CENP-A/C reloading at the centromere. Our results imply a model in which CENP-B drives the initiation of the epigenetic centromere assembly loop. This can be achieved either by direct recruitment of CENP-A via CENP-B-DBD (Fujita et al., 2015; Okada et al., 2007), or via CENP-C that in turn interacts with components of the Mis18 complex (Moree et al., 2011; Stellfox et al., 2016). Alternatively, CENP-B could directly promote the recruitment of the Mis18 complex (**Figure 29A**), although previous evidence argued that the M18BP1 complex is insufficient to recruit endogenous CENP-A at an ectopic site (Shono et al., 2015). To test this hypothesis, we measured the number of endogenous M18BP1 centromere foci in early G1 RPE-1 cells following rapid depletion of CENP-A, CENP-C or co-deletion of CENP-A and CENP-B (**Figure 29B**). CENP-A removal or depletion of CENP-B alone did not alter M18BP1 recruitment at centromeric regions, while CENP-A/CENP-B co-depletion led to a reduction on the total number of M18BP1 foci (**Figure 29B-D**). Rapid and complete removal of CENP-C even further perturbed M18BP1 foci at most centromeres (**Figure 29C, D**). Since CENP-C depletion showed the most drastic effect, and co-depletion of CENP-A/CENP-B lead to a strong, although not complete, loss of centromeric CENP-C signal (Hoffmann et al., 2016), we favor the model that CENP-C, but not CENP-B, promotes M18BP1 recruitment, as previously observed (Dambacher et al., 2012; Moree et al., 2011).

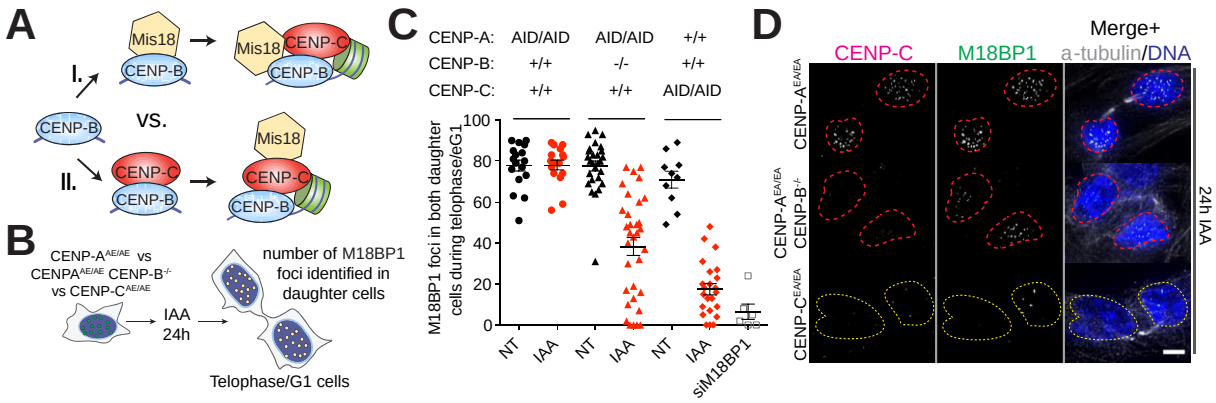


Figure 29: M18BP1 recruitment is CENP-C dependent.

*(**A**) Models of CENP-B induced CENP-A reloading via (I.) initial Mis18 recruitment or (II.) initial CENP-C recruitment.*

*(**B**) Schematic of the experiment analyzed in C, D.*

*(**C**) Quantification of M18BP1 foci in late M-phase/early G1 (eG1) in both daughter cells in the indicated conditions. Each dot represents two daughter cells. Error bars show SEM. Unpaired t test: ****p<0.0001.*

*(**D**) Representative immunofluorescence images showing M18BP1 foci in different cell lines after 24 hr IAA treatment. Cells with centromeric CENP-A are marked with a red dashed contour line while a yellow contour line marks cell without centromeric CENP-A. Scale bar, 5 µm.*

Our data emphasize that endogenous centromeric CENP-B can initiate the epigenetic self-assembly loop via CENP-C recruitment independently of CENP-A. Hence, we next dissected the temporal events that control centromere formation at endogenous CENP-B positive centromeres (**Figure 30A**). To do this, we developed a system to rapidly co-deplete CENP-A and CENP-C and separately induce the expression of either exogenous (e) CENP-A or CENP-C tagged with mRFP and a DHFR destabilization domain (**Figure 30B-C**). Expression of ectopic eCENP-A or eCENP-C was placed under the control of a doxycycline-regulatable promoter and stabilized by addition of TMP molecule (**Figure 30C-D**).

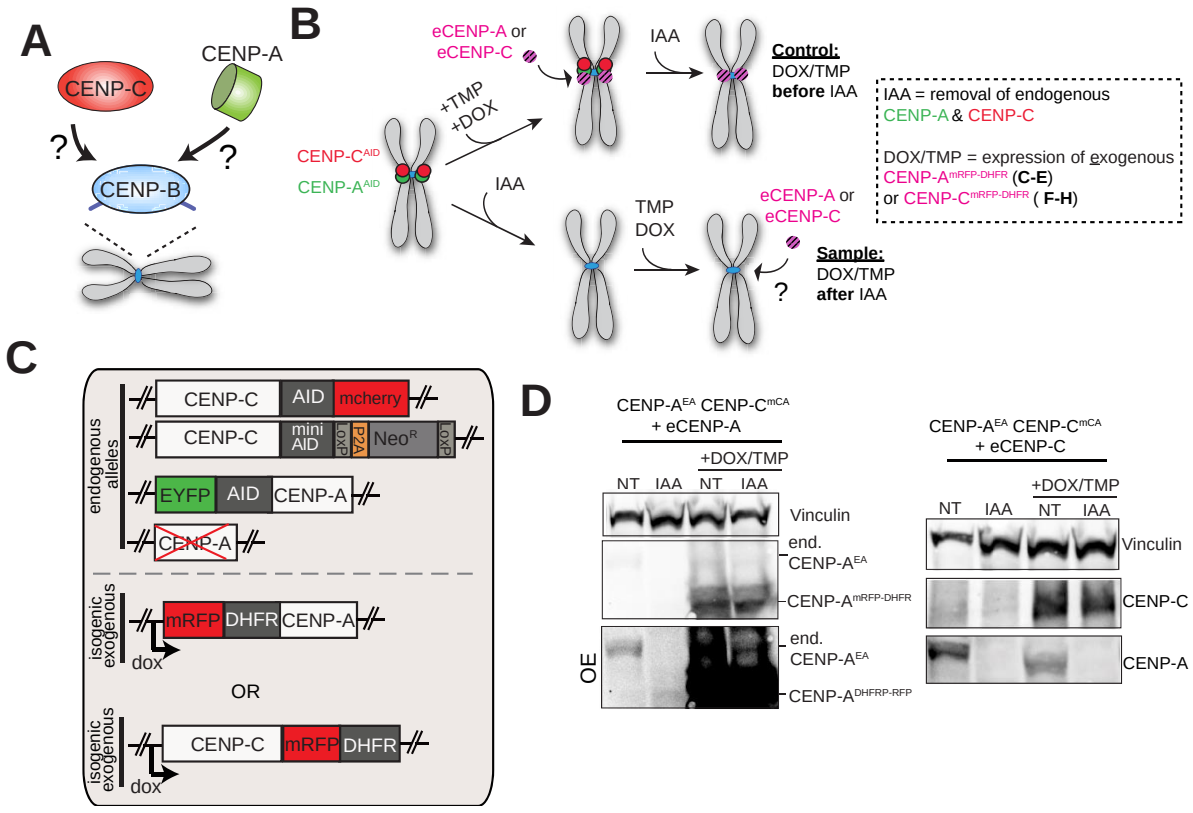


Figure 30: Exogenous CENP-A^{DHFR} or CENP-C^{DHFR} system to study chronology of partial CENP-A/C loading.

*(**A**) Schematic to illustrate the goal of this experiment.*

*(**B**) Schematic illustration of the experiments performed in Figure 31 and 32.*

*(**C**) Illustration of the genomic make-up of DLD-1 cells used to test CENP-A or CENP-C reloading in absence of endogenous CENPA/C.*

*(**D**) Immunoblot showing exogenous CENP-A (left) and CENP-C (right) expression upon addition of doxycycline (DOX) and TMP to induce CENP-A$^{DHFR\text{-}mRFP}$ or CENP-C$^{DHFR\text{-}mRFP}$ overexpression in DLD-1 cells.*

We first tested if eCENP-A can be loaded at the CENP-B-positive centromeres that lack preexisting CENP-A and CENP-C. We synchronized cells to be able to monitor eCENP-A loading in G1 prior to or after endogenous CENP-A/C removal by IAA (**Figure 31A**). Surprisingly, we did not observe any eCENP-A loading when endogenous CENP-A/CENP-C were removed prior to its deposition, in contrast to control cells in which CENP-A/C removal was performed after eCENP-A loading (**Figure 31D, E**). The absence of eCENP-A loading to centromeric regions depleted of CENP-A/C was also confirmed by CUT&RUN-qPCR (**Figure 31B-C**). This lack of signal was only partly explained by CENP-A reduced stability due to the absence of CENP-C (Falk et al., 2015; 2016), as even prolonged IAA treatment (24 hr) in control cells to remove endogenous CENP-A/C did not abolish the enrichment of eCENP-A at centromeric regions (**Figure 31F**).

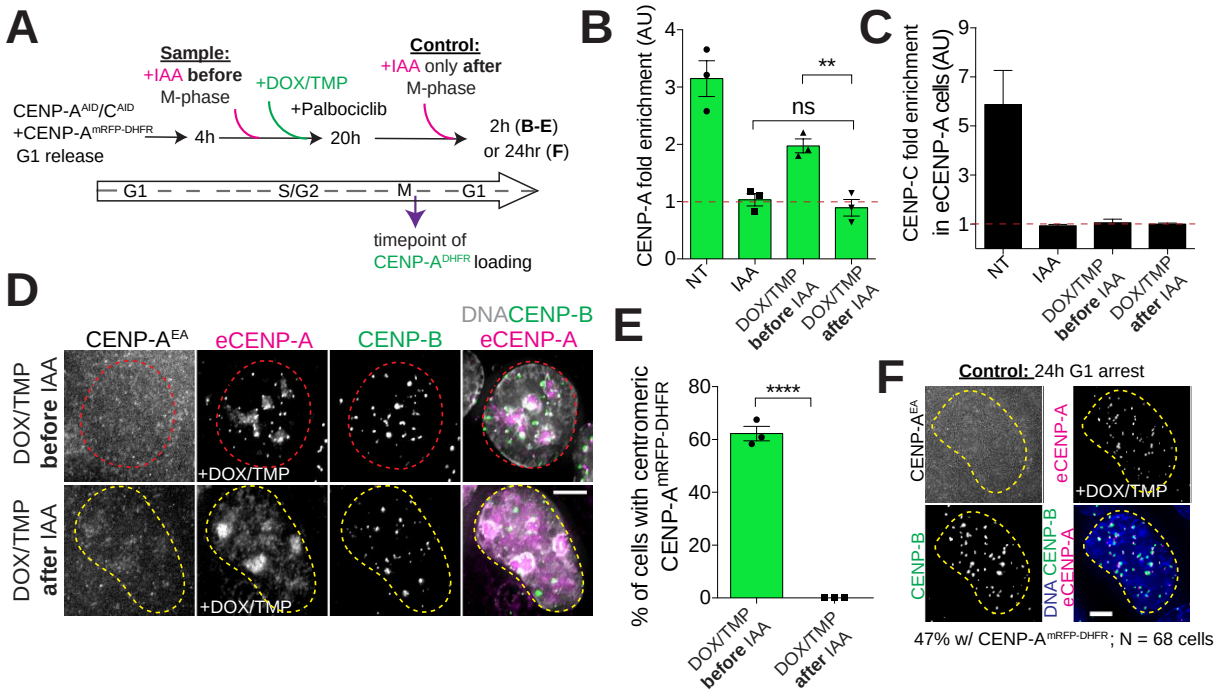


Figure 31: Exogenously expressed CENP-A does not reloaded at the original centromere position in absence of endogenous CENP-A/C.

*(**A**) Experimental design of experiments shown in B-F.*

*(**B**) Bar plot of CUT&RUN qPCR quantification using CENP-A antibody and primers binding at the centromere of chromosome 4. Enrichment is measured relative to the IgG control and normalized to Alu repeats. Error bars represent SEM from 3 independent experiments. Each dot represents one independent experiment. Unpaired t test, **p=0.0048.*

*(**C**) Bar graphs showing the control of CUT&RUN qPCR on CENP-C antibody and primers binding to the centromere of chromosome 4 in the indicated conditions. Enrichment is measured relative to the IgG control and normalized to Alu repeats. Error bars represent of SD of 3 independent experiments.*

*(**D**) Representative immunofluorescence images showing exogenous CENP-A$^{mRFP-DHFR}$ reloading in the presence (control), but not in the absence (sample) of endogenous CENP-A and CENP-C. Scale bar; 5 µm. Cells with centromeric CENP-A are marked with a red dashed contour line while a yellow contour line marks cell without centromeric CENP-A.*

*(**E**) Quantification of relative number of cells showing centromeric eCENP-A$^{mRFP-DHFR}$ in the indicated conditions. Error bars represent SEM from 3 independent experiments. Each dot*

*represents one independent experiment with at least 30 cells per condition. Unpaired t test, ****p<0.0001.*

*(**F**) Representative immunofluorescence image showing 24 h CENP-A$^{mRFP\text{-}DHFR}$ stability in the absence of endogenous CENP-A and CENP-C in G1 arrested cells.*

Next, we tested the capacity of eCENP-C to be reloaded at the centromere in absence of centromeric CENP-A/C. Here, we used asynchronous cells (**Figure 32A**) as CENP-C loading is not restricted to early G1 (Hoffmann et al., 2016). In contrast to eCENP-A, we found that initial depletion of endogenous CENP-A/C did not prevent eCENP-C loading at CENP-B-positive centromeres in ~ 60% of interphase or mitotic cells (**Figure 32B-C, F-G)**. As previously noted (**Figure 28**), we observed only partial centromeric eCENP-C loading (**Figure 32G**). We confirmed the ability of eCENP-C to reload at the centromere by CUT&RUN-qPCR in the absence of CENP-A reloading (**Figure 32D-E**).

Altogether, our results show that CENP-C, but not CENP-A, can be loaded, at least partially, at centromeres that lack previously deposited CENP-A/CENP-C.

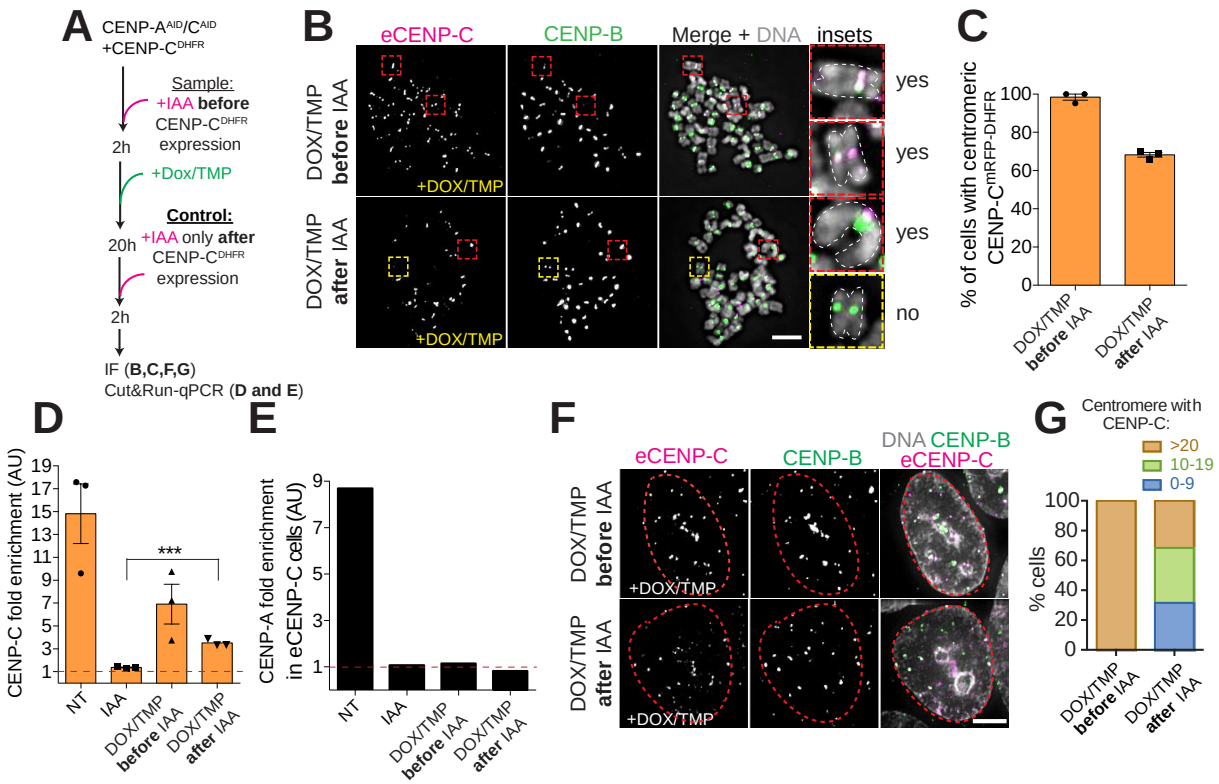


Figure 32: Exogenously expressed CENP-C reloaded partially at the original centromere position in absence of endogenous CENP-A/C.

*(**A**) Experimental design of experiments shown in B-G.*

*(**B**) Representative immunofluorescence chromosome spreads showing reloading of exogenous eCENP-C$^{mRFP\text{-}DHFR}$ in the presence (control) and in the absence (sample) of endogenous CENP-A and CENP-C. Cells were arrested with Colcemid for 3 h prior to spread. Cells with centromeric CENP-C are marked with a red dashed contour line while a yellow contour line marks cell without centromeric CENP-C. Chromosomes in inset are highlighted using a white dashed line. Scale bar, 5 µm.*

*(**C**) Quantification of relative number of cells with eCENP-C$^{mRFP\text{-}DHFR}$ in the indicated conditions. Each dot represents one independent experiment with at least 30 cells per condition. Error bars represent SEM from 3 independent experiments.*

*(**D**) Bar plot of CUT&RUN qPCR results using CENP-C antibody and primers binding at the centromere of chromosome 4. Enrichment is measured relative to the IgG control and normalized to Alu repeats. Each dot represents one independent experiment. Error bars represent SEM from 3 independent experiments. Unpaired t test, ***p=0.0003.*

*(**E**) Bar graph showing the control of CUT&RUN qPCR on CENP-A antibody and primers binding to the centromere of chromosome 4 in the indicated conditions. Enrichment is measured relative to the IgG control and normalized to Alu repeats.*

*(**F**) Representative immunofluorescence image showing CENP-C reloading in interphase cells (nuclei are highlighted by a dashed red contour line) in presence (control) or absence (sample) of endogenous CENP-A/C. Scale bar, 5 µm.*

*(**G**) Quantification of relative number of exogenous CENP-C-positive centromeres per cell in the indicated conditions quantified on chromosome spreads.*

CENP-A-negative resting CD4^{+} T cells re-assemble CENP-A *de novo* upon cell cycle entry

Our results imply that CENP-B provides memory of centromere identity. We next aimed to assess the physiological relevance of this mechanism. While dividing cells maintain CENP-A expression, terminally differentiated non-dividing cells lacking CENP-A have been described

(Swartz et al., 2019). Subsets of differentiated lymphocytes are quiescent in the peripheral blood from healthy individual, and they can re-enter the cell cycle upon activation through T Cell Receptor (TCR) activation. We hypothesized that a sub-population of circulating T lymphocytes may contain a fraction of CENP-A-negative cells, and they might re-acquire CENP-A expression and centromere deposition upon activation. We measured the level of CENP-A and CENP-B in resting, non-dividing $CD14^{-}/CD4^{+}$ T cells isolated from peripheral blood of healthy human donors (**Figure 33A**). While all the analyzed $CD4^{+}$ T cells were CENP-B-positive, two distinct populations were detected based on their CENP-A expression, referred to as CENP-A^{high} and CENP-A^{low} (**Figure 33B**).

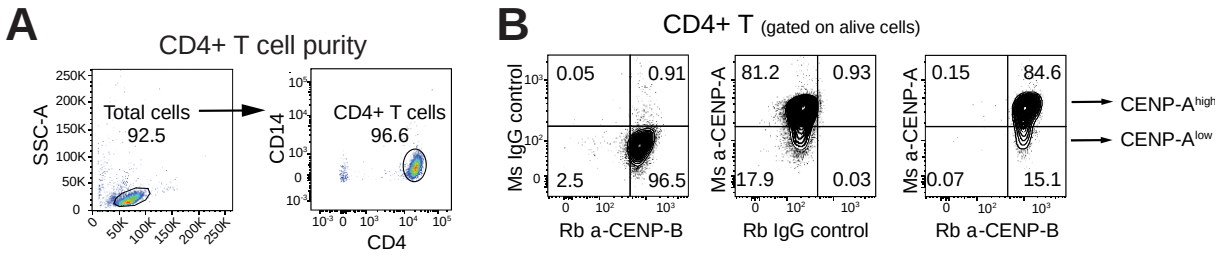


Figure 33: Identification of CENP-B positive but CENP-A deprived viable $CD4^{+}$ T in human blood samples.

*(**A**) Representative FACS plots showing the efficiency of total CD4+ T cell purification from human blood PBMCs.*

*(**B**) Representative FACS plots showing gating of CENP-B positive/CENP-A-high and CENP-B positive/CENP-A-low populations of freshly purified human CD4+ T cells based on isotype controls. Ms = mouse; Rb = rabbit.*

Next, we sorted CENP-A^{high} and CENP-A^{low} $CD4^{+}$ T cells to investigate CENP-A localization. While in CENP-A^{high} cells CENP-A colocalized with centromeric CENP-B, in most CENP-A^{low} cells CENP-A was undetectable or barely detectable at CENP-B-marked centromeres (**Figure 34**). In contrast, CENP-C localized at these CENP-A-depleted centromeres (**Figure 34E**).

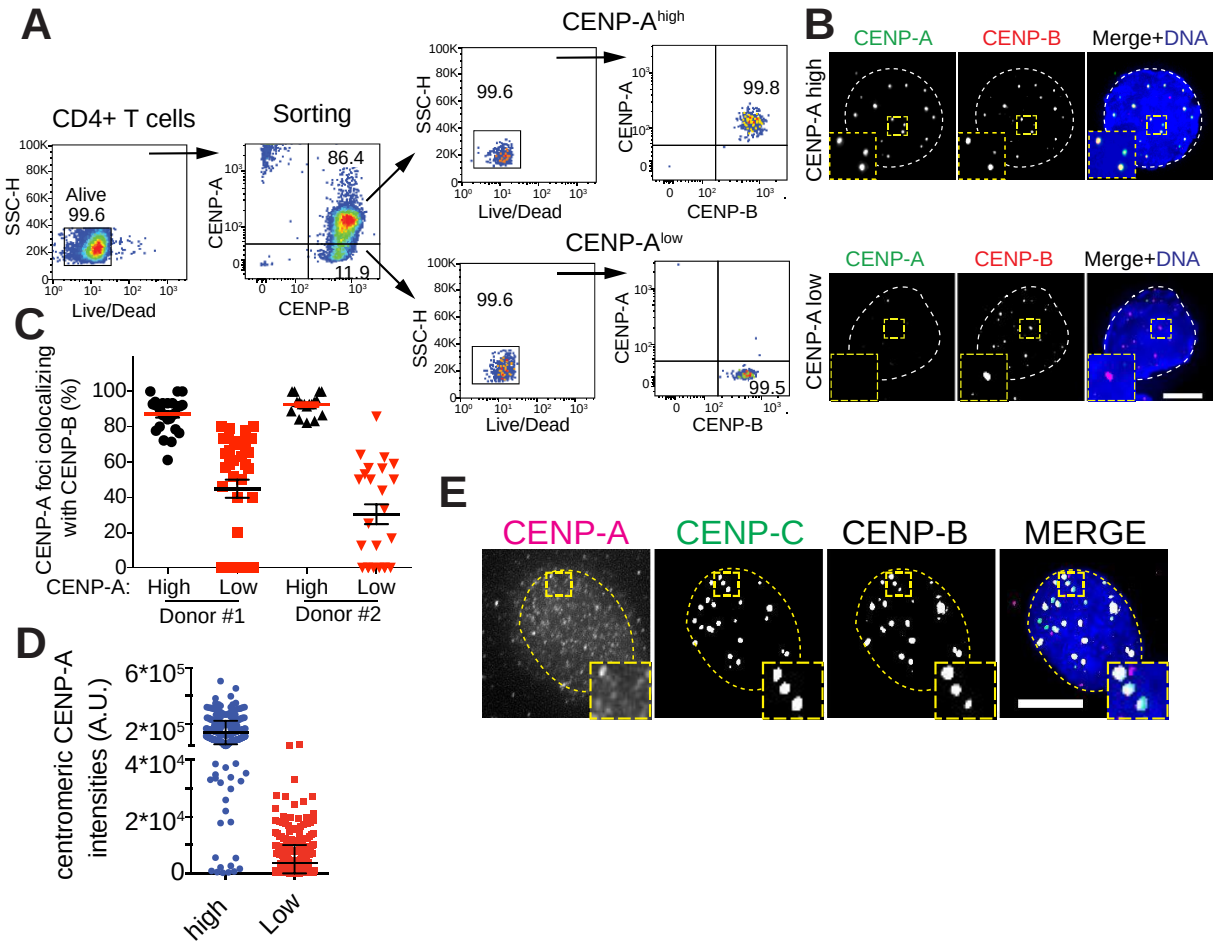


Figure 34. Centromere identity is maintained in resting CENP-A deprived CENP-B/C positive CD4$^+$ T cells.

*(**A**) Gating strategy for the sorting of freshly purified resting human CD4+ T cells based on CENP-A expression. Dead cells were excluded (first plot) using Fixable Viability Dye eFluor 780, CENP-A^{high} and CENP-A^{low} cells (both CENP-B+) were sorted (second plot). Post-sort analysis of each population is then shown (third and fourth plots).*

*(**B**) Representative images of sorted CD4+ T cells with high or low CENP-A level. The nucleus is contoured by a white dashed line. Scale bar, 5 µm.*

*(**C**) Graph showing CENP-A foci identified in high or low CENP-A expressing cells co-localizing with CENP-B foci in two donors. Each dot represents the percentage of CENP-A/B colocalizing in one cell. Error bar shows SEM.*

*(**D**) Quantification of centromeric CENP-A level in high or low CENP-A expressing CD4$^+$ T cells. Each dot represents one centromere. Error bars show standard deviation.*

*(**E**) Representative immunofluorescence images showing CENP-B and CENP-C positive centromeres, but lacking CENP-A in a CD4$^+$ T cell. Nucleus is contoured by a dashed yellow line. Scale bar, 5 µm*

This result prompted us to investigate CENP-A expression and localization upon T cell activation (**Figure 35A**). To track the number of cell divisions across time, purified CD4$^+$ T cells were labeled with the fluorescent dye CFSE (**Figure 35B**). In undivided cells (CFSE^{high}), we observed a gradual increase in CENP-A expression from day 1 to day 3 upon T cell activation (**Figure 35C-D**), consistent with the exit from quiescence and re-entry in the cell cycle (**Figure 35B-D**). The frequency of dead cells did not increase between day 1 and day 3 (**Figure 35H**), suggesting that CENP-A^{low} T cells were not lost as a result of compromised viability. In cells that had divided at least once (CFSE^{low}, day 3 to day 5), the CENP-A level was high and homogenous (**Figure 35B-D**). At day 3 only the remaining population that had not yet divided (div 0) contained cells expressing lower levels of CENP-A, compared to cells that had divided (**Figure 35I**). These results were consistent across independent blood donors (**Figure 35E-F, J**) and were confirmed by immunofluorescence analysis (**Figure 35G, K**). Our data thus suggest that CENP-A^{low} T lymphocytes convert into CENP-A^{high} cells upon T cell activation before cell division. Overall, we conclude that a physiologic sub-population of quiescent resting human CD4$^+$ T cells expresses CENP-B and CENP-C, but lacks detectable centromeric CENP-A. Upon cell cycle entry CENP-A is reloaded at endogenous centromeres, in agreement with the essential role of CENP-B/CENP-C in the maintenance of centromere identity.

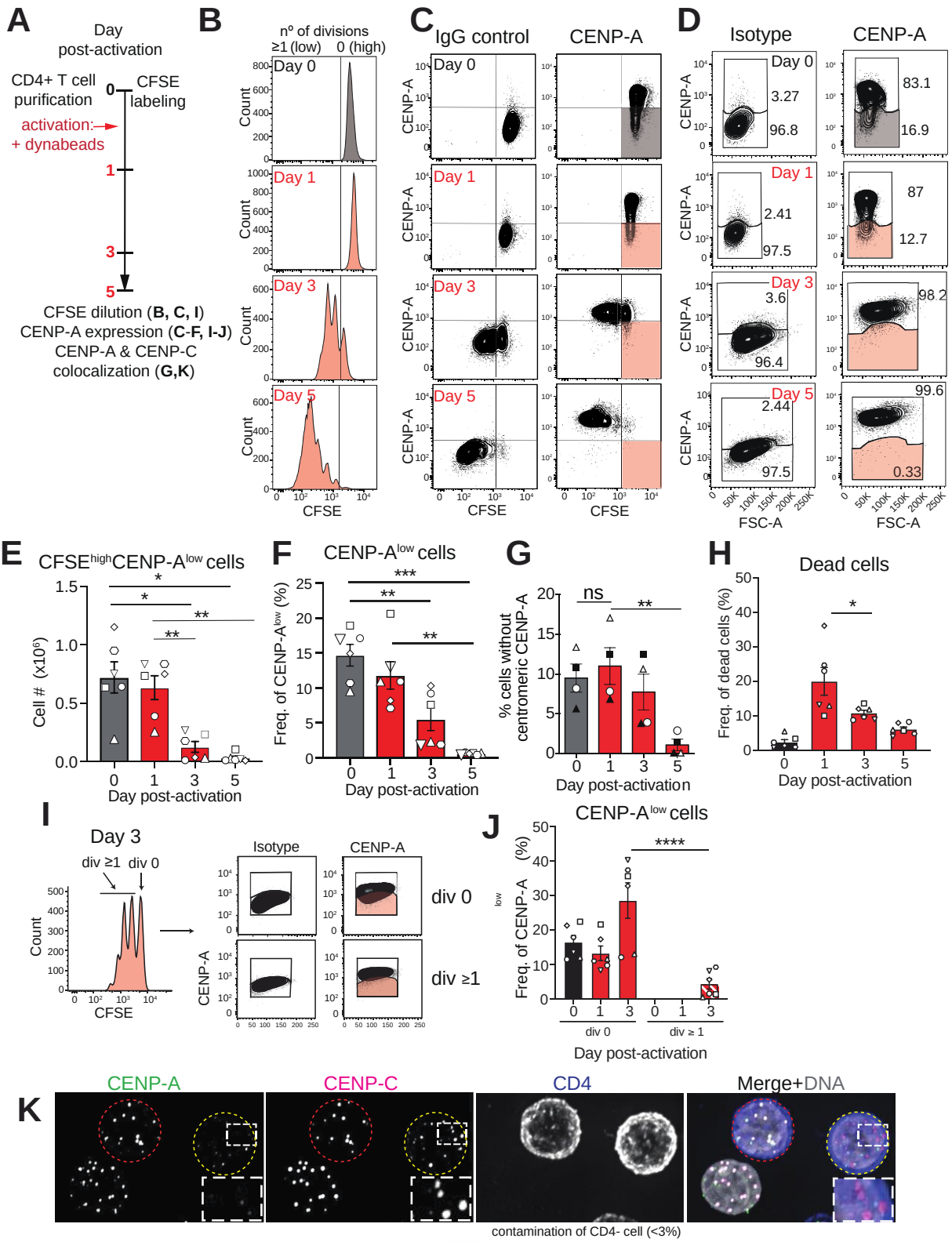


Figure 35: CENP-A deprived CD4⁺ T cells found in cells from human donors disappear upon T cell activation.

*(**A**) Schematic representation of the analysis of CENP-A expression over time in CD4+ T cells, shown in B-K. T cell were activated with anti-CD3/anti-CD28-coated beads.*

*(**B**) Tracking of cell division by CFSE dilution. The black line separates cells that have not divided (nº of div=0, CFSE-high) from those that have undergone at least one division (nº of div≥1, CFSE-low).*

*(**C**) CENP-A expression and CFSE dilution. Black lines were set based on CENP-A isotype control and CFSE maximal staining at day 0 (cells in shaded quadrants are CFSE-high/CENP-A-low). One representative donor is depicted.*

*(**D**) Representative plots showing CENP-A expression vs. forward scatter area (FSC-A), which determines the relative size of $CD4^+$ T cells after activation. Gates represent the frequency of CENP-A high and low populations in total $CD4^+$ T (shaded gates represent CENP-A^{low} cells).*

*(**E**) Absolute number of CFSE-high/CENP-A-low CD4+ T cells (shaded gate in C). Symbols represent individual donors. Error bars represent SEM from 6 independent donors. One-way ANOVA, multiple comparisons, n=6. *p < 0.05; **p < 0.01.*

*(**F**) Graph representing the frequency of CENP-A-low $CD4^+$ T cells during the experimental kinetics. One-way ANOVA, multiple comparisons, n=6 (each symbol represents a different donor). **p < 0.01, ***p < 0.001.*

*(**G**) Frequency of cells without centromeric CENP-A determined by immunofluorescence microscopy before and post activation. Symbols represent individual donors, n=4. Error bars represent SEM from 6 independent donors. Unpaired T test: **p=0.0064.*

*(**H**) Graph showing the frequency of dead cells in CD4+ T cell cultures over time. One-way ANOVA, multiple comparisons, n=6 (each symbol represents a different donor). Error bars show SEM. *p < 0.05*

*(**I**) CFSE dilution and CENP-A expression at day 3 post-activation. Representative FACS plots showing CENP-A expression in CD4+ T cells that have not divided (div 0) and those that have divided once or more times (div ≥1), gated based on CFSE dilution. Gates in CENP-A plots were set based on isotype control for each specific population (cells in shaded gates are CENP-A-low CD4+ T cells).*

*(**J**) Graph representing the frequency of CENP-A low cells (shaded gate in I). One-way ANOVA, multiple comparisons, n=6 (each symbol represents a different donor). Error bars show SEM. ****p < 0.0001.*

*(**K**) Representative immunofluorescence images showing CENP-A, CENP-C and CD4 staining after FACS. Cells with centromeric CENP-A are marked with a red dashed contour line while a yellow contour line marks cell without centromeric CENP-A. Scale bar, 5 µm.*

6. RESULTATS (Résumé en français)

Dans le cadre de mon projet de thèse, nos données indiquent que la position du centromère est préservée même en l'absence de CENP-A. Nous démontrons que CENP-A précédemment assemblé n'est pas essentiel pour le dépôt de CENP-A *de novo* ni pour contrôler son abondance, car les quantités de CENP-A augmentent très rapidement au niveau des centromères dépourvus de CENP-A, ce qui est en désaccord avec le modèle actuel (**Figure 11**). En outre, tout comme le rechargement canonique du CENP-A en présence de CENP-A précédemment déposé, le rechargement *de novo* du CENP-A est régulé par le cycle cellulaire et se produit exclusivement en sortie de mitose (**Figure 12/13**).

Nous montrons également que CENP-A est déposé su niveau du même HOR même en l'absence de CENP-A, ce qui montre que des mécanismes (épi)génétiques autres que la protéine CENP-A sont impliqués dans le maintien de la position du centromère (**Figure 16**).

Nous avons déjà démontré que la protéine CENP-B joue un rôle majeur dans la stabilisation des protéines des centromères, y compris une fraction du CENP-C, sur les centromères appauvris en CENP-A (Hoffmann et al, 2016). Nous avons donc émis l'hypothèse que CENP-B pourrait être important pour le dépôt *de novo* de CENP-A. Pour le vérifier, nous avons évalué le dépôt *de novo* de CENP-A chez le mutant CENP-B knock-out (KO) en utilisant le système CENP-A$^{OFF/ON}$ (Hoffmann et al, 2016). Nous avons constaté que CENP-B est nécessaire pour un rechargement efficace du CENP-A *de novo* en l'absence de tout CENP-A centromérique déposé précédemment (**Figure 18/20**).

Comme observé dans les cellules KO CENP-B, le rechargement *de novo* de CENP-A s'est également avéré altéré au niveau du chromosome Y qui ne contient naturellement pas de CENP-B (**Figure 21**). Nous avons ensuite étudié si le dépôt de CENP-A peut se produire à un endroit différent, en dehors de l'alpha-satellite, après la ré-expression de CENP-A. Cependant, le chromosome Y n'est pas un chromosome essentiel dans le modèle de culture cellulaire *in vitro*, et comme la formation de néocentromère se produit à très basse fréquence (Shang et al, 2013), le rechargement altéré de CENP-A devrait entraîner la perte du chromosome Y dans la majorité des

cellules (Ly et al, 2017). Pour prévenir sa perte, nous avons choisi de conserver le chromosome Y en insérant une cassette sélectionnable (Néomycine) (**Figure 22A-C**). Après épuisement et réactivation de CENP-A, nous avons soumis les cellules à un traitement G418 pour sélectionner les clones qui conservaient le chromosome Y. De manière intrigante, nous avons observé dans 15 % des cas la présence de CENP-C à un endroit ectopique du chromosome Y, et ne coïncidant pas avec la sonde FISH centromérique Y, indiquant la présence d'un néocentromère (**Figure 22D-E**).

Ces données indiquent que des éléments de type néocentromère peuvent se former sur un chromosome déficient en CENP-B suite à une rapide ré-expression de CENP-A, mais seulement lorsque l'identité du centromère médié par CENP-A a été perturbée de façon transitoire. Néanmoins, l'emplacement du centromère natif reste le site préférentiel pour la reformation du centromère.

Nous avons ensuite étudié comment CENP-B maintient la position des centromères à leur emplacement d'origine. Des études antérieures ont montré que CENP-C interagit directement avec CENP-B (Fachinetti et al., 2015 ; Suzuki et al., 2004), en plus de son interaction bien connue avec CENP-A (Fachinetti et al., 2013 ; Guse et al., 2011 ; Kato et al., 2013). Ici, nous avons remarqué que l'expression exogène de CENP-B a, non seulement, restauré le rechargement de CENP-A *de novo* dans les cellules CENP-$B^{-/-}$ (**Figure 23**), mais a également conduit à une augmentation des niveaux de CENP-C aux centromères qui se sont maintenus, pour lamajorité, même après l'épuisement de CENP-A (**Figure 23**). Ces observations indiquent que CENP-B, non seulement, maintient et stabilise CENP-C (Hoffmann et al., 2016), mais pourrait également recruter CENP-C *de novo* au centromère.

Pour tester cette hypothèse, nous avons utilisé le système LacO-LacI dans une lignée cellulaire U2-OS décrite précédemment (Janicki et al., 2004) dans laquelle nous avons en outre intégré le système CENP-$A^{OFF/ON}$ au locus endogène CENP-A. Ce système nous permet de tester si CENP-B peut recruter CENP-C à un locus ectopique LacO indépendamment de CENP-A (**Figure 24A**). Après la transfection de CENP-B-LacI-mCherry, le recrutement de CENP-C vers les grands sites ectopiques LacI/CENP-B a été observé dans la majorité des cellules, même après un traitement par IAA pour épuiser CENP-A, contrairement au contrôle LacI/mCherry (**Figure 24B-D**).

Étant donné le fort enrichissement de CENP-B-LacI au niveau du réseau LacO, il n'est pas clair si l'interaction avec CENP-C est pertinente pour les nouveaux dépôts de CENP-A aux centromères endogènes. En outre, le CENP-B-LacI du réseau LacO pourrait potentiellement se regrouper avec les centromères endogènes et, dans une certaine mesure, influencer notre observation du recrutement du CENP-C. Nous avons donc vérifié si CENP-B pouvait favoriser la formation de centromères également au niveau du centromère natif, où les niveaux du CENP-B sont bien plus

faibles que ceux du système LacI. Pour tester cela, nous avons généré une lignée cellulaire DLD-1 dans laquelle les protéines CENP-A et CENP-C endogènes peuvent être rapidement épuisées (et ré-exprimées) simultanément en utilisant le système de dégradation des protéines inductibles à l'auxine (CENP-A/C$^{OFF/ON}$; **Figure 27**).

Après CENP-A/C$^{OFF/ON}$, nous avons observé un rechargement des deux protéines à certains centromères positifs en CENP-B dans ~30% des cellules (**Figure 28**). Nous avons remarqué que seule une fraction de centromères par cellule présentait une recharge efficace (**Figure 28C**). Après la suppression additionnelle de CENP-B par ARN interférant (**Figure 28F**), le rechargement *de novo* de CENP-A/C a cependant été presque complètement perdu (**Figure 28G-I**).

Ainsi, nos résultats montrent que, parfois CENP-B peut déclencher le dépôt de CENP-A/C pour maintenir la position du centromère dans les cellules dépourvues de CENP-A/C.

Nous avons ensuite voulu identifier le mécanisme par lequel le CENP-B favorise *de novo* le rechargement de CENP-A/C au niveau du centromère. Nous avons donc disséqué les événements temporels qui contrôlent la formation du centromère au niveau des centromères positifs en CENP-B endogènes (**Figure 30A**). Pour ce faire, nous avons développé un système permettant de co-déprimer rapidement CENP-A et CENP-C et d'induire séparément l'expression des CENP-A ou CENP-C exogènes (e) (**Figure 30B-C**).

En utilisant ce système, nous n'avons pu observer le rechargement de CENP-A exprimé ectopiquement au niveau du centromère (**Figure 31**).

En revanche, lors de l'expression exclusive de l'eCENP-C, nous avons observé un rechargement partiel du CENP-C malgré l'absence de tout CENP-A/C endogène (**Figure 32**). Cela indique que CENP-C est recruté en premier au centromère alors que CENP-A a besoin de CENP-C pour son rechargement *de novo*.

Nos résultats impliquent que CENP-B incarne la mémoire de l'identité du centromère. Nous avons donc cherché à évaluer la pertinence physiologique de ce mécanisme. Alors que les cellules en division maintiennent l'expression de CENP-A, des cellules différenciées et non en division dépourvue de CENP-A ont été décrites (Swartz et al., 2019). Il existe des sous-ensembles de lymphocytes différenciés au repos dans le sang périphérique d'un individu sain, et ceux-ci peuvent réintégrer le cycle cellulaire lors de leur activation par le récepteur des cellules T (TCR). Nous avons émis l'hypothèse qu'une sous-population de lymphocytes T en circulation peut contenir une fraction de cellules CENP-A-négatives, et qu'elle peut acquérir à nouveau l'expression CENP-A et le dépôt de

centromère lors de son activation. Nous avons mesuré le niveau de CENP-A et CENP-B dans des cellules T $CD14^{-}/CD4^{+}$ au repos et non divisées, isolées du sang périphérique de donneurs humains sains (**Figure 33A**). Alors que toutes les cellules T CD4+ analysées étaient CENP-B-positives, deux populations distinctes ont été détectées sur la base de leur expression de CENP-A, appelée CENP-A^{high} (haute) et CENP-A^{low} (faible) (**Figure 33B**).

Ce résultat nous a incité à étudier l'expression et la localisation de CENP-A lors de l'activation des cellules T (**Figure 35A**). Dans les cellules indivisées, nous avons observé une augmentation progressive de l'expression de CENP-A du jour 1 au jour 3 lors de l'activation des cellules T (**Figure 35C-D**), ce qui correspond à la sortie de la quiescence et à leur retour dans le cycle cellulaire (**Figure 35B-D**). Dans les cellules qui s'étaient divisées au moins une fois (du jour 3 au jour 5), le niveau de CENP-A était élevé et homogène (**Figure 35B-D**). Au jour 3, seule la population restante qui ne s'était pas encore divisée contenait des cellules exprimant des niveaux de CENP-A inférieurs à ceux des cellules qui s'étaient divisées (**Figure 35I**). Nos données suggèrent donc que les lymphocytes T CENP-A^{low} se transforment en cellules CENP-A^{high} lors de l'activation des cellules T avant la division cellulaire. Dans l'ensemble, nous concluons qu'une sous-population physiologique de cellules T $CD4^{+}$ humaines au repos exprime CENP-B et CENP-C, mais ne possède pas de CENP-A centromérique détectable. Lors de l'entrée dans le cycle cellulaire, CENP-A est rechargée aux centromères endogènes, en accord avec le rôle essentiel de CENP-B/CENP-C dans le maintien de l'identité des centromères.

7. DISCUSSION

In this work, we identify CENP-B as a key contributor to maintain centromere position together with CENP-A (**Figure 36(I)**). By preserving a critical level of CENP-C at native centromeres, CENP-B provides memory for maintenance of native human centromeres by promoting *de novo* CENP-A assembly. This has a physiological impact in cells that have temporarily lost CENP-A (e.g. a subpopulation of resting $CD4^{+}$ T cells) where the CENP-B/CENP-C connection is key to preserving the original centromere identity. Additionally, under selective pressure, neocentromeres-like elements form at CENP-B-negative centromeres. We demonstrated that CENP-B bound to centromeric DNA can trigger the recruitment of CENP-C, not only at ectopic LacO site, but also at native centromeres (**Figure 36(II)**). We show that CENP-C is key factor for CENP-A loading which, subsequently, initiates the maintenance of centromere position in a CENP-A-

dependent manner. While CENP-C has been mainly described as the reader of the epigenetic mark CENP-A, now we have demonstrated its capability to recognize and mark centromere position via CENP-B and independently of CENP-A. Moreover, we showed that CENP-A itself, in the absence of CENP-C, lacks this ability to reinstate centromere memory. In contrast to previous models, our observations suggest that *de novo* centromere formation is likely initiated by a critical level of CENP-C rather than CENP-A (**Figure 36(III)**). However, our data still emphasize the importance of the centromeric self-assembly loop to efficiently maintain centromere position for indefinite cell cycles, a function that cannot be fulfilled exclusively by CENP-B and CENP-C themselves. This is in agreement with the well-studied essentiality of CENP-A in proliferating somatic cells in all CENP-A-containing species studied so far (Fukagawa and Earnshaw, 2014; McKinley and Cheeseman, 2016; Stankovic and Jansen, 2017).

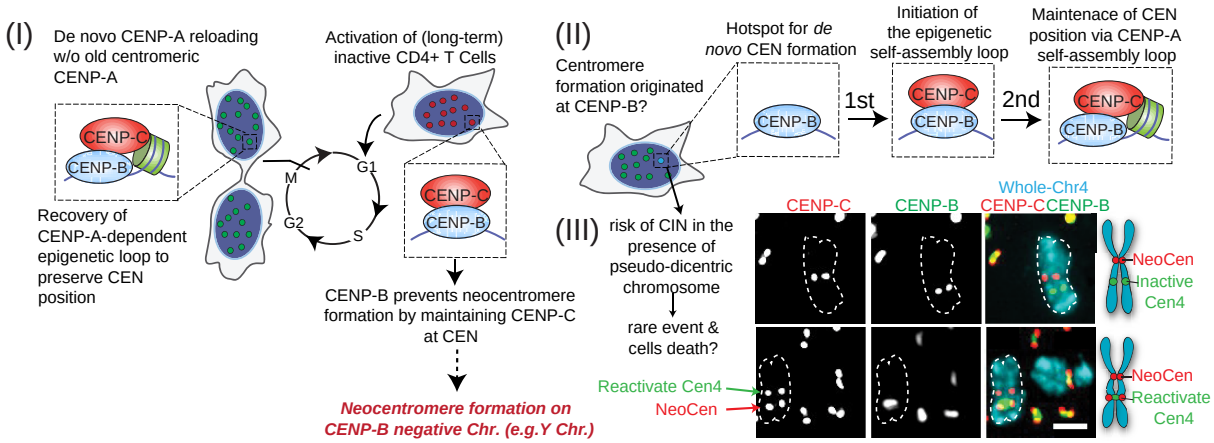


Figure 36 Model of the role of CENP-B in the maintenance of centromere identity. *I-III: See text for details.*

Our data have several implications. First, as CENP-A also incorporates outside the canonical centromeric sites (Bodor et al., 2014; Nechemia-Arbely et al., 2019), we propose that CENP-B is required to prevent neocentromere formation by preserving CENP-C at the original centromere location and favoring the maintenance of centromere identity in a CENP-C-dependent manner (**Figure 36(I)**). This is even more critical in the event of CENP-A up-regulation, as observed in some types of cancer, although to date is still unclear how massive CENP-A mis-incorporation promotes the emergence of genome instability (Barra and Fachinetti, 2018). Certainly, in contrast to what has been observed in flies (Heun et al., 2006; Olszak et al., 2011), ectopic CENP-A is unable to drive neocentromere formation in human cells. This could be explained by the lack of CENP-C-mediated nucleosome protection mechanism of ectopic CENP-A during DNA replication (Nechemia-Arbely et

al., 2019) that we argue is strongly supported by CENP-B. Furthermore, we have proven that, under selective pressure, neocentromere-like elements can form on the CENP-B-deficient centromere of the Y chromosome (**Figure 22**).

The contribution of the CENP-B/CENP-C connection to centromere identity is most relevant at CENP-A-depleted centromeres (**Figure 36(I)**). We identified a population of resting $CD4^+$ T cells characterized by low CENP-A expression with undetectable localization at centromeres (**Figure 34A-E)**. The origin and properties of T cells lacking CENP-A warrants further study. We hypothesize that CENP-A loss under physiological conditions could be a consequence of prolonged cell cycle withdrawal, a feature of quiescent cells. This suggests the inability to reload CENP-A during quiescence, a phenomenon recently described (Swartz et al., 2019). Importantly, this loss is reversible in $CD4^+$ T cells and we find that resuming the cell cycle restores CENP-A expression and centromere localization. In this scenario, upon cell activation, T cells are more likely to reactivate CENP-A transcription during G2 (Shelby et al., 1997) and to reload it at mitotic exit via its canonical pathways (Zasadzińska and Foltz, 2017), as the absence of CENP-A in the first mitosis can be tolerated (Hoffmann et al., 2016). This type of *de novo* CENP-A reloading was also observed in our CENP-$A^{OFF/ON}$ system using human cell lines as we found full recovery of centromeric CENP-A level within one **(Figure 11H)** or two cell cycles (**Figure 11F**). This suggests the existence of a quantitative transmission likely determined by CENP-A expression levels mechanism that preserves the number of CENP-A molecules independently of the presence of previously deposited CENP-A. In T cells, one cell cycle was sufficient to recover full and homogenous CENP-A expression, irrespective of the initial fraction of CENP-A-low cells among independent donors. This demonstrates that the amount of previously deposited CENP-A molecules is not the key determinant of total centromere CENP-A. A complete turnover of all preexisting CENP-A molecules has been only previously observed in holocentric organisms (Gassmann et al., 2012). Our findings further demonstrate that post-translational modifications on preexisting centromeric CENP-A are not essential to guide new CENP-A deposition at the centromere in contrast to previously described models (Niikura et al., 2016, 2015).

Our results also show that endogenous CENP-B level enables centromere formation via CENP-C recruitment (**Figure 36(II)**). This finding has implications for the mechanisms that drive HAC formation via CENP-B, as we dissected the temporal events necessary for maintenance of centromere position. A previous report proposed that CENP-B promotes HAC formation via direct recruitment of CENP-A (Okada et al., 2007), since, to a certain extent, they stabilize each other (Fachinetti et al., 2015; Fujita et al., 2015). However, our data on ectopically tethered CENP-B and

at native centromeres show that this interaction appears insufficient to load and stabilize CENP-A in the absence of CENP-C (**Figure 31**). Importantly, CENP-C is a key factor for centromere formation as it is both able to recruit CENP-A via the Mis18 complex (Dambacher et al., 2012; Moree et al., 2011; Pan et al., 2017; Stellfox et al., 2016) and stabilize it (Falk et al., 2016, 2015). Interestingly, the Mis18 complex was not sufficient to mediate detectable endogenous CENP-A assembly when tethered to an ectopic location (LacO) (Ohzeki et al., 2019), reinforcing the importance of CENP-C to maintain centromere position. However, our data do not entirely rule out a contribution of centromere components other than CENP-C to stabilize the Mis18 complex in human centromeres.

Finally, the existence of stably inherited pseudodicentric neocentric chromosomes – in which both the non-repetitive neocentromere bound by CENP-A and the inactive centromeric locus carrying satellite DNA and CENP-B are found on the same chromosome [e.g. PD-NC4 cells (Amor et al., 2004)] – disfavors the notion that CENP-B initiates centromere formation. In this scenario, the presence of ectopic CENP-B should represent a threat for the cells as it can occasionally lead to a dicentric chromosome. Indeed, we sporadically observed the spontaneous re-activation of the inactive native centromere in PD-NC4 cell line and the formation of a dicentric chromosome (**Figure 36(III)**). However, in these cells the abundance of repetitive DNA and CENP-B at the inactive centromere were particularly low (Fachinetti et al., 2015), perhaps reducing its ability to promote centromere formation via CENP-C recruitment. The potential to become a dicentric chromosome could further explain why the occurrence of pseudodicentric chromosomes remains a very rare event and that CENP-B only inefficiently promotes centromere formation on HACs (Ohzeki et al., 2002; Okada et al., 2007). Interestingly, to date, 3 out of 8 cases of described pseudodicentric neocentric chromosomes in humans occur on the Y chromosome (Marshall et al., 2008) where centromere re-activation is predicted to be less frequent due to lack of CENP-B. Previous reports have suggested that CENP-B is removed from non-centromeric regions via its chaperone Nap1 (Tachiwana et al., 2013). Considering our findings, such regulation appears also to be important to prevent CENP-B-driven centromere formation outside of alpha-satellite DNA.

Altogether, our data suggest cooperation of CENP-A/CENP-B/CENP-C in maintenance of centromere position and function. CENP-C via CENP-B is the initiation factor for new centromere formation capable to promote de novo CENP-A deposition, that in turn stabilized CENP-C in a temporal manner. What causes the heterogeneity of CENP-B mediated CENP-C recruitment remains to be identified. As the percentage of cells that reload CENP-C doubles following ectopic CENP-C expression (**Figure 32C**), it is possible that the inefficient reloading is due to low CENP-C protein level following IAA WO. On this regard, the importance of CENP-C for endogenous CENP-A

loading was previously observed at ectopic chromatin-integrated LacO sites (Shono et al, 2015; Hori et al, 2013), in which protein levels at these ectopic site are tremendously higher compared to the native centromere. As protein level recover gradually over time following IAA wash-out, this can explain why we initially failed to identify rapid new CENP-C loading (within 3hr) at endogenous centromeres in the absence of CENP-A (Hoffmann et al, 2016). CENP-C is also absent at the CENP-B-bound inner centromere during metaphase. Possibly, the chromatin environment or post-translational modifications might also regulate CENP-C recruitment. On this regard, the acidic domain of CENP-B was recently described to function as an interface for several chromatin remodeling proteins with either chromatin compaction or decompaction activities (Otake et al., 2020). Although the precise regulation of this CENP-B-mediated chromatin alteration remains unclear, chromatin compaction could impair CENP-C recruitment and thus explain partial CENP-C recruitment.

Another question that remains to be elucidated is the role of CENP-I in centromere formation as it was shown to recruit CENP-A at an ectopic site (Hori et al., 2012; Shono et al., 2015) and to display an extended localization profile compared to the other CCAN proteins in a similar, but not identical pattern to CENP-B (Kyriacou and Heun, 2018). However, *in vitro* assays favor the hypothesis of a cooperative binding between CENP-I and the CCAN likely via a direct interaction with CENP-C (Klare et al., 2015; McKinley et al., 2015; Pesenti et al., 2018; Weir et al., 2016) that in turn promotes and stabilizes centromeric CENP-A.

Finally, our model does not exclude that CENP-B and repetitive DNA could maintain the centromere in a manner that is independent of the direct interaction with CENP-B and CENP-C. A recent computational study suggested that CENP-B induces the formation of specific non-B DNA structure (Kasinathan and Henikoff, 2018) that could potentially facilitate CENP-A incorporation. Interestingly, BACs containing centromeric DNA favor CENP-A assembly in contrast to non-centromeric DNA (Aze et al., 2016). Non-B DNA structures are also observed at CENP-B-free centromeres (neocentromeres and the Y chromosome) and could explain the formation of *de novo* centromere on some types of DNA in a HAC formation assay (Logsdon et al., 2019). In support of this hypothesis, our data on CENP-A reloading at the Y chromosome show that the alpha-satellite DNA sequences provide CENP-B independent features that favor centromere re-formation. More studies are needed to shed light on the centromere DNA secondary structures and their possible role in centromere biology. Interestingly, in *S. pombe*, centromere DNA inherently destabilizes H3 nucleosomes to favor CENP-A deposition (Shukla et al., 2018b). Previous studies have also reported that CENP-C has DNA binding activity, that could potentially promote CENP-C loading to site of *de*

novo centromere formation in several species, although its DNA sequence specificity remains uncertain (Politi et al., 2002; Sugimoto et al., 1994; Xiao et al., 2017; Yang et al., 1996). Additionally, the alpha satellite locus bears marks of active transcription (B. A. Sullivan and Karpen, 2004) and satellite transcripts might play a direct role in centromere formation by mediating CENP-A and CENP-C incorporation (Bergmann et al., 2011; Bobkov et al., 2018; Chan et al., 2012; Chen et al., 2015; McNulty et al., 2017; Quénet and Dalal, 2014). CENP-B might play an additional role in transcript regulation possibly by competing with DNA methylation and methylated DNA binding proteins at CENP-B boxes (Scelfo and Fachinetti, 2019). Lastly, CENP-B was proposed to act as a chromatin regulator by directly modulating heterochromatin formation (Morozov et al., 2017; Okada et al., 2007). Indeed, the acetyltransferase KAT7 via M18BP1 has been implicated in CENP-A deposition (Ohzeki et al., 2016). Future studies will be important for defining the interplay among DNA sequence, transcription and epigenetic modifications in establishing and maintaining centromere identity via CENP-B and CENP-C.

8. DISCUSSION (Résumé en français)

Dans le chapitre de la discussion, nous examinons la contribution de la connexion CENP-B/CENP-C dans l'identité du centromère dans des conditions physiologiques. Nous émettons l'hypothèse que la perte de CENP-A dans des conditions physiologiques pourrait être la conséquence d'un retrait prolongé du cycle cellulaire, une caractéristique des cellules quiescentes. Il est important de noter que cette perte est réversible dans les cellules T $CD4^+$ et nous constatons que la reprise du cycle cellulaire rétablit l'expression de CENP-A et la localisation du centromère. Nous discutons en outre du fait qu'un rechargement *de novo* de CENP-A pourrait dans ce cas se produire à la sortie de la mitose via sa voie canonique (Zasadzińska & Foltz, 2017), car l'absence de CENP-A peut être tolérée au cours de la première mitose (Hoffmann et al, 2016). Ce type de rechargement *de novo* de CENP-A a également été observé dans notre système CENP-$A^{OFF/ON}$ utilisant des lignées cellulaires humaines, car nous avons constaté une récupération complète du niveau de CENP-A centromérique après un (**Figure 11H**) ou deux cycles cellulaires (**Figure 11F**). Sur la base de nos observations, nous émettons l'hypothèse qu'il existe un mécanisme de transmission quantitatif, probablement déterminé par les niveaux d'expression de CENP-A, qui préserve le nombre de molécules CENP-A indépendamment de la présence de protéines CENP-A précédemment déposées.

Nous discutons également du rôle de CENP-B dans l'identité des centromères et examinons les premières étapes de la formation des centromères via CENP-B. Celle-ci est nécessaire pour prévenir la formation de néocentromères en préservant CENP-C à l'emplacement original du centromère et en favorisant le maintien de l'identité des centromères de manière dépendante de CENP-C (**Figure 36(I)**). Nos résultats montrent que le niveau endogène de CENP-B permet la formation de centromère par le recrutement de CENP-C (**Figure 36(II)**).

Nous commentons également l'existence de chromosomes néocentriques pseudodicentriques hérités de manière stable - dans lesquels le néocentromère non répétitif lié à la présence de CENP-A et le locus centromérique inactif portant l'ADN satellite et le CENP-B se trouvent sur le même chromosome [par exemple les cellules PD-NC4 (Amor et al, 2004)]. Ces cas vont à l'encontre de l'idée selon laquelle le CENP-B initie la formation de centromère. En effet, nous avons sporadiquement observé la réactivation spontanée du centromère natif inactif dans la lignée cellulaire PD-NC4 et la formation d'un chromosome dicentrique (**Figure 36(III)**). Cependant, dans ces cellules, l'abondance de l'ADN répétitif et du CENP-B au niveau du centromère inactif était particulièrement faible (Fachinetti et al, 2015), ce qui a peut-être réduit sa capacité à favoriser la formation du centromère par le recrutement de CENP-C.

Nous discutons plus en détails l'hétérogénéité du recrutement de CENP-C par CENP-B. En effet, l'échec du recrutement de CENP-C par CENP-B pourrait avoir plusieurs raisons. Par exemple, l'environnement chromatinien ou des modifications post-traductionnelles pourraient réguler le recrutement de CENP-C. Il est intéressant de noter que CENP-C est également absent au niveau du centromère interne lié au CENP-B pendant la métaphase, ce qui pourrait laisser entrevoir un rôle important de l'environnement chromatinien dans le recrutement du CENP-C par CENP-B.

Dans ce chapitre, nous développons également des modèles alternatifs sur la formation de centromère par CENP-B qui ne sont pas liés à un recrutement direct de CENP-C par CENP-B. CENP-B pourrait par exemple induire la formation d'une structure spécifique d'ADN dite « non B » (Kasinathan & Henikoff, 2018) qui pourrait potentiellement faciliter l'incorporation CENP-A (ou CENP-C). Nous avons conclu que d'autres études sont nécessaires pour faire dévoiler les structures secondaires de l'ADN des centromères et leur rôle possible dans la biologie des centromères.

Dans l'ensemble, nous avons conclu que les études futures seront importantes pour définir l'interaction entre la séquence d'ADN, la transcription et les modifications épigénétiques dans l'établissement et le maintien de l'identité des centromères via CENP-B et CENP-C.

9. METHODS

Cell culture

Cells were cultivated at 37°C in a 5% CO2 atmosphere. Flp-In TRex DLD-1 cells and U-2OS 2-6-3 R.I.K LacO cells (Janicki et al., 2004) were maintained in Dulbecco's modified essential medium (DMEM) medium containing 10% tetracycline free Fetal Bovine Serum (FBS, Pan Biotech). U-2OS cells were cultured with 0.1 mg/ml hygromycin (Invitrogen). Immortalized hTERT RPE-1 cells were maintained in DMEM:F12 medium containing 10% Fetal Bovine Serum (BioSera), 0.123% sodium bicarbonate and 2 mM L-glutamine.

IAA (I5148; Sigma) dissolved in ddH_2O was used at 500 µM, doxycycline (1 µg/ml), TMP (10 µM), Reversine (0.5 µM), Colcemid (0.1 mg/ml, Roche), Palbociclib (1 µM), G418^{TM} (0.3 mg/ml, Gibco^{TM}), Flavopiridol (5 µM, Sigma) and NM-PP1 (5µM, Sigma). IAA was washed-out three times using culture medium. Efficient IAA removal was ensured by first repeating the washes after 15 min and second after 30 min to allow excess compound to diffuse from cells. For DOX/TMP wash-outs, cells were additionally detached (TrypLE^{TM}, Gibco), washed by centrifugation and then re-seeded on glass slides in normal culture medium.

Gene Targeting

Gene targeting was performed using either CRISPR/Cas9 to target CENP-B (5'-CACCGCGCGATCTCGCCCTTGCGCAAAC-3') and RPS4Y (guide RNAs: 5'-caccgTCCGTCGCAGAGTTTCGCCA-3') or TALENs to target CENP-A (5'-GTCATGGGCCCGCGCC-3' and 5'-GGCCCCGAGGAGGCGCA-3') and CENP-C (5'-GAGGAAAGTGTTCTTC-3' and 5'-GGTTGATCTTTCATC-3'), as described previously (Fachinetti et al., 2015; Hoffmann et al., 2016; Hoffmann and Fachinetti, 2018). Here, we introduced a mini-AID-LoxP-P2A-Neomycin-LoxP sequence at the C-terminus of the CENP-C gene and a mCherry sequence at the C-terminus of the CENP-B gene using modified pUC57 plasmids harboring the inserts and homology arms (assembled via Gibson cloning). A start codon deprived neomycin resistance cassette was integrated in frame with the start codon of the non-essential RPS4Y gene on chromosome Y. Cells were selected either by neomycin treatment (0.3 mg/ml) or FACS sorted as

described previously (Hoffmann and Fachinetti, 2018) (Hoffmann & Fachinetti, 2018). Successful integration was confirmed by PCR, immunoblot and/or immunofluorescence microscopy. Analog Sensitive (AS) CDK1 expressing RPE-1 cells were generated using a one-shot isolation strategy as described in Saldivar et al. 2018. In brief, puromycin sensitive RPE-1 CENP-A$^{EA/EA}$ were transfected with a CDK1^{AS} expression vector, a Sleeping Beauty transposase expression vector and an endogenous CDK1 inactivation vector (0.5 µg for each plasmid) using nucleofection (Lonza). Successful integration of CDK1^{AS} give rise to puromycin (5 µg/ml) resistance. G2 phase cell cycle block was confirmed in selected and isolated clones by FACS analysis after overnight NM-PP1 (5µM) treatment.

Generation of Stable Cell Lines

The FRT/Flp-in system was used to generate stable cell lines as described previously (Fachinetti et al., 2013). Briefly, we integrated cDNA of carboxy-terminal tagged CENP-B^{mcherry} and CENP-C$^{DHFR-RFP}$ as well as amino-terminal tagged CENP-A$^{DHFR-RFP}$ or CENP-A$^{EYFP-AID}$ into a pcDNA5/FRT plasmid harboring a promoter and start-codon lacking hygromycin resistance cassette and co-transfected this plasmid with a pOG44 plasmid (Flp-recombinase) in a 9:1 ratio into Flp-In TRex DLD-1 cells using FugeneHD (Promega). Correct integration of the insert at the isogenic FRT site gives rise to hygromycin resistance. After selection (0.3 mg/ml hygromycin, Invitrogen) single clones were isolated and tested for successful integration by immunofluorescence microscopy.

os-TIR1-9-myc was integrated using a recombinant retrovirus as described previously (Holland et al., 2012). Single cells were isolated using 5 µg/ml puromycin.

siRNA, transient transfections, EdU staining and Colony Formation Assay

Lipofectamine RNAiMax (Invitrogen) was used to introduce siRNAs as described previously (Fachinetti *et al*, 2013). siRNAs SMARTpool against Luciferase, M18BP1, CENP-B, CENP-C were purchased from Dharmacon. siRNA pools against HJURP and DAXX were a kindly gift from the G. Almouzni laboratory (Lacoste et al., 2014). Lipofectamine RNAiMAX and siRNAs were mixed in serum-free OptiMEM (Gibco^{TM}). LacI constructs were transfected into U-2OS cells using FugeneHD (Promega) following the manufacturer's instructions. FugeneHD and plasmids were mixed in serum

free OptiMEM. For this study, we generated different LacI-CENP-B constructs: HA-LacI or mCherry-LacI-CENP-B$^{full\text{-}length}$, HA-LacI-CENP-B$^{\Delta acidic}$ (1-405; 539-599), HA-LacI-^{NLS}CENP-B$^{\Delta DBD}$ (125-599), ^{NLS}CENP-B$^{\Delta DBD+\Delta acidic}$ (125-405; 539-599) using Gibson assembly cloning. Plasmids for HA-LacI-H.3.3 and mCherry-LacI control constructs were kindly gifts from B. Black laboratory. CENP-T (1-374) LacI-NLS (Gascoigne et al., 2011) was purchased from addgene.
EdU click labeling was performed using Click-iT® labelling technologies (Thermo-Fisher scientific) and colony formation assays was performed as previously described (Barra et al., 2019).

Immunoblotting

Cell pellets were suspended in protein sample buffer and samples were separated by SDS-PAGE, transferred onto nitrocellulose membranes (Bio-Rad), and revealed with the following antibodies: DM1A (α-tubulin, 1:5,000), CENP-A (1:1,000; Cell Signaling), GFP (1:1,000; Chromotek), HJURP (1:1,000; a kindly gift from D. Foltz), DAXX (a kindly gift from G. Almouzni), CENP-B (1:1,000; Abcam), GAPDH (1:10,000; Abcam), CENP-C (a kindly gift from I. Cheeseman and B. Black) and Vinculin (1:2000, Sigma).

Immunofluorescence, Chromosome Spreads, IF-FISH and Live-Cell Microscopy

Cells were fixed in 4% formaldehyde 0.1% Triton X-100 at room temperature for 10 min or in methanol at -20°C for 15 min and subsequently blocked in blocking buffer (2.5% FBS (v/v), 0.2 M Glycine, 0,1% triton X-100 (v/v) in PBS) for 30 min at room temperature or overnight at 4°C. Primary antibodies were incubated in blocking buffer for 2 h at room temperature. The following antibodies were used: CENP-A (1:1,000; clone 3-19, ENZO, ref. ADI-KAM-CC006-E), CENP-C (1:1,000; Clinisiences, ref. PD030PD030clinisciences), CENP-B (1:1,000; Abcam, ref. ab25734), ACA (1:500; Antibodies), DM1A (α-tubulin, 1:2,000), HA (1:500; Institut Curie antibody platform), CENP-A (1:1000, ref. #2186S, Cell Signaling), GFP and RFP booster (1:200, Chromotek), APC-CD4 (1:50, ref. 555349, BD). Immunofluorescence in combination with fluorescence in situ hybridization (FISH) was performed as described previously(Dumont *et al*, 2020). Images of DAPI counterstained and immuno-/FISH-stained cells were collected using a Deltavision Core system (Applied Precision). If necessary, microscope stage coordinates were recorded for sequential FISH

after immunofluorescence microscopy. For M18BP1 staining in hRPE-1 cells were fixed in methanol (-20°C) and extracted prior to blocking (in BSA) as previously described (Fachinetti et al., 2013). M18BP1 antibody was a kind gift from Paul Maddox, University of North Carolina. All secondary antibodies used for immunofluorescence microscopy were purchased from Jackson Immuno Research. After immunostaining cells were DAPI counterstained and mounted using anti-fading reagent (Life technologies).

Movies of live-cell were acquired using Deltavision Core system (Applied Precision) as described previously (Hoffmann et al., 2016). Cells were grown on high optical quality plastic slides (ibidi) for this purpose.

Single molecule microscopy

For Single Molecule Microscopy (SMM) cells were seeded on optical quality glass bottom dishes (FluoroDish^{TM}, World Precision Instruments). SMM was performed under conditions of 37°C and 5% CO_2 using a Tokai Hit heating system on an epifluorescence inverted microscope (IX71, Olympus). A Highly Inclined and Laminated Optical sheet (HILO) configuration was employed to reduce background fluorescence and enable the detection of single molecules (Tokunaga *et al*, 2008). Specifically, a 500 mm achromatic lens was used to focus the beam at the back focal plane of a 150x objective (UApo N 150X TIRF 1.45 NA, O.I., Olympus, France). A metallic mirror on a translation stage was used to displace the focal point of the beam from the center of the back focal plane of the objective and determine the angle of the beam. A slit allowed us to regulate the thickness of the HILO illumination. Illumination was performed using a 488 nm laser (488LM-200, ERROL, France) and a 564 nm laser (Sapphire 561, Coherent, Santa Clara, CA, USA). Illumination was controlled by an acousto-optical tunable filter (AOTFnC-400-650-TN, A&A Optoelectronic, France). A quad band dichroic mirror (FF409/493/573/652-Di02-25x36, SEMROCK) was used to separate emission light from excitation light. Use of a quad band dichroic mirror led to small, but significant excitation of mCherry molecules using 488 nm laser. To block this signal, an additional 520/70 emission filter (SEMROCK) was used when illumination with the 488 nm laser was applied. Images were recorded using an EM-CCD camera (iXonEM DV860DCS-BV, Andor, Ireland) run in frame-transfer mode using a 10 millisecond time interval for a total of 2000 frames. In case of overexpression of the GR^{EYFP} protein, the signal was bleached until individual fluorescence could be distinguished. To capture a single molecule of CENP-A^{EYFP} and avoid bleaching great care was taken to avoid any

illumination of the cell prior to the recording of the images. To that end, the CENP-B^{mcherry} signal was used to obtain proper focus on the centromeres prior to imaging of CENP-A^{EYFP}. Post-acquisition, integrated densities at the centromere or of single GR^{EYFP} spots were determined by manually drawing a 6x6-pixel square around the centromere marked by CENP-B^{mcherry}/CENP-C$^{AID-mcherry}$ or GR^{EYFP} spots using the software FIJI. For background corrections, we drew a bigger square (8x8-pixel) and generated a band of 1 pixel width that surrounds the first 6x6 pixel square as illustrated in Figure 15A. The integrated density of the background band was subtracted from the integrated density of the inner square.

CUT&RUN-sequencing and -qPCR

CUT&RUN was performed according to the procedure reported by Skene and Henikoff (2017) (Skene and Henikoff, 2017) starting from one million cells and using anti-CENP-A (Ozyme, 2186S), -CENP-B (Abcam, ab25734), or -CENP-C (Abcam, ab33034) antibodies. Rabbit IgG isotype control antibodies (ThermoFisher, 10500C) were used for background detection and spike-in with yeast DNA was performed to allow comparison across samples (Skene and Henikoff, 2017). Illumina sequencing library was prepared using the Illumina TruSeq ChIP Library Preparation Kit according to the manufacturer's instructions. Sequencing was performed with an Illumina Hiseq 2500 system. qPCR was performed using the LightCycler 480 (Roche) system with primers reported below. Fold enrichement at centromeres was calculated using alpha satellite or D4Z1 primers, with the ΔΔCt method. The rabbit IgG isotype sample was used as control sample and normalization was performed to the Ct values from the ALU repeat primers.

Target	**Reference**	**Forward**	**Reverse**
Alpha satellite	Hoffmann et al 2016	TCATTCCCACAA ACTGCGTTG	TCCAACGAAGGCCA CAAGA
D4Z1	Contreras-Galindo et al 2017	CTGTAGTATCTG GAAGTGGACATT	GGTTCAACTGTGTTC GTTTAGG
ALU repeat	Lou et al 2015	TACAAAAAATTA GCCGGGCG	GATCTCGGCTCACT GCAAG

Bioinformatic analysis

Read mapping was performed as previously described (Dumont et al., 2020). Briefly, reads were mapped using the bwa-mem algorithm of the BWA software package (Li, 2013; Li and Durbin, 2009) on the human reference genome GRCh38.p12 which includes centromere reference models (Miga et al., 2014; Nechemia-Arbely et al., 2017; Schneider et al., 2017).

To generate the full centromere coverage plot of Figure 2B, CENP-A CUT&RUN reads were normalized to the IgG control reads using bamCompare tool from the deepTools2 package (Ramírez *et al*, 2016) with the operation set to "ratio". Peak calling was performed using SICER1.1. (Xu et al., 2015). The centromeric peaks from the untreated and the auxin wash-out CENP-A CUT&RUN samples were compared using UCSC table browser (Karolchik et al., 2004) in two ways: first (Figure 16D) by counting the number of peaks overlapping by at least 20% of their length; second (Figure 16E) by measuring the total length in megabases of all the overlaps between untreated and auxin wash-out samples.

Read mapping to HOR arrays was carried out as previously described (Dumont et al., 2020). Briefly, the reads mapped on the centromere reference models of GRCh38 were extracted using samtools (Li and Durbin, 2009) and mapped with bwa-mem on a reference composed of 64 centromeric HOR array consensus sequences (Nechemia-Arbely et al., 2019). Due to the high similarities between the HOR sequence of acrocentric chromosomes 13, 14, 21, 22, reads could not be specifically assigned to these centromeres and they were excluded from the analysis. Read counts for cen1+5+19, cen16, cen2 and cen18 were corrected as previously described (Dumont et al., 2020) and each value was normalized to the total read number and to the spike-in control. Table 1 lists normalized read counts for each HOR consensus sequence, excluding acrocentric chromosomes 13,14,21,22.

Table 1: *Normalized read counts of CENP-A CUT&RUN-seq on each HOR consensus sequence except for acrocentric chromosomes 13, 14, 21, 22. Read counts were normalized to the spike-in control and to the total read count. Cen1+15+19 represents the consensus sequence of a HOR that is present on chromosomes 1, 15 and 19.*

	Untreated	Auxin wash out
cen1+5+19	5411	4818
cen1_2	89	78
cen1_3	21	16
cen1_4	35	30
cen1_5	13	12
cen2	1597	1465
cen3_1	1464	1327
cen3_2	92	64
cen3_3	11	8
cen4	2322	2038
cen5_1	105	68
cen5_2	29	23
cen5_3	236	213
cen5_4	25	21
cen5_5	17	12
cen5_6	25	19
cen6	1777	1564
cen7_1	2024	1664
cen7_2	62	39
cen8	1555	1491
cen9	1101	954
cen10_1	1102	1014
cen10_2	98	90
cen10_3	323	289
cen10_4	8	7
cen11_1	1712	1519
cen11_2	41	34
cen11_3	8	7
cen12_1	1453	1335
cen12_2	25	22
cen15_1	1366	1257
cen15_2	87	53
cen15_3	37	27
cen16_1	1581	1498
cen16_2	20	20
cen17_1	1148	1064
cen17_2	232	205
cen17_3	198	185
cen18_1	3123	2865

	Untreated	Auxin wash out
cen18_2	83	72
cen18_3	32	29
cen18_4	86	77
cen18_5	132	120
cen18_6	30	25
cen19_1	29	22
cen20_1	2130	2039
cen20_2	105	93
cen20_3	21	17
cen20_4	25	20
cen20_5	41	26
cen20_6	17	10
cenX	1435	1140

IF-FISH chromatin fiber

Cells were harvested for extended chromatin fiber preparation as previously described (Sullivan, 2010). Briefly, cells were harvested by trypsinization and washed in 1X PBS before dilution to $1x10^5$ cells/mL and swelling in hypotonic buffer (1:1:1 75 mM KCl: 0.8% sodium citrate: dH_2O) for 10 min. Cells were cytospun using a Shandon Cytospin 4 (ThermoFisher) onto slides at 800 rpm. Slides were then immersed in lysis buffer (25 mM Tris-HCl, pH7.5, 0.5 M NaCl, 1% Triton X-100, 0.5 M urea) for 10 min (RPE1 cells) to 17 min (DLD1 CENP-A^{AID} cells) and slowly removed from lysis buffer to stretch DNA into long fibers. Fibers were fixed in 4% PFA and permeabilized in KCM. Slides were then subjected to CENP-A immunostaining and FISH to detect the X alpha satellite array DXZ1. CENP-A antibody (ab13939; Abcam) concentration was increased to 1:75 and DXZ1 HOR DNA FISH probe (pBAMX7B) amount used at a concentration of ~500 ng per 22 mm x 40 mm slide area.

All images were acquired using an inverted Olympus IX-71 microscope controlled by the Deltavision Elite Imaging System (Applied Precision) equipped with a Photometric CoolSNAP HQ^2 CCD camera. Fiber images were collected using the 100X objective. Fibers extending through multiple fields of view were captured using the Panels option in the softWoRx Acquire 3D program and merged into single images using the 'Stitch' function. All images were exported for analysis and visualization into Adobe Photoshop and ImageJ.

The 'Measure Distances' tool of SoftWorx was used to calculate lengths of fluorescent signals representing euchromatic probe, alpha satellite probes or CENP-A immunostaining as previously

described (Sullivan et al., 2011). CENP-A domain size was measured by comparing the length of CENP-A antibody staining (in micrometers) to the length of overlapping DXZ1 alpha satellite FISH probe (pBAMX7B). Alpha satellite FISH probe signal length represented total satellite array size that had been determined by pulsed field gel electrophoresis. CENP-A chromatin domain size was calculated from the ratio of the length of CENP-A antibody signal over the total length of alpha satellite FISH signal.

Intensity quantifications in microscopy experiments

We quantified signal intensities of centromeres on interphase cells manually or automatically as described previously (Fachinetti et al., 2015, 2013). Integrated densities were determined using the software FIJI. In the case of manual quantification, a 15x15-pixel circle was drawn around the centromere (marked by CENP-B or CENP-B box staining), and an identical circle was drawn nearby (background). The integrated intensity of each centromere was calculated by subtracting the background signal from the centromeric signal. 15 centromeres per cell were averaged to determine the average centromere intensity for each cell. The same quantifications were performed using a 30x30-pixel circle to determine CENP-C and CENP-A intensities at the LacO array avoiding inclusion of adjacent native centromeres.
To quantify total nuclear signal, the nuclear area was identified via the DAPI staining. The nuclear integrated intensity of EYFP signal was calculated by subtracting the background signal measured outside the DAPI signal. To calculate the difference in signal intensity between nuclear and centromeric CENP-A, the integrated intensity of EYFP at centromeres were quantified as described above, but the background was also taken outside the DAPI signal. The sum of the measurement of all centromeres was then taken into consideration and subtracted from the total nuclear signal.

Cloning, Expression, and protein Purification

pET30-His-CENP-B (161-599) and His-CENP-B (161-404 ; 465-599) generated using the Gibson Assembly® kit (Neb) were used for protein expression in bacteria and pFB-GST-CENP-C (1-727) was used for protein expression in Sf9 cells as described previously (Fachinetti *et al*, 2015). CENP-B constructs were expressed in BL21 pLysS (Agilent) at 37°C in 2-YT medium supplemented with Kanamycin (50 µg/ml) and chloramphenicol (35 µg/ml) until OD_{600nm} between 0,6 and 0,8.

Recombinant protein expression was induced by addition of 0,5 mM IPTG (Isopropyl β- D -1-thiogalactopyranoside). Cell cultures were incubated overnight at 20°C and then harvested by centrifugation 4500 g 30 min. Cell culture pellets from bacteria or insect cells were suspended in Buffer A (His-CENP-B constructs: 20 mM Tris pH 8, 200 mM NaCl, 2 mM β-mercapto-ethanol, protease inhibitor (Roche), 1 mM phenylmethylsulfonyl fluoride (PMSF), 20 mM Imidazole pH 8, GST-CENP-C: 20 mM HEPES pH 7,5, 500 mM NaCl, 1 mM DTT, 1 mM PMSF, protease inhibitor (Roche), 0,1% Triton X-100). His-CENP-B constructs from bacteria and CENP-C Sf9 cells were lysed respectively by sonication or cell homogenizer, then centrifuged for 30 min at 45000 g. The clear lysate was loaded onto a 5 ml GST-Trap (GE Healthcare) for GST-CENP-C or a 5 ml His-Trap crude FF (GE Healthcare) for His tagged CENP-B variants connected to an Äkta Pure (GE Healthcare). Nonspecific proteins were removed by washing the column with buffer A, the proteins of interest were eluted with a linear gradient Buffer B (His-CENP-B constructs: 20 mM Tris pH 8, 200 mM NaCl, 2 mM β-mercapto-ethanol, 1 mM phenylmethylsulfonyl fluoride (PMSF), 350 mM Imidazole pH 8; GST-CENP-C: 20 mM HEPES pH 7,5, 500 mM NaCl, 1 mM DTT, 1 mM PMSF, 20 mM reduced glutathione pH 7,5). For His-CENP-B (161-599), an anion exchange chromatography was performed subsequently. The protein of interest was loaded onto a capto Q impress 1mL (GE Healthcare) connected to an Äkta Pure (GE Healthcare). Nonspecific proteins were removed by washing the column with buffer C (20 mM Tris pH 8, 200 mM NaCl, 1 mM DTT), the proteins of interest were eluted with a linear gradient Buffer D (20 mM Tris pH8, 1M NaCl, 1mM DTT). Size exclusion chromatography was then performed for all constructs on the fractions containing eluted proteins using buffer E (His-CENP-B constructs: 20 mM Tris pH 8, 200 mM NaCl, 1 mM DTT; GST-CENP-C: 20 mM HEPES pH 7,5, 500 mM NaCl, 1 mM DTT, 1 mM PMSF) on a Superdex 200 column connected to an Äkta Pure (GE Healthcare).

GST pull down assay

100 µl of glutathione sepharose beads 50% (GE healthcare) were washed 4 times with 1000 µl of interaction Buffer (20 mM HEPES pH 7,5, 300 mM KCl, 1 mM DTT, 5% glycerol). Recombinant His-CENP-B variants and/or GST-CENP-C were added at a final concentration of 2 µM in 100 µl then incubated for 1h at RT. Beads were then washed 3 times with 1000 µl of buffer A then 1 time with 1000 µl of PBS. Proteins are finally eluted with 50 µl of elution buffer (20 mM HEPES pH 7,5, 500 mM NaCl, 1 mM DTT, 5% glycerol, 20 mM reduced glutathione). 25 µl of SB5X are added to

25 µl of INPUT and Elution, boiled 5 min at 95°C then 12 µl are loaded on mini protein precast gels (Biorad). The presence of the proteins was revealed by immunoblot.

CD4$^+$ T cell staining and sorting

Peripheral blood mononuclear cells (PBMCs) were isolated from platelet apheresis blood from healthy human donors (approved by the Institut National de la Santé et de la Recherche Médicale committee) using Ficoll-Paque PLUS (GE Healthcare). Total CD4+ T cell enrichment was performed by negative selection with an EasySep Human CD4+ T Cell Isolation Kit (Stem Cell Technologies, ref. 17952).

Cell surface staining to assess CD4+ T cell enrichment. Staining was performed in FACS buffer: 1% BSA (Sigma, ref. A7030-10G) and 1 mM EDTA (ThermoFisher, ref. 15575020) in PBS. Antibodies used were: anti-human CD14 (clone M5E2, BD Biosciences, cat. 560919) and anti-human CD4 (clone RPA-T4, BD Biosciences, cat. 557871). Cells were stained for 15 min at 4 ºC and washed twice. Data were acquired on a FACSVerse flow cytometer (BD) and analyzed in FlowJo (TreeStar). CD4+ T cell enrichment was always superior to 95%.

Intracellular staining of CENP-A & CENP-B, and cell sorting. Freshly purified CD4+ T cells were washed once with PBS and stained with Fixable Viability dye eFluor 780 (eBiosciences, ref. 65-0865-14) for 30 min at 4ºC in PBS. Cells were washed twice in PBS and fixed using FoxP3 Staining buffer set (ThermoFisher, ref. 00-5523-00) according to the manufacturer's instructions. Intracellular staining was performed in permeabilization buffer using mouse anti-human CENP-A (clone 3-19, ENZO, ref. ADI-KAM-CC006-E) and rabbit anti-human CENP-B (Abcam, ref. ab25734). Mouse IgG1 kappa (eBiosciences, ref. 14-4714-82) and rabbit IgG (TermoFisher, ref. 10500C) were used as isotype controls. Cells were washed twice with permeabilization buffer and stained with secondary antibodies: F(ab')2-Goat anti-mouse IgG (H+L) (ThermoFischer, ref. A-11017) and goat anti-rabbit IgG (H+L) antibody (ThermoFishcer, ref. A-21245). Each staining was performed for 30 min at 4ºC. Cells were washed twice and resuspended in FACS buffer for FACS acquisition or cell sorting. Dead cells were gated as eFluor780+ and CENP-A & CENP-B expression was assessed in alive cells, gated as eFluor780-. Cell sorting was performed on a MoFlo Astrios (Beckman Coulter). Sorted populations purity was superior to 98%.

To track cell division, purified CD4+ T cells were labeled with carboxyfluorescein succinimidyl ester (CFSE, eBioscience, ref. 850850-84). Cells were washed once in PBS and stained with 1.6 µM

CFSE in PBS at a concentration of 10 millions cells per ml. Cells were incubated at 37ºC for 3 min in the dark, and staining was stopped by adding cold 50% serum in PBS and letting cells on ice for 3 min in the dark. Cells were centrifuged and washed with cold 50% serum and X-VIVO15 medium (Lonza, ref. BE02-060F). Cells were then plated onto 12-well plates at 1 million per ml in X-VIVO15 containing penicillin/streptomycin (Invitrogen) and 100U/ml IL-2 (ImmunoTools, ref. 11340027). Cells were activated with Dynabeads Human T-Activator CD3/CD28 (ThermoFishher, ref. 11131D) and incubated at 37 ºC for 5 days. Cells were counted every day with LUNA automated cell counter (Logos biosystems). At day 3 post-activation, media was changed and cells were plated at 1 million per ml.

Sorted CENP-A-high and low expressing cells were cytospun using a Shandon Cytospin 4 (ThermoFisher) onto coverglasses at 1500 rpm for 5 min. Cells were fixed with 4% Formaldehyde for 5 min at room temperature, DAPI counterstained and mounted using anti-fading mounting reagent (Life technologies).

Data availability

Sequencing data for the CENP-A CUT&RUN on non-treated and IAA wash-out samples, with corresponding negative controls are available at GEO (https://www.ncbi.nlm.nih.gov/geo/) under the accession numbers GSM3852804, GSE148187, GSM3852807 and GSM3852808 respectively.

10. CONTRIBUTIONS

For this work I received help from Riccardo Gamba (Insitut Curie) who performed all chromatin-protein interactions experiments and bioinformatic analysis (Figure 16A-E, 28I, 31B-C, 32D-3). Experiments on CD4+ T cells were conceived and performed together with Helena M. Izquierdo (Institut Curie) under the supervision of Nicolas Manel (Institut Curie) (Figure 33-35); Florian Chardon performed *in vitro* protein purification and interaction (Figure 26B-D); Marie Dumont performed and analyzed some of LacO/I experiments and helped in other assays (Figure 26A). Together with Veer Keizer we performed and analyzed the single molecule experiments shown in Figure 15; Shannon M. McNulty (Duke University) performed and analyzed the chromatin fiber and metaphase chromosome line scan experiments under the supervision of Beth A. Sullivan (Duke University) (Figure 16F-G).; Solene Hervé (Institut Curie) helped me with image analysis. All other

experiments were performed by me with support from Daniele Fachinetti (Figure 18 and 22). The result, methods and discussion section included in this thesis were drafted for publication in the EMBO Journal by D. Fachinetti and me. D. Fachinetti supervised me and directed the research. All authors mentioned above contributed to editing the EMBO Journal manuscript.

11. REFERENCES

Akera, T., Chmátal, L., Trimm, E., Yang, K., Aonbangkhen, C., Chenoweth, D.M., Janke, C., Schultz, R.M., Lampson, M.A., 2017. Spindle asymmetry drives non-Mendelian chromosome segregation. Science 358, 668–672. https://doi.org/10.1126/science.aan0092

Akera, T., Trimm, E., Lampson, M.A., 2019. Molecular Strategies of Meiotic Cheating by Selfish Centromeres. Cell 178, 1132-1144.e10. https://doi.org/10.1016/j.cell.2019.07.001

Akiyoshi, B., Sarangapani, K.K., Powers, A.F., Nelson, C.R., Reichow, S.L., Arellano-Santoyo, H., Gonen, T., Ranish, J.A., Asbury, C.L., Biggins, S., 2010. Tension directly stabilizes reconstituted kinetochore-microtubule attachments. Nature 468, 576–9. https://doi.org/10.1038/nature09594

Alabert, C., Groth, A., 2012. Chromatin replication and epigenome maintenance. Nat. Rev. Mol. Cell Biol. 13, 153–167. https://doi.org/10.1038/nrm3288

Aldrup-Macdonald, M.E., Sullivan, B.A., 2014. The past, present, and future of human centromere genomics. Genes 5, 33–50. https://doi.org/10.3390/genes5010033

Ali-Ahmad, A., Bilokapić, S., Schäfer, I.B., Halić, M., Sekulić, N., 2019. CENP-C unwraps the human CENP-A nucleosome through the H2A C-terminal tail. EMBO Rep. 20, e48913. https://doi.org/10.15252/embr.201948913

Alkan, C., Cardone, M.F., Catacchio, C.R., Antonacci, F., O'Brien, S.J., Ryder, O.A., Purgato, S., Zoli, M., Della Valle, G., Eichler, E.E., Ventura, M., 2011. Genome-wide characterization of centromeric satellites from multiple mammalian genomes. Genome Res. 21, 137–145. https://doi.org/10.1101/gr.111278.110

Allu, P.K., Dawicki-McKenna, J.M., Eeuwen, T.V., Slavin, M., Braitbard, M., Xu, C., Kalisman, N., Murakami, K., Black, B.E., 2019. Structure of the Human Core Centromeric Nucleosome Complex. Curr. Biol. 29, 2625-2639.e5. https://doi.org/10.1016/j.cub.2019.06.062

Alonso, A., Fritz, B., Hasson, D., Abrusan, G., Cheung, F., Yoda, K., Radlwimmer, B., Ladurner, A.G., Warburton, P.E., 2007. Co-localization of CENP-C and CENP-H to discontinuous domains of CENP-A chromatin at human neocentromeres. Genome Biol. 8, R148. https://doi.org/10.1186/gb-2007-8-7-r148

Alonso, A., Hasson, D., Cheung, F., Warburton, P.E., 2010. A paucity of heterochromatin at functional human neocentromeres. Epigenetics Chromatin 3, 6. https://doi.org/10.1186/1756-8935-3-6

Alushin, G.M., Ramey, V.H., Pasqualato, S., Ball, D.A., Grigorieff, N., Musacchio, A., Nogales, E., 2010. The Ndc80 kinetochore complex forms oligomeric arrays along microtubules. Nature 467, 805–10. https://doi.org/10.1038/nature09423

Amor, D.J., Bentley, K., Ryan, J., Perry, J., Wong, L., Slater, H., Choo, K.H.A., 2004. Human centromere repositioning "in progress." Proc. Natl. Acad. Sci. 101, 6542–6547. https://doi.org/10.1073/pnas.0308637101

Ando, S., Yang, H., Nozaki, N., Okazaki, T., Yoda, K., 2002. CENP-A, -B, and -C Chromatin Complex That Contains the I-Type α-Satellite Array Constitutes the Prekinetochore in HeLa Cells. Mol. Cell. Biol. 22, 2229–2241. https://doi.org/10.1128/mcb.22.7.2229-2241.2002

Andronov, L., Ouararhni, K., Stoll, I., Klaholz, B.P., Hamiche, A., 2019. CENP-A nucleosome clusters form rosette-like structures around HJURP during G1. Nat. Commun. 10, 4436. https://doi.org/10.1038/s41467-019-12383-3

Annunziato, A.T., Schindler, R.K., Riggs, M.G., Seale, R.L., 1982. Association of newly synthesized histones with replicating and nonreplicating regions of chromatin. J. Biol. Chem. 257, 8507–15.

Arco, Ag., Edgar, B., Erhardt, S., 2018. In Vivo Analysis of Centromeric Proteins Reveals a Stem Cell-Specific Asymmetry and an Essential Role in Differentiated, Non-proliferating Cells. Cell Rep. 22, 1982–1993. https://doi.org/10.1016/j.celrep.2018.01.079

Arimura, Y., Tachiwana, H., Takagi, H., Hori, T., Kimura, H., Fukagawa, T., Kurumizaka, H., 2019. The CENP-A centromere targeting domain facilitates H4K20 monomethylation in the nucleosome by structural polymorphism. Nat. Commun. 10, 576. https://doi.org/10.1038/s41467-019-08314-x

Aze, A., Sannino, V., Soffientini, P., Bachi, A., Costanzo, V., 2016. Centromeric DNA replication reconstitution reveals DNA loops and ATR checkpoint suppression. Nat. Cell Biol. 18, 684–691. https://doi.org/10.1038/ncb3344

Bailey, A.O., Panchenko, T., Sathyan, K.M., Petkowski, J.J., Pai, P.-J., Bai, D.L., Russell, D.H., Macara, I.G., Shabanowitz, J., Hunt, D.F., Black, B.E., Foltz, D.R., 2013. Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. Proc. Natl. Acad. Sci. U. S. A. 110, 11827–11832. https://doi.org/10.1073/pnas.1300325110

Bailey, A.O., Panchenko, T., Shabanowitz, J., Lehman, S.M., Bai, D.L., Hunt, D.F., Black, B.E., Foltz, D.R., 2016. Identification of the Post-translational Modifications Present in Centromeric Chromatin. Mol. Cell. Proteomics MCP 15, 918–931. https://doi.org/10.1074/mcp.M115.053710

Barnhart, M.C., Kuich, P.H.J.L., Stellfox, M.E., Ward, J.A., Bassett, E.A., Black, B.E., Foltz, D.R., 2011. HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. J. Cell Biol. 194, 229–243. https://doi.org/10.1083/jcb.201012017

Barnhart-Dailey, M.C., Trivedi, P., Stukenberg, P.T., Foltz, D.R., 2017. HJURP interaction with the condensin II complex during G1 promotes CENP-A deposition. Mol. Biol. Cell 28, 54–64. https://doi.org/10.1091/mbc.e15-12-0843

Barra, V., Fachinetti, D., 2018. The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. Nat. Commun. 9, 4340. https://doi.org/10.1038/s41467-018-06545-y

Barra, V., Logsdon, G.A., Scelfo, A., Hoffmann, S., Hervé, S., Aslanian, A., Nechemia-Arbely, Y., Cleveland, D.W., Black, B.E., Fachinetti, D., 2019. Phosphorylation of CENP-A on serine 7 does not control centromere function. Nat. Commun. 10, 175. https://doi.org/10.1038/s41467-018-08073-1

Basilico, F., Maffini, S., Weir, J.R., Prumbaum, D., Rojas, A.M., Zimniak, T., Antoni, A.D., Jeganathan, S., Voss, B., Gerwen, S. van, Krenn, V., Massimiliano, L., Valencia, A., Vetter, I.R., Herzog, F., Raunser, S., Pasqualato, S., Musacchio, A., 2014. The pseudo GTPase CENP-M drives human kinetochore assembly. eLife 3, e02978. https://doi.org/10.7554/elife.02978

Bassett, E.A., DeNizio, J., Barnhart-Dailey, M.C., Panchenko, T., Sekulic, N., Rogers, D.J., Foltz, D.R., Black, B.E., 2012. HJURP uses distinct CENP-A surfaces to recognize and to stabilize CENP-A/histone H4 for centromere assembly. Dev. Cell 22, 749–62. https://doi.org/10.1016/j.devcel.2012.02.001

Bassett, E.A., Wood, S., Salimian, K.J., Ajith, S., Foltz, D.R., Black, B.E., 2010. Epigenetic centromere specification directs aurora B accumulation but is insufficient to efficiently correct mitotic errors. J. Cell Biol. 190, 177–185. https://doi.org/10.1083/jcb.201001035

Basu, J., Stromberg, G., Compitello, G., Willard, H.F., Van Bokkelen, G., 2005. Rapid creation of BAC-based human artificial chromosome vectors by transposition with synthetic alpha-satellite arrays. Nucleic Acids Res. 33, 587–596. https://doi.org/10.1093/nar/gki207

Bergmann, J.H., Jakubsche, J.N., Martins, N.M., Kagansky, A., Nakano, M., Kimura, H., Kelly, D.A., Turner, B.M., Masumoto, H., Larionov, V., Earnshaw, W.C., 2012. Epigenetic engineering: histone H3K9 acetylation is compatible with kinetochore structure and function. J. Cell Sci. 125, 411–421. https://doi.org/10.1242/jcs.090639

Bergmann, J.H., Rodríguez, M.G., Martins, N.M.C., Kimura, H., Kelly, D.A., Masumoto, H., Larionov, V., Jansen, L.E.T., Earnshaw, W.C., 2011. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. EMBO J. 30, 328–340. https://doi.org/10.1038/emboj.2010.329

Bernad, R., Sánchez, P., Rivera, T., Rodríguez-Corsino, M., Boyarchuk, E., Vassias, I., Ray-Gallet, D., Arnaoutov, A., Dasso, M., Almouzni, G., Losada, A., 2011. Xenopus HJURP and condensin II are required for CENP-A assembly. J. Cell Biol. 192, 569–82. https://doi.org/10.1083/jcb.201005136

Bird, A., 2007. Perceptions of epigenetics. Nature 447, 396–398. https://doi.org/10.1038/nature05913

Black, B.E., Cleveland, D.W., 2011. Epigenetic centromere propagation and the nature of CENP-a nucleosomes. Cell 144, 471–479. https://doi.org/10.1016/j.cell.2011.02.002

Black, B.E., Foltz, D.R., Chakravarthy, S., Luger, K., Woods, V.L., Cleveland, D.W., 2004. Structural determinants for generating centromeric chromatin. Nature 430, 578–582. https://doi.org/10.1038/nature02766

Black, B.E., Jansen, L.E.T., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J.V., Cleveland, D.W., 2007. Centromere Identity Maintained by Nucleosomes Assembled with Histone H3 Containing the CENP-A Targeting Domain. Mol. Cell 25, 309–322. https://doi.org/10.1016/j.molcel.2006.12.018

Bloom, K.S., 2014. Centromeric heterochromatin: the primordial segregation machine. Annu. Rev. Genet. 48, 457–484. https://doi.org/10.1146/annurev-genet-120213-092033

Blower, M.D., 2016. Centromeric Transcription Regulates Aurora-B Localization and Activation. Cell Rep. 15, 1624–1633. https://doi.org/10.1016/j.celrep.2016.04.054

Blower, M.D., Sullivan, B.A., Karpen, G.H., 2002. Conserved organization of centromeric chromatin in flies and humans. Dev. Cell 2, 319–330. https://doi.org/10.1016/s1534-5807(02)00135-1

Bobkov, G.O., Huang, A., Berg, S.J. van den, Mitra, S., Anselm, E., Lazou, V., Schunter, S., Federle, R., Imhof, A., Lusser, A., Jansen, L.E., Heun, P., 2019. Spt6 is a maintenance factor for centromeric CENP-A. bioRxiv 560300. https://doi.org/10.1101/560300

Bobkov, G.O.M., Gilbert, N., Heun, P., 2018. Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation. J. Cell Biol. 217, 1957–1972. https://doi.org/10.1083/jcb.201611087

Bodor, D.L., Mata, J.F., Sergeev, M., David, A.F., Salimian, K.J., Panchenko, T., Cleveland, D.W., Black, B.E., Shah, J.V., Jansen, L.E., 2014. The quantitative architecture of centromeric chromatin. eLife 3, e02137. https://doi.org/10.7554/elife.02137

Bodor, D.L., Valente, L.P., Mata, J.F., Black, B.E., Jansen, L.E.T., 2013. Assembly in G1 phase and long-term stability are unique intrinsic features of CENP-A nucleosomes. Mol. Biol. Cell 24, 923–32. https://doi.org/10.1091/mbc.e13-01-0034

Boeke, J.D., Church, G., Hessel, A., Kelley, N.J., Arkin, A., Cai, Y., Carlson, R., Chakravarti, A., Cornish, V.W., Holt, L., Isaacs, F.J., Kuiken, T., Lajoie, M., Lessor, T., Lunshof, J., Maurano, M.T., Mitchell, L.A., Rine, J., Rosser, S., Sanjana, N.E., Silver, P.A., Valle, D., Wang, H., Way, J.C., Yang, L., 2016. GENOME ENGINEERING. The Genome Project-Write. Science 353, 126–127. https://doi.org/10.1126/science.aaf6850

Buchwalter, A., Kaneshiro, J.M., Hetzer, M.W., 2018. Coaching from the sidelines: the nuclear periphery in genome regulation. Nat. Rev. Genet. 20, 39–50. https://doi.org/10.1038/s41576-018-0063-5

Bui, M., Dimitriadis, E.K., Hoischen, C., An, E., Quénet, D., Giebe, S., Nita-Lazar, A., Diekmann, S., Dalal, Y., 2012. Cell-cycle-dependent structural transitions in the human CENP-A nucleosome in vivo. Cell 150, 317–326. https://doi.org/10.1016/j.cell.2012.05.035

Burgess, R.J., Zhang, Z., 2013. Histone chaperones in nucleosome assembly and human disease. Nat. Struct. Mol. Biol. 20, 14–22. https://doi.org/10.1038/nsmb.2461

Bury, L., Moodie, B., McKay, L.S., Miga, K.H., Cheeseman, I.M., 2020. Alpha-satellite RNA transcripts are repressed by centromere-nucleolus associations. bioRxiv 2020.04.14.040766. https://doi.org/10.1101/2020.04.14.040766

Camahort, R., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., Gerton, J.L., 2007. Scm3 is essential to recruit the histone h3 variant cse4 to centromeres and to maintain a functional kinetochore. Mol. Cell 26, 853–865. https://doi.org/10.1016/j.molcel.2007.05.013

Cao, S., Zhou, K., Zhang, Z., Luger, K., Straight, A.F., 2018. Constitutive centromere-associated network contacts confer differential stability on CENP-A nucleosomes in vitro and in the cell. Mol. Biol. Cell 29, mbc.E17-10-0596. https://doi.org/10.1091/mbc.e17-10-0596

Cardinale, S., Bergmann, J.H., Kelly, D., Nakano, M., Valdivia, M.M., Kimura, H., Masumoto, H., Larionov, V., Earnshaw, W.C., 2009. Hierarchical inactivation of a synthetic human kinetochore by a chromatin modifier. Mol. Biol. Cell 20, 4194–4204. https://doi.org/10.1091/mbc.e09-06-0489

Carone, D.M., Zhang, C., Hall, L.E., Obergfell, C., Carone, B.R., O'Neill, M.J., O'Neill, R.J., 2013. Hypermorphic expression of centromeric retroelement-encoded small RNAs impairs CENP-A loading. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. 21, 49–62. https://doi.org/10.1007/s10577-013-9337-0

Carroll, C.W., Milks, K.J., Straight, A.F., 2010. Dual recognition of CENP-A nucleosomes is required for centromere assembly. J. Cell Biol. 189, 1143–55. https://doi.org/10.1083/jcb.201001013

Carroll, C.W., Silva, M.C.C., Godek, K.M., Jansen, L.E.T., Straight, A.F., 2009. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. Nat. Cell Biol. 11, 896–902. https://doi.org/10.1038/ncb1899

Casola, C., Hucks, D., Feschotte, C., 2008. Convergent Domestication of pogo-like Transposases into Centromere-Binding Proteins in Fission Yeast and Mammals. Mol. Biol. Evol. 25, 29–41. https://doi.org/10.1093/molbev/msm221

Chan, F.L., Marshall, O.J., Saffery, R., Kim, B.W., Earle, E., Choo, K.H.A., Wong, L.H., 2012. Active transcription and essential role of RNA polymerase II at the centromere during mitosis. Proc. Natl. Acad. Sci. U. S. A. 109, 1979–1984. https://doi.org/10.1073/pnas.1108705109

Chan, F.L., Wong, L.H., 2012. Transcription in the maintenance of centromere chromatin identity. Nucleic Acids Res. 40, 11178–11188. https://doi.org/10.1093/nar/gks921

Cheeseman, I.M., Hori, T., Fukagawa, T., Desai, A., 2008. KNL1 and the CENP-H/I/K Complex Coordinately Direct Kinetochore Assembly in Vertebrates. Mol. Biol. Cell 19, 587–594. https://doi.org/10.1091/mbc.e07-10-1051

Chen, C.-C., Bowers, S., Lipinszki, Z., Palladino, J., Trusiak, S., Bettini, E., Rosin, L., Przewloka, M.R., Glover, D.M., O'Neill, R.J., Mellone, B.G., 2015. Establishment of Centromeric Chromatin by the CENP-A Assembly Factor CAL1 Requires FACT-Mediated Transcription. Dev. Cell 34, 73–84. https://doi.org/10.1016/j.devcel.2015.05.012

Chittori, S., Hong, J., Saunders, H., Feng, H., Ghirlando, R., Kelly, A.E., Bai, Y., Subramaniam, S., 2017. Structural mechanisms of centromeric nucleosome recognition by the kinetochore protein CENP-N. Science 359, 339–343. https://doi.org/10.1126/science.aar2781

Chmátal, L., Gabriel, S.I., Mitsainas, G.P., Martínez-Vargas, J., Ventura, J., Searle, J.B., Schultz, R.M., Lampson, M.A., 2014. Centromere strength provides the cell biological basis for meiotic drive and karyotype evolution in mice. Curr. Biol. CB 24, 2295–2300. https://doi.org/10.1016/j.cub.2014.08.017

Chunduri, N.K., Storchová, Z., 2019. The diverse consequences of aneuploidy. Nat. Cell Biol. 21, 54–62. https://doi.org/10.1038/s41556-018-0243-8

Cohen, R.L., Espelin, C.W., Wulf, P.D., Sorger, P.K., Harrison, S.C., Simons, K.T., 2008. Structural and functional dissection of Mif2p, a conserved DNA-binding kinetochore protein. Mol. Biol. Cell 19, 4480–91. https://doi.org/10.1091/mbc.e08-03-0297

Cooke, C.A., Bernat, R.L., Earnshaw, W.C., 1990. CENP-B: a major human centromere protein located beneath the kinetochore. J. Cell Biol. 110, 1475–1488. https://doi.org/10.1083/jcb.110.5.1475

Crasta, K., Ganem, N.J., Dagher, R., Lantermann, A.B., Ivanova, E.V., Pan, Y., Nezi, L., Protopopov, A., Chowdhury, D., Pellman, D., 2012. DNA breaks and chromosome pulverization from errors in mitosis. Nature 482, 53–8. https://doi.org/10.1038/nature10802

Cutter, A.R., Hayes, J.J., 2015. A brief review of nucleosome structure. FEBS Lett. 589, 2914–22. https://doi.org/10.1016/j.febslet.2015.05.016

Dai, J., Higgins, J.M.G., 2005. Haspin: a mitotic histone kinase required for metaphase chromosome alignment. Cell Cycle Georget. Tex 4, 665–668. https://doi.org/10.4161/cc.4.5.1683

Dai, X., Otake, K., You, C., Cai, Q., Wang, Z., Masumoto, H., Wang, Y., 2013. Identification of novel α-n-methylation of CENP-B that regulates its binding to the centromeric DNA. J. Proteome Res. 12, 4167–4175. https://doi.org/10.1021/pr400498y

Dalal, Y., Wang, H., Lindsay, S., Henikoff, S., 2007. Tetrameric Structure of Centromeric Nucleosomes in Interphase Drosophila Cells. PLoS Biol. 5, e218. https://doi.org/10.1371/journal.pbio.0050218

Dambacher, S., Deng, W., Hahn, M., Sadic, D., Fröhlich, J., Nuber, A., Hoischen, C., Diekmann, S., Leonhardt, H., Schotta, G., 2012. CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin. Nucl. Austin Tex 3, 101–10. https://doi.org/10.4161/nucl.18955

Dawe, R.K., Lowry, E.G., Gent, J.I., Stitzer, M.C., Swentowsky, K.W., Higgins, D.M., Ross-Ibarra, J., Wallace, J.G., Kanizay, L.B., Alabady, M., Qiu, W., Tseng, K.-F., Wang, N., Gao, Z., Birchler, J.A., Harkess, A.E., Hodges, A.L., Hiatt, E.N., 2018. A Kinesin-14 Motor Activates Neocentromeres to Promote Meiotic Drive in Maize. Cell 173, 839-850.e18. https://doi.org/10.1016/j.cell.2018.03.009

Déjardin, J., 2015. Switching between Epigenetic States at Pericentromeric Heterochromatin. Trends Genet. TIG 31, 661–672. https://doi.org/10.1016/j.tig.2015.09.003

Demirdizen, E., Spiller-Becker, M., Förtsch, A., Wilhelm, A., Corless, S., Bade, D., Bergner, A., Hessling, B., Erhardt, S., 2019. Localization of Drosophila CENP-A to non-centromeric sites depends on the NuRD complex. Nucleic Acids Res. 47, 11589–11608. https://doi.org/10.1093/nar/gkz962

Ding, M., Jiang, J., Yang, F., Zheng, F., Fang, J., Wang, Q., Wang, J., Yao, W., Liu, Xu, Gao, X., Mullen, M., He, P., Rono, C., Ding, X., Hong, J., Fu, C., Liu, Xing, Yao, X., 2019. Holliday junction recognition protein interacts with and specifies the centromeric assembly of CENP-T. J. Biol. Chem. 294, 968–980. https://doi.org/10.1074/jbc.ra118.004688

Drpic, D., Almeida, A.C., Aguiar, P., Renda, F., Damas, J., Lewin, H.A., Larkin, D.M., Khodjakov, A., Maiato, H., 2018. Chromosome Segregation Is Biased by Kinetochore Size. Curr. Biol. CB 28, 1344-1356.e5. https://doi.org/10.1016/j.cub.2018.03.023

Du, Y., Topp, C.N., Dawe, R.K., 2010. DNA Binding of Centromere Protein C (CENPC) Is Stabilized by Single-Stranded RNA. PLoS Genet. 6, e1000835. https://doi.org/10.1371/journal.pgen.1000835

Dumont, M., Fachinetti, D., 2017. DNA Sequences in Centromere Formation and Function. Prog. Mol. Subcell. Biol. 56, 305–336. https://doi.org/10.1007/978-3-319-58592-5_13

Dumont, M., Gamba, R., Gestraud, P., Klaasen, S., Worrall, J.T., Vries, S.G.D., Boudreau, V., Salinas-Luypaert, C., Maddox, P.S., Lens, S.M., Kops, G.J., McClelland, S.E., Miga, K.H., Fachinetti, D., 2020. Human chromosome-specific aneuploidy is influenced by DNA-dependent centromeric features. EMBO J. 39. https://doi.org/10.15252/embj.2019102924

Dunleavy, E.M., Almouzni, G., Karpen, G.H., 2011. H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G_1 phase. Nucl. Austin Tex 2, 146–57. https://doi.org/10.4161/nucl.2.2.15211

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., Almouzni-Pettinotti, G., 2009. HJURP Is a Cell-Cycle-Dependent Maintenance and Deposition Factor of CENP-A at Centromeres. Cell 137, 485–497. https://doi.org/10.1016/j.cell.2009.02.040

Earnshaw, W.C., Machlin, P.S., Bordwell, B.J., Rothfield, N.F., Cleveland, D.W., 1987. Analysis of anticentromere autoantibodies using cloned autoantigen CENP-B. Proc. Natl. Acad. Sci. U. S. A. 84, 4979–4983. https://doi.org/10.1073/pnas.84.14.4979

Earnshaw, W.C., Migeon, B.R., 1985. Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. Chromosoma 92, 290–296. https://doi.org/10.1007/BF00329812

Earnshaw, W.C., Ratrie, H., Stetten, G., 1989. Visualization of centromere proteins CENP-B and CENP-C on a stable dicentric chromosome in cytological spreads. Chromosoma 98, 1–12. https://doi.org/10.1007/BF00293329

Earnshaw, W.C., Rothfield, N., 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. Chromosoma 91, 313–321. https://doi.org/10.1007/bf00328227

Eot-Houllier, G., Magnaghi-Jaulin, L., Fulcrand, G., Moyroud, F.-X., Monier, S., Jaulin, C., 2018. Aurora A-dependent CENP-A phosphorylation at inner centromeres protects bioriented chromosomes against cohesion fatigue. Nat. Commun. 9, 1888. https://doi.org/10.1038/s41467-018-04089-9

Erhardt, S., Mellone, B.G., Betts, C.M., Zhang, W., Karpen, G.H., Straight, A.F., 2008. Genome-wide analysis reveals a cell cycle-dependent mechanism controlling centromere propagation. J. Cell Biol. 183, 805–18. https://doi.org/10.1083/jcb.200806038

Erliandri, I., Fu, H., Nakano, M., Kim, J.-H., Miga, K.H., Liskovykh, M., Earnshaw, W.C., Masumoto, H., Kouprina, N., Aladjem, M.I., Larionov, V., 2014. Replication of alpha-satellite DNA arrays in endogenous human centromeric regions and in human artificial chromosome. Nucleic Acids Res. 42, 11502–11516. https://doi.org/10.1093/nar/gku835

Fachinetti, D., Folco, H.D., Nechemia-Arbely, Y., Valente, L.P., Nguyen, K., Wong, A.J., Zhu, Q., Holland, A.J., Desai, A., Jansen, L.E.T., Cleveland, D.W., 2013. A two-step mechanism for epigenetic specification of centromere identity and function. Nat. Cell Biol. 15, 1056–1066. https://doi.org/10.1038/ncb2805

Fachinetti, D., Han, J.S., McMahon, M.A., Ly, P., Abdullah, A., Wong, A.J., Cleveland, D.W., 2015. DNA Sequence-Specific Binding of CENP-B Enhances the Fidelity of Human Centromere Function. Dev. Cell 33, 314–327. https://doi.org/10.1016/j.devcel.2015.03.020

Fachinetti, D., Logsdon, G.A., Abdullah, A., Selzer, E.B., Cleveland, D.W., Black, B.E., 2017. CENP-A Modifications on Ser68 and Lys124 Are Dispensable for Establishment, Maintenance, and Long-Term Function of Human Centromeres. Dev. Cell 40, 104–113. https://doi.org/10.1016/j.devcel.2016.12.014

Falk, S.J., Guo, L.Y., Sekulic, N., Smoak, E.M., Mani, T., Logsdon, G.A., Gupta, K., Jansen, L.E.T., Duyne, G.D.V., Vinogradov, S.A., Lampson, M.A., Black, B.E., 2015. CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. Science 348, 699–703. https://doi.org/10.1126/science.1259308

Falk, S.J., Lee, J., Sekulic, N., Sennett, M.A., Lee, T.-H., Black, B.E., 2016. CENP-C directs a structural transition of CENP-A nucleosomes mainly through sliding of DNA gyres. Nat. Struct. Mol. Biol. 23, 204–208. https://doi.org/10.1038/nsmb.3175

Fang, J., Liu, Y., Wei, Y., Deng, W., Yu, Z., Huang, L., Teng, Y., Yao, T., You, Q., Ruan, H., Chen, P., Xu, R.-M., Li, G., 2015. Structural transitions of centromeric chromatin regulate the cell

cycle-dependent recruitment of CENP-N. Genes Dev. 29, 1058–1073. https://doi.org/10.1101/gad.259432.115
Ferri, F., Bouzinba-Segard, H., Velasco, G., Hubé, F., Francastel, C., 2009. Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. Nucleic Acids Res. 37, 5071–5080. https://doi.org/10.1093/nar/gkp529
Fishman, L., Kelly, J.K., 2015. Centromere-associated meiotic drive and female fitness variation in Mimulus. Evol. Int. J. Org. Evol. 69, 1208–1218. https://doi.org/10.1111/evo.12661
Flemming, W., 1882. Zellsubstanz, Kern und Zelltheilung /. https://doi.org/10.5962/bhl.title.168645
Folco, H.D., Campbell, C.S., May, K.M., Espinoza, C.A., Oegema, K., Hardwick, K.G., Grewal, S.I.S., Desai, A., 2015. The CENP-A N-tail confers epigenetic stability to centromeres via the CENP-T branch of the CCAN in fission yeast. Curr. Biol. CB 25, 348–356. https://doi.org/10.1016/j.cub.2014.11.060
Foltz, D.R., Jansen, L.E.T., Bailey, A.O., Yates, J.R., Bassett, E.A., Wood, S., Black, B.E., Cleveland, D.W., 2009. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell 137, 472–84. https://doi.org/10.1016/j.cell.2009.02.039
Foltz, D.R., Jansen, L.E.T., Black, B.E., Bailey, A.O., Yates, J.R., Cleveland, D.W., 2006. The human CENP-A centromeric nucleosome-associated complex. Nat. Cell Biol. 8, 458–469. https://doi.org/10.1038/ncb1397
Forsberg, L.A., Rasi, C., Malmqvist, N., Davies, H., Pasupulati, S., Pakalapati, G., Sandgren, J., Diaz de Ståhl, T., Zaghlool, A., Giedraitis, V., Lannfelt, L., Score, J., Cross, N.C.P., Absher, D., Janson, E.T., Lindgren, C.M., Morris, A.P., Ingelsson, E., Lind, L., Dumanski, J.P., 2014. Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. Nat. Genet. 46, 624–628. https://doi.org/10.1038/ng.2966
Fowler, K.J., Hudson, D.F., Salamonsen, L.A., Edmondson, S.R., Earle, E., Sibson, M.C., Choo, K.H., 2000. Uterine dysfunction and genetic modifiers in centromere protein B-deficient mice. Genome Res. 10, 30–41.
French, B.T., Westhorpe, F.G., Limouse, C., Straight, A.F., 2017. Xenopus laevis M18BP1 Directly Binds Existing CENP-A Nucleosomes to Promote Centromeric Chromatin Assembly. Dev. Cell 42, 190-199.e10. https://doi.org/10.1016/j.devcel.2017.06.021
Fujita, R., Otake, K., Arimura, Y., Horikoshi, N., Miya, Y., Shiga, T., Osakabe, A., Tachiwana, H., Ohzeki, J., Larionov, V., Masumoto, H., Kurumizaka, H., 2015. Stable complex formation of CENP-B with the CENP-A nucleosome. Nucleic Acids Res. 43, 4909–4922. https://doi.org/10.1093/nar/gkv405
Fujita, Y., Hayashi, T., Kiyomitsu, T., Toyoda, Y., Kokubu, A., Obuse, C., Yanagida, M., 2007. Priming of Centromere for CENP-A Recruitment by Human hMis18α, hMis18β, and M18BP1. Dev. Cell 12, 17–30. https://doi.org/10.1016/j.devcel.2006.11.002
Fukagawa, T., Earnshaw, W.C., 2014. The centromere: chromatin foundation for the kinetochore machinery. Dev. Cell 30, 496–508. https://doi.org/10.1016/j.devcel.2014.08.016
Fukagawa, T., Mikami, Y., Nishihashi, A., Regnier, V., Haraguchi, T., Hiraoka, Y., Sugata, N., Todokoro, K., Brown, W., Ikemura, T., 2001. CENP-H, a constitutive centromere component, is required for centromere targeting of CENP-C in vertebrate cells. EMBO J. 20, 4603–4617. https://doi.org/10.1093/emboj/20.16.4603
Fukagawa, T., Pendon, C., Morris, J., Brown, W., 1999. CENP-C is necessary but not sufficient to induce formation of a functional centromere. EMBO J. 18, 4196–4209. https://doi.org/10.1093/emboj/18.15.4196
Gallego, J., Chou, S.H., Reid, B.R., 1997. Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel T:G:G:T tetrad: solution structure of the d(TCCCGTTTCCA) dimer. J. Mol. Biol. 273, 840–856. https://doi.org/10.1006/jmbi.1997.1361

Gamba, R., Fachinetti, D., 2020. From evolution to function: Two sides of the same CENP-B coin? Exp. Cell Res. 390, 111959. https://doi.org/10.1016/j.yexcr.2020.111959
Ganem, N.J., Godinho, S.A., Pellman, D., 2009. A mechanism linking extra centrosomes to chromosomal instability. Nature 460, 278–82. https://doi.org/10.1038/nature08136
Garavís, M., Escaja, N., Gabelica, V., Villasante, A., González, C., 2015a. Centromeric Alpha-Satellite DNA Adopts Dimeric i-Motif Structures Capped by AT Hoogsteen Base Pairs. Chem. Weinh. Bergstr. Ger. 21, 9816–9824. https://doi.org/10.1002/chem.201500448
Garavís, M., Méndez-Lago, M., Gabelica, V., Whitehead, S.L., González, C., Villasante, A., 2015b. The structure of an endogenous Drosophila centromere reveals the prevalence of tandemly repeated sequences able to form i-motifs. Sci. Rep. 5, 13307. https://doi.org/10.1038/srep13307
García Del Arco, A., Erhardt, S., 2017. Post-translational Modifications of Centromeric Chromatin. Prog. Mol. Subcell. Biol. 56, 213–231. https://doi.org/10.1007/978-3-319-58592-5_9
Gascoigne, K.E., Takeuchi, K., Suzuki, A., Hori, T., Fukagawa, T., Cheeseman, I.M., 2011. Induced Ectopic Kinetochore Assembly Bypasses the Requirement for CENP-A Nucleosomes. Cell 145, 410–422. https://doi.org/10.1016/j.cell.2011.03.031
Gassmann, R., Rechtsteiner, A., Yuen, K.W., Muroyama, A., Egelhofer, T., Gaydos, L., Barron, F., Maddox, P., Essex, A., Monen, J., Ercan, S., Lieb, J.D., Oegema, K., Strome, S., Desai, A., 2012. An inverse relationship to germline transcription defines centromeric chromatin in C. elegans. Nature 484, 534–537. https://doi.org/10.1038/nature10973
Gérard, A., Koundrioukoff, S., Ramillon, V., Sergère, J., Mailand, N., Quivy, J., Almouzni, G., 2006. The replication kinase Cdc7-Dbf4 promotes the interaction of the p150 subunit of chromatin assembly factor 1 with proliferating cell nuclear antigen. EMBO Rep. 7, 817–823. https://doi.org/10.1038/sj.embor.7400750
Giulotto, E., Raimondi, E., Sullivan, K.F., 2017. The Unique DNA Sequences Underlying Equine Centromeres. Prog. Mol. Subcell. Biol. 56, 337–354. https://doi.org/10.1007/978-3-319-58592-5_14
Giunta, S., Funabiki, H., 2017. Integrity of the human centromere DNA repeats is protected by CENP-A, CENP-C, and CENP-T. Proc. Natl. Acad. Sci. U. S. A. 114, 1928–1933. https://doi.org/10.1073/pnas.1615133114
Goutte-Gattat, D., Shuaib, M., Ouararhni, K., Gautier, T., Skoufias, D.A., Hamiche, A., Dimitrov, S., 2013. Phosphorylation of the CENP-A amino-terminus in mitotic centromeric chromatin is required for kinetochore function. Proc. Natl. Acad. Sci. U. S. A. 110, 8579–8584. https://doi.org/10.1073/pnas.1302955110
Grenfell, A.W., Heald, R., Strzelecka, M., 2016. Mitotic noncoding RNA processing promotes kinetochore and spindle assembly in Xenopus. J. Cell Biol. 214, 133–141. https://doi.org/10.1083/jcb.201604029
Gruenbaum, Y., Foisner, R., 2015. Lamins: nuclear intermediate filament proteins with fundamental functions in nuclear mechanics and genome regulation. Annu. Rev. Biochem. 84, 131–64. https://doi.org/10.1146/annurev-biochem-060614-034115
Guldner, H.H., Lakomek, H.J., Bautz, F.A., 1984. Human anti-centromere sera recognise a 19.5 kD non-histone chromosomal protein from HeLa cells. Clin. Exp. Immunol. 58, 13–20.
Guo, L.Y., Allu, P.K., Zandarashvili, L., McKinley, K.L., Sekulic, N., Dawicki-McKenna, J.M., Fachinetti, D., Logsdon, G.A., Jamiolkowski, R.M., Cleveland, D.W., Cheeseman, I.M., Black, B.E., 2017. Centromeres are maintained by fastening CENP-A to DNA and directing an arginine anchor-dependent nucleosome transition. Nat. Commun. 8, 15775. https://doi.org/10.1038/ncomms15775
Guse, A., Carroll, C.W., Moree, B., Fuller, C.J., Straight, A.F., 2011. In vitro centromere and kinetochore assembly on defined chromatin templates. Nature 477, 354–8. https://doi.org/10.1038/nature10379

Haaf, T., Schmid, M., 1991. Chromosome topology in mammalian interphase nuclei. Exp. Cell Res. 192, 325–332. https://doi.org/10.1016/0014-4827(91)90048-y
Haaf, T., Warburton, P.E., Willard, H.F., 1992. Integration of human alpha-satellite DNA into simian chromosomes: centromere protein binding and disruption of normal chromosome segregation. Cell 70, 681–696. https://doi.org/10.1016/0092-8674(92)90436-g
Haarhuis, J.H.I., Elbatsh, A.M.O., Rowland, B.D., 2014. Cohesin and its regulation: on the logic of X-shaped chromosomes. Dev. Cell 31, 7–18. https://doi.org/10.1016/j.devcel.2014.09.010
Haase, J., Mishra, P.K., Stephens, A., Haggerty, R., Quammen, C., Taylor, R.M., Yeh, E., Basrai, M.A., Bloom, K., 2013. A 3D map of the yeast kinetochore reveals the presence of core and accessory centromere-specific histone. Curr. Biol. CB 23, 1939–44. https://doi.org/10.1016/j.cub.2013.07.083
Han, F., Gao, Z., Birchler, J.A., 2009. Reactivation of an inactive centromere reveals epigenetic and structural components for centromere specification in maize. Plant Cell 21, 1929–39. https://doi.org/10.1105/tpc.109.066662
Hara, M., Fukagawa, T., 2018. Kinetochore assembly and disassembly during mitotic entry and exit. Curr. Opin. Cell Biol. 52, 73–81. https://doi.org/10.1016/j.ceb.2018.02.005
Harkes, R., Keizer, V.I.P., Schaaf, M.J.M., Schmidt, T., 2015. Depth-of-Focus Correction in Single-Molecule Data Allows Analysis of 3D Diffusion of the Glucocorticoid Receptor in the Nucleus. PloS One 10, e0141080. https://doi.org/10.1371/journal.pone.0141080
Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., Willard, H.F., 1997. Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nat. Genet. 15, 345–355. https://doi.org/10.1038/ng0497-345
Hasson, D., Alonso, A., Cheung, F., Tepperberg, J.H., Papenhausen, P.R., Engelen, J.J.M., Warburton, P.E., 2011. Formation of novel CENP-A domains on tandem repetitive DNA and across chromosome breakpoints on human chromosome 8q21 neocentromeres. Chromosoma 120, 621–32. https://doi.org/10.1007/s00412-011-0337-6
Hasson, D., Panchenko, T., Salimian, K.J., Salman, M.U., Sekulic, N., Alonso, A., Warburton, P.E., Black, B.E., 2013. The octamer is the major form of CENP-A nucleosomes at human centromeres. Nat. Struct. Mol. Biol. 20, 687–95. https://doi.org/10.1038/nsmb.2562
Hatch, E.M., Fischer, A.H., Deerinck, T.J., Hetzer, M.W., 2013. Catastrophic nuclear envelope collapse in cancer cell micronuclei. Cell 154, 47–60. https://doi.org/10.1016/j.cell.2013.06.007
Hayashi, T., Fujita, Y., Iwasaki, O., Adachi, Y., Takahashi, K., Yanagida, M., 2004. Mis16 and Mis18 Are Required for CENP-A Loading and Histone Deacetylation at Centromeres. Cell 118, 715–729. https://doi.org/10.1016/j.cell.2004.09.002
Hayden, K.E., Strome, E.D., Merrett, S.L., Lee, H.-R., Rudd, M.K., Willard, H.F., 2013. Sequences associated with centromere competency in the human genome. Mol. Cell. Biol. 33, 763–772. https://doi.org/10.1128/MCB.01198-12
Hellwig, D., Münch, S., Orthaus, S., Hoischen, C., Hemmerich, P., Diekmann, S., 2008. Live-cell imaging reveals sustained centromere binding of CENP-T via CENP-A and CENP-B. J. Biophotonics 1, 245–254. https://doi.org/10.1002/jbio.200810014
Hemmerich, P., Weidtkamp-Peters, S., Hoischen, C., Schmiedeberg, L., Erliandri, I., Diekmann, S., 2008. Dynamics of inner kinetochore assembly and maintenance in living cells. J. Cell Biol. 180, 1101–1114. https://doi.org/10.1083/jcb.200710052
Henikoff, S., Ahmad, K., Malik, H.S., 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293, 1098–1102. https://doi.org/10.1126/science.1062939
Henikoff, S., Ahmad, K., Platero, J.S., Steensel, B. van, 2000. Heterochromatic deposition of centromeric histone H3-like proteins. Proc. Natl. Acad. Sci. 97, 716–721. https://doi.org/10.1073/pnas.97.2.716

Henikoff, S., Ramachandran, S., Krassovsky, K., Bryson, T.D., Codomo, C.A., Brogaard, K., Widom, J., Wang, J.-P., Henikoff, J.G., 2014. The budding yeast Centromere DNA Element II wraps a stable Cse4 hemisome in either orientation in vivo. eLife 3, e01861. https://doi.org/10.7554/elife.01861

Henikoff, S., Smith, M.M., 2015. Histone Variants and Epigenetics. Cold Spring Harb. Perspect. Biol. 7, a019364. https://doi.org/10.1101/cshperspect.a019364

Heun, P., Erhardt, S., Blower, M., Weiss, S., Skora, A., Karpen, G., 2006. Mislocalization of the Drosophila Centromere-Specific Histone CID Promotes Formation of Functional Ectopic Kinetochores. Dev. Cell 10, 303–315. https://doi.org/10.1016/j.devcel.2006.01.014

Higgins, A.W., Gustashaw, K.M., Willard, H.F., 2005. Engineered human dicentric chromosomes show centromere plasticity. Chromosome Res. 13, 745–762. https://doi.org/10.1007/s10577-005-1009-2

Hill, A., Bloom, K., 1987. Genetic manipulation of centromere function. Mol. Cell. Biol. 7, 2397–2405. https://doi.org/10.1128/mcb.7.7.2397

Hochegger, H., Dejsuphong, D., Sonoda, E., Saberi, A., Rajendra, E., Kirk, J., Hunt, T., Takeda, S., 2007. An essential role for Cdk1 in S phase control is revealed via chemical genetics in vertebrate cells. J. Cell Biol. 178, 257–268. https://doi.org/10.1083/jcb.200702034

Hoffmann, S., Dumont, M., Barra, V., Ly, P., Nechemia-Arbely, Y., McMahon, M.A., Hervé, S., Cleveland, D.W., Fachinetti, D., 2016. CENP-A Is Dispensable for Mitotic Centromere Function after Initial Centromere/Kinetochore Assembly. Cell Rep. 17, 2394–2404. https://doi.org/10.1016/j.celrep.2016.10.084

Hoffmann, S., Fachinetti, D., 2018. Real-Time De Novo Deposition of Centromeric Histone-Associated Proteins Using the Auxin-Inducible Degradation System. Methods Mol. Biol. Clifton NJ 1832, 223–241. https://doi.org/10.1007/978-1-4939-8663-7_12

Holland, A.J., Fachinetti, D., Han, J.S., Cleveland, D.W., 2012. Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. 109, E3350-3357. https://doi.org/10.1073/pnas.1216880109

Hooser, A.A.V., Ouspenski, I.I., Gregson, H.C., Starr, D.A., Yen, T.J., Goldberg, M.L., Yokomori, K., Earnshaw, W.C., Sullivan, K.F., Brinkley, B.R., 2001. Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. J. Cell Sci. 114, 3529–42.

Hori, T., Amano, M., Suzuki, A., Backer, C.B., Welburn, J.P., Dong, Y., McEwen, B.F., Shang, W.-H., Suzuki, E., Okawa, K., Cheeseman, I.M., Fukagawa, T., 2008a. CCAN Makes Multiple Contacts with Centromeric DNA to Provide Distinct Pathways to the Outer Kinetochore. Cell 135, 1039–1052. https://doi.org/10.1016/j.cell.2008.10.019

Hori, T., Haraguchi, T., Hiraoka, Y., Kimura, H., Fukagawa, T., 2003. Dynamic behavior of Nuf2-Hec1 complex that localizes to the centrosome and centromere and is essential for mitotic progression in vertebrate cells. J. Cell Sci. 116, 3347–3362. https://doi.org/10.1242/jcs.00645

Hori, T., Kagawa, N., Toyoda, A., Fujiyama, A., Misu, S., Monma, N., Makino, F., Ikeo, K., Fukagawa, T., 2016. Constitutive centromere-associated network controls centromere drift in vertebrate cells. J. Cell Biol. 216, 101–113. https://doi.org/10.1083/jcb.201605001

Hori, T., Okada, M., Maenaka, K., Fukagawa, T., 2008b. CENP-O Class Proteins Form a Stable Complex and Are Required for Proper Kinetochore Function. Mol. Biol. Cell 19, 843–854. https://doi.org/10.1091/mbc.e07-06-0556

Hori, T., Shang, W.-H., Hara, M., Ariyoshi, M., Arimura, Y., Fujita, R., Kurumizaka, H., Fukagawa, T., 2017. Association of M18BP1/KNL2 with CENP-A Nucleosome Is Essential for Centromere Formation in Non-mammalian Vertebrates. Dev. Cell 42, 181-189.e3. https://doi.org/10.1016/j.devcel.2017.06.019

Hori, T., Shang, W.-H., Takeuchi, K., Fukagawa, T., 2012. The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly. J. Cell Biol. 200, 45–60. https://doi.org/10.1083/jcb.201210106

Hori, T., Shang, W.-H., Toyoda, A., Misu, S., Monma, N., Ikeo, K., Molina, O., Vargiu, G., Fujiyama, A., Kimura, H., Earnshaw, W.C., Fukagawa, T., 2014. Histone H4 Lys 20 monomethylation of the CENP-A nucleosome is essential for kinetochore assembly. Dev. Cell 29, 740–749. https://doi.org/10.1016/j.devcel.2014.05.001

Howman, E.V., Fowler, K.J., Newson, A.J., Redward, S., MacDonald, A.C., Kalitsis, P., Choo, K.H., 2000. Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. Proc. Natl. Acad. Sci. U. S. A. 97, 1148–1153. https://doi.org/10.1073/pnas.97.3.1148

Hoyt, M.A., 2001. A new view of the spindle checkpoint. J. Cell Biol. 154, 909–912. https://doi.org/10.1083/jcb.200108010

Hu, H., Liu, Y., Wang, M., Fang, J., Huang, H., Yang, N., Li, Y., Wang, J., Yao, X., Shi, Y., Li, G., Xu, R.-M., 2011. Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP. Genes Dev. 25, 901–906. https://doi.org/10.1101/gad.2045111

Hu, Z., Huang, G., Sadanandam, A., Gu, S., Lenburg, M.E., Pai, M., Bayani, N., Blakely, E.A., Gray, J.W., Mao, J.-H., 2010. The expression level of HJURP has an independent prognostic impact and predicts the sensitivity to radiotherapy in breast cancer. Breast Cancer Res. BCR 12, R18. https://doi.org/10.1186/bcr2487

Huang, C., Cheng, J., Bawa-Khalfe, T., Yao, X., Chin, Y.E., Yeh, E.T.H., 2016. SUMOylated ORC2 Recruits a Histone Demethylase to Regulate Centromeric Histone Modification and Genomic Stability. Cell Rep. 15, 147–157. https://doi.org/10.1016/j.celrep.2016.02.091

Hudson, D.F., Fowler, K.J., Earle, E., Saffery, R., Kalitsis, P., Trowell, H., Hill, J., Wreford, N.G., de Kretser, D.M., Cancilla, M.R., Howman, E., Hii, L., Cutts, S.M., Irvine, D.V., Choo, K.H., 1998. Centromere protein B null mice are mitotically and meiotically normal but have lower body and testis weights. J. Cell Biol. 141, 309–319. https://doi.org/10.1083/jcb.141.2.309

Ideue, T., Cho, Y., Nishimura, K., Tani, T., 2014. Involvement of satellite I noncoding RNA in regulation of chromosome segregation. Genes Cells Devoted Mol. Cell. Mech. 19, 528–538. https://doi.org/10.1111/gtc.12149

Ikeno, M., Grimes, B., Okazaki, T., Nakano, M., Saitoh, K., Hoshino, H., McGill, N.I., Cooke, H., Masumoto, H., 1998. Construction of YAC-based mammalian artificial chromosomes. Nat. Biotechnol. 16, 431–439. https://doi.org/10.1038/nbt0598-431

Ishii, K., Ogiyama, Y., Chikashige, Y., Soejima, S., Masuda, F., Kakuma, T., Hiraoka, Y., Takahashi, K., 2008. Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. Science 321, 1088–91. https://doi.org/10.1126/science.1158699

Iwamoto, M., Björklund, T., Lundberg, C., Kirik, D., Wandless, T.J., 2010. A general chemical method to regulate protein stability in the mammalian central nervous system. Chem. Biol. 17, 981–988. https://doi.org/10.1016/j.chembiol.2010.07.009

Iwata-Otsubo, A., Dawicki-McKenna, J.M., Akera, T., Falk, S.J., Chmátal, L., Yang, K., Sullivan, B.A., Schultz, R.M., Lampson, M.A., Black, B.E., 2017. Expanded Satellite Repeats Amplify a Discrete CENP-A Nucleosome Assembly Site on Chromosomes that Drive in Female Meiosis. Curr. Biol. CB 27, 2365-2373.e8. https://doi.org/10.1016/j.cub.2017.06.069

Izuta, H., Ikeno, M., Suzuki, N., Tomonaga, T., Nozaki, N., Obuse, C., Kisu, Y., Goshima, N., Nomura, F., Nomura, N., Yoda, K., 2006. Comprehensive analysis of the ICEN (Interphase Centromere Complex) components enriched in the CENP-A chromatin of human cells. Genes Cells 11, 673–684. https://doi.org/10.1111/j.1365-2443.2006.00969.x

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H., Spector, D.L., 2004. From Silencing to Gene Expression Real-Time Analysis in Single Cells. Cell 116, 683–698. https://doi.org/10.1016/s0092-8674(04)00171-0

Jansen, L.E.T., Black, B.E., Foltz, D.R., Cleveland, D.W., 2007a. Propagation of centromeric chromatin requires exit from mitosis. J. Cell Biol. 176, 795–805. https://doi.org/10.1083/jcb.200701066
Jansen, L.E.T., Black, B.E., Foltz, D.R., Cleveland, D.W., 2007b. Propagation of centromeric chromatin requires exit from mitosis. J. Cell Biol. 176, 795–805. https://doi.org/10.1083/jcb.200701066
Janssen, A., Burg, M. van der, Szuhai, K., Kops, G.J.P.L., Medema, R.H., 2011. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. Science 333, 1895–8. https://doi.org/10.1126/science.1210214
Kabeche, L., Nguyen, H.D., Buisson, R., Zou, L., 2018. A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. Science 359, 108–114. https://doi.org/10.1126/science.aan6490
Kapoor, M., Montes de Oca Luna, R., Liu, G., Lozano, G., Cummings, C., Mancini, M., Ouspenski, I., Brinkley, B.R., May, G.S., 1998. The cenpB gene is not essential in mice. Chromosoma 107, 570–576. https://doi.org/10.1007/s004120050343
Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., Kent, W.J., 2004. The UCSC Table Browser data retrieval tool. Nucleic Acids Res. 32, D493-496. https://doi.org/10.1093/nar/gkh103
Kasinathan, S., Henikoff, S., 2018. Non-B-form DNA is enriched at centromeres. Mol. Biol. Evol. 35, 949–962. https://doi.org/10.1093/molbev/msy010
Kato, H., Jiang, J., Zhou, B.-R., Rozendaal, M., Feng, H., Ghirlando, R., Xiao, T.S., Straight, A.F., Bai, Y., 2013. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. Science 340, 1110–3. https://doi.org/10.1126/science.1235532
Kelly, A.E., Ghenoiu, C., Xue, J.Z., Zierhut, C., Kimura, H., Funabiki, H., 2010. Survivin reads phosphorylated histone H3 threonine 3 to activate the mitotic kinase Aurora B. Science 330, 235–239. https://doi.org/10.1126/science.1189505
Kim, I.S., Lee, M., Park, J.H., Jeon, R., Baek, S.H., Kim, K.I., 2014. βTrCP-mediated ubiquitylation regulates protein stability of Mis18β in a cell cycle-dependent manner. Biochem. Biophys. Res. Commun. 443, 62–67. https://doi.org/10.1016/j.bbrc.2013.11.058
Kim, J.-H., Lee, H.-S., Lee, N.C.O., Goncharov, N.V., Kumeiko, V., Masumoto, H., Earnshaw, W.C., Kouprina, N., Larionov, V., 2016. Development of a novel HAC-based "gain of signal" quantitative assay for measuring chromosome instability (CIN) in cancer cells. Oncotarget 7, 14841–14856. https://doi.org/10.18632/oncotarget.7854
Kim, S., Yu, H., 2015. Multiple assembly mechanisms anchor the KMN spindle checkpoint platform at human mitotic kinetochores. J. Cell Biol. 208, 181–96. https://doi.org/10.1083/jcb.201407074
Kipling, D., Warburton, P.E., 1997. Centromeres, CENP-B and Tigger too. Trends Genet. TIG 13, 141–145. https://doi.org/10.1016/s0168-9525(97)01098-6
Klare, K., Weir, J.R., Basilico, F., Zimniak, T., Massimiliano, L., Ludwigs, N., Herzog, F., Musacchio, A., 2015. CENP-C is a blueprint for constitutive centromere-associated network assembly within human kinetochores. J. Cell Biol. 210, 11–22. https://doi.org/10.1083/jcb.201412028
Kononenko, A.V., Lee, N.C.O., Liskovykh, M., Masumoto, H., Earnshaw, W.C., Larionov, V., Kouprina, N., 2015. Generation of a conditionally self-eliminating HAC gene delivery vector through incorporation of a tTAVP64 expression cassette. Nucleic Acids Res. 43, e57. https://doi.org/10.1093/nar/gkv124
Kouzine, F., Wojtowicz, D., Baranello, L., Yamane, A., Nelson, S., Resch, W., Kieffer-Kwon, K.-R., Benham, C.J., Casellas, R., Przytycka, T.M., Levens, D., 2017. Permanganate/S1 Nuclease Footprinting Reveals Non-B DNA Structures with Regulatory Potential across a Mammalian Genome. Cell Syst. 4, 344-356.e7. https://doi.org/10.1016/j.cels.2017.01.013

Kunitoku, N., Sasayama, T., Marumoto, T., Zhang, D., Honda, S., Kobayashi, O., Hatakeyama, K., Ushio, Y., Saya, H., Hirota, T., 2003. CENP-A Phosphorylation by Aurora-A in Prophase Is Required for Enrichment of Aurora-B at Inner Centromeres and for Kinetochore Function. Dev. Cell 5, 853–864. https://doi.org/10.1016/s1534-5807(03)00364-2

Kwon, M.-S., Hori, T., Okada, M., Fukagawa, T., 2007. CENP-C Is Involved in Chromosome Segregation, Mitotic Checkpoint Function, and Kinetochore Assembly. Mol. Biol. Cell 18, 2155–2168. https://doi.org/10.1091/mbc.e07-01-0045

Kyriacou, E., Heun, P., 2018. High-resolution mapping of centromeric protein association using APEX-chromatin fibers. Epigenetics Chromatin 11, 68. https://doi.org/10.1186/s13072-018-0237-6

Lacoste, N., Woolfe, A., Tachiwana, H., Garea, A.V., Barth, T., Cantaloube, S., Kurumizaka, H., Imhof, A., Almouzni, G., 2014. Mislocalization of the Centromeric Histone Variant CenH3/CENP-A in Human Cells Depends on the Chaperone DAXX. Mol. Cell 53, 631–644. https://doi.org/10.1016/j.molcel.2014.01.018

Lagana, A., Dorn, J.F., De Rop, V., Ladouceur, A.-M., Maddox, A.S., Maddox, P.S., 2010. A small GTPase molecular switch regulates epigenetic centromere maintenance by stabilizing newly incorporated CENP-A. Nat. Cell Biol. 12, 1186–1193. https://doi.org/10.1038/ncb2129

Lam, A.L., Boivin, C.D., Bonney, C.F., Rudd, M.K., Sullivan, B.A., 2006. Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. Proc. Natl. Acad. Sci. U. S. A. 103, 4186–4191. https://doi.org/10.1073/pnas.0507947103

Lando, D., Endesfelder, U., Berger, H., Subramanian, L., Dunne, P.D., McColl, J., Klenerman, D., Carr, A.M., Sauer, M., Allshire, R.C., Heilemann, M., Laue, E.D., 2012. Quantitative single-molecule microscopy reveals that CENP-A(Cnp1) deposition occurs during G2 in fission yeast. Open Biol. 2, 120078. https://doi.org/10.1098/rsob.120078

Lara-Gonzalez, P., Westhorpe, F.G., Taylor, S.S., 2012. The Spindle Assembly Checkpoint. Curr. Biol. 22, R966–R980. https://doi.org/10.1016/j.cub.2012.10.006

Lee, N.C.O., Kononenko, A.V., Lee, H.-S., Tolkunova, E.N., Liskovykh, M.A., Masumoto, H., Earnshaw, W.C., Tomilin, A.N., Larionov, V., Kouprina, N., 2013. Protecting a transgene expression from the HAC-based vector by different chromatin insulators. Cell. Mol. Life Sci. CMLS 70, 3723–3737. https://doi.org/10.1007/s00018-013-1362-9

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv13033997 Q-Bio.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liebelt, F., Jansen, N.S., Kumar, S., Gracheva, E., Claessens, L.A., Verlaan-de Vries, M., Willemstein, E., Vertegaal, A.C.O., 2019. The poly-SUMO2/3 protease SENP6 enables assembly of the constitutive centromere-associated network by group deSUMOylation. Nat. Commun. 10, 3987. https://doi.org/10.1038/s41467-019-11773-x

Liu, C., Mao, Y., 2016. Diaphanous formin mDia2 regulates CENP-A levels at centromeres. J. Cell Biol. 213, 415–24. https://doi.org/10.1083/jcb.201512034

Liu, C., Zhu, R., Mao, Y., 2018. Nuclear Actin Polymerized by mDia2 Confines Centromere Movement during CENP-A Loading. iScience 9, 314–327. https://doi.org/10.1016/j.isci.2018.10.031

Logsdon, G.A., Barrey, E.J., Bassett, E.A., DeNizio, J.E., Guo, L.Y., Panchenko, T., Dawicki-McKenna, J.M., Heun, P., Black, B.E., 2015. Both tails and the centromere targeting domain of CENP-A are required for centromere establishmentNew roles for the CENP-A N-terminal tail and CATD. J. Cell Biol. 208, 521–531. https://doi.org/10.1083/jcb.201412011

Logsdon, G.A., Gambogi, C.W., Liskovykh, M.A., Barrey, E.J., Larionov, V., Miga, K.H., Heun, P., Black, B.E., 2019. Human Artificial Chromosomes that Bypass Centromeric DNA. Cell 178, 624-639.e19. https://doi.org/10.1016/j.cell.2019.06.006

London, N., Biggins, S., 2014. Signalling dynamics in the spindle checkpoint response. Nat. Rev. Mol. Cell Biol. 15, 736–47. https://doi.org/10.1038/nrm3888
Loyola, A., Almouzni, G., 2004. Histone chaperones, a supporting role in the limelight. Biochim. Biophys. Acta BBA - Gene Struct. Expr. 1677, 3–11. https://doi.org/10.1016/j.bbaexp.2003.09.012
Ly, P., Teitz, L.S., Kim, D.H., Shoshani, O., Skaletsky, H., Fachinetti, D., Page, D.C., Cleveland, D.W., 2017. Selective Y centromere inactivation triggers chromosome shattering in micronuclei and repair by non-homologous end joining. Nat. Cell Biol. 19, 68–75. https://doi.org/10.1038/ncb3450
Lyubitelev, A.V., Nikitin, D.V., Shaytan, A.K., Studitsky, V.M., Kirpichnikov, M.P., 2016. Structure and functions of linker histones. Biochem. Mosc. 81, 213–223. https://doi.org/10.1134/s0006297916030032
Maalouf, J.E., Texier, P., Erliandri, I., Cohen, C., Corpet, A., Catez, F., Boutell, C., Lomonte, P., 2018. CENP-B dynamics at centromeres is regulated by a SUMOylation/ubiquitination and proteasomal-dependent degradation mechanism involving the SUMO-targeted ubiquitin E3 ligase RNF4. bioRxiv 245597. https://doi.org/10.1101/245597
Maddox, P., Desai, A., Oegema, K., Mitchison, T.J., Salmon, E.D., 2002. Poleward Microtubule Flux Is a Major Component of Spindle Dynamics and Anaphase A in Mitotic Drosophila Embryos. Curr. Biol. 12, 1670–1674. https://doi.org/10.1016/s0960-9822(02)01183-1
Maddox, P.S., Bloom, K.S., Salmon, E.D., 1999. The polarity and dynamics of microtubule assembly in the budding yeast Saccharomyces cerevisiae. Nat. Cell Biol. 2, 36–41. https://doi.org/10.1038/71357
Maddox, P.S., Hyndman, F., Monen, J., Oegema, K., Desai, A., 2007. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. J. Cell Biol. 176, 757–763. https://doi.org/10.1083/jcb.200701065
Maiato, H., Gomes, A., Sousa, F., Barisic, M., 2017. Mechanisms of Chromosome Congression during Mitosis. Biology 6, 13. https://doi.org/10.3390/biology6010013
Maloney, K.A., Sullivan, L.L., Matheny, J.E., Strome, E.D., Merrett, S.L., Ferris, A., Sullivan, B.A., 2012. Functional epialleles at an endogenous human centromere. Proc. Natl. Acad. Sci. 109, 13704–13709. https://doi.org/10.1073/pnas.1203126109
Marshall, Owen J., Chueh, A.C., Wong, L.H., Choo, K.H.A., 2008. Neocentromeres: New Insights into Centromere Structure, Disease Development, and Karyotype Evolution. Am. J. Hum. Genet. 82, 261–282. https://doi.org/10.1016/j.ajhg.2007.11.009
Marshall, Owen J, Marshall, A.T., Choo, K.H.A., 2008. Three-dimensional localization of CENP-A suggests a complex higher order structure of centromeric chromatin. J. Cell Biol. 183, 1193–202. https://doi.org/10.1083/jcb.200804078
Martins, N.M.C., Bergmann, J.H., Shono, N., Kimura, H., Larionov, V., Masumoto, H., Earnshaw, W.C., 2016. Epigenetic engineering shows that a human centromere resists silencing mediated by H3K27me3/K9me3. Mol. Biol. Cell 27, 177–196. https://doi.org/10.1091/mbc.E15-08-0605
Masumoto, H., Masukata, H., Muro, Y., Nozaki, N., Okazaki, T., 1989. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. J. Cell Biol. 109, 1963–1973. https://doi.org/10.1083/jcb.109.5.1963
McClintock, B., 1941. The Stability of Broken Ends of Chromosomes in Zea Mays. Genetics 26, 234–82.
McClintock, B., 1939. The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis. Proc. Natl. Acad. Sci. 25, 405–416. https://doi.org/10.1073/pnas.25.8.405
McGrew, J., Diehl, B., Fitzgerald-Hayes, M., 1986. Single base-pair mutations in centromere element III cause aberrant chromosome segregation in Saccharomyces cerevisiae. Mol. Cell. Biol. 6, 530–538. https://doi.org/10.1128/mcb.6.2.530

McKinley, K.L., Cheeseman, I.M., 2016. The molecular basis for centromere identity and function. Nat. Rev. Mol. Cell Biol. 17, 16–29. https://doi.org/10.1038/nrm.2015.5
McKinley, K.L., Cheeseman, I.M., 2014. Polo-like Kinase 1 Licenses CENP-A Deposition at Centromeres. Cell 158, 397–411. https://doi.org/10.1016/j.cell.2014.06.016
McKinley, K.L., Sekulic, N., Guo, L.Y., Tsinman, T., Black, B.E., Cheeseman, I.M., 2015. The CENP-L-N Complex Forms a Critical Node in an Integrated Meshwork of Interactions at the Centromere-Kinetochore Interface. Mol. Cell 60, 886–898. https://doi.org/10.1016/j.molcel.2015.10.027
McKnight, S.L., Miller, O.L., 1977. Electron microscopic analysis of chromatin replication in the cellular blastoderm drosophila melanogaster embryo. Cell 12, 795–804. https://doi.org/10.1016/0092-8674(77)90278-1
McNulty, S.M., Sullivan, L.L., Sullivan, B.A., 2017. Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are Complexed with CENP-A and CENP-C. Dev. Cell 42, 226-240.e6. https://doi.org/10.1016/j.devcel.2017.07.001
Melters, D.P., Paliulis, L.V., Korf, I.F., Chan, S.W.L., 2012. Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. 20, 579–93. https://doi.org/10.1007/s10577-012-9292-1
Melters, D.P., Rakshit, T., Bui, M., Grigoryev, S.A., Sturgill, D., Dalal, Y., 2019. The ratio between centromeric proteins CENP-A and CENP-C maintains homeostasis of human centromeres. bioRxiv 604223. https://doi.org/10.1101/604223
Mendiburo, M., 2011. Drosophila CENH3 Is Sufficient for Centromere Formation. Science 334, 686–690. https://doi.org/10.1126/science.1206880
Miga, K.H., 2017. The Promises and Challenges of Genomic Studies of Human Centromeres. Prog. Mol. Subcell. Biol. 56, 285–304. https://doi.org/10.1007/978-3-319-58592-5_12
Miga, K.H., Newton, Y., Jain, M., Altemose, N., Willard, H.F., Kent, W.J., 2014. Centromere reference models for human chromosomes X and Y satellite arrays. Genome Res. 24, 697–707. https://doi.org/10.1101/gr.159624.113
Milks, K.J., Moree, B., Straight, A.F., 2009. Dissection of CENP-C-directed centromere and kinetochore assembly. Mol. Biol. Cell 20, 4246–55. https://doi.org/10.1091/mbc.e09-05-0378
Mitchell, A.R., Jeppesen, P., Nicol, L., Morrison, H., Kipling, D., 1996. Epigenetic control of mammalian centromere protein binding: does DNA methylation have a role? J. Cell Sci. 109 (Pt 9), 2199–2206.
Mitchison, T., Kirschner, M., 1984. Microtubule assembly nucleated by isolated centrosomes. Nature 312, 232–237. https://doi.org/10.1038/312232a0
Mitra, S., Bodor, D.L., David, A.F., Abdul-Zani, I., Mata, J.F., Neumann, B., Reither, S., Tischer, C., Jansen, L.E.T., 2020. Genetic screening identifies a SUMO protease dynamically maintaining centromeric chromatin. Nat. Commun. 11, 501. https://doi.org/10.1038/s41467-019-14276-x
Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M.M., Wu, C., 2007. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. Cell 129, 1153–1164. https://doi.org/10.1016/j.cell.2007.04.026
Molina, O., Vargiu, G., Abad, M.A., Zhiteneva, A., Jeyaprakash, A.A., Masumoto, H., Kouprina, N., Larionov, V., Earnshaw, W.C., 2016. Epigenetic engineering reveals a balance between histone modifications and transcription in kinetochore maintenance. Nat. Commun. 7, 13334. https://doi.org/10.1038/ncomms13334
Moree, B., Meyer, C.B., Fuller, C.J., Straight, A.F., 2011. CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. J. Cell Biol. 194, 855–871. https://doi.org/10.1083/jcb.201106079

Moroi, Y., Peebles, C., Fritzler, M.J., Steigerwald, J., Tan, E.M., 1980. Autoantibody to centromere (kinetochore) in scleroderma sera. Proc. Natl. Acad. Sci. 77, 1627–1631. https://doi.org/10.1073/pnas.77.3.1627
Morozov, V.M., Giovinazzi, S., Ishov, A.M., 2017. CENP-B protects centromere chromatin integrity by facilitating histone deposition via the H3.3-specific chaperone Daxx. Epigenetics Chromatin 10, 63. https://doi.org/10.1186/s13072-017-0164-y
Mravinac, B., Sullivan, L.L., Reeves, J.W., Yan, C.M., Kopf, K.S., Farr, C.J., Schueler, M.G., Sullivan, B.A., 2009. Histone modifications within the human X centromere region. PloS One 4, e6602. https://doi.org/10.1371/journal.pone.0006602
Müller, S., Montes de Oca, R., Lacoste, N., Dingli, F., Loew, D., Almouzni, G., 2014. Phosphorylation and DNA Binding of HJURP Determine Its Centromeric Recruitment and Function in CenH3CENP-A Loading. Cell Rep. 8, 190–203. https://doi.org/10.1016/j.celrep.2014.06.002
Murillo-Pineda, M., Jansen, L.E.T., 2020. Genetics, epigenetics and back again: Lessons learned from neocentromeres. Exp. Cell Res. 389, 111909. https://doi.org/10.1016/j.yexcr.2020.111909
Muro, Y., Masumoto, H., Yoda, K., Nozaki, N., Ohashi, M., Okazaki, T., 1992. Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. J. Cell Biol. 116, 585–596. https://doi.org/10.1083/jcb.116.3.585
Murray, A.W., 2011. A brief history of error. Nat. Cell Biol. 13, 1178–1182. https://doi.org/10.1038/ncb2348
Musacchio, A., Desai, A., 2017. A Molecular View of Kinetochore Assembly and Function. Biology 6, 5. https://doi.org/10.3390/biology6010005
Nagpal, H., Hori, T., Furukawa, A., Sugase, K., Osakabe, A., Kurumizaka, H., Fukagawa, T., 2015. Dynamic changes in CCAN organization through CENP-C during cell-cycle progression. Mol. Biol. Cell 26, 3768–76. https://doi.org/10.1091/mbc.e15-07-0531
Nakano, M., Cardinale, S., Noskov, V.N., Gassmann, R., Vagnarelli, P., Kandels-Lewis, S., Larionov, V., Earnshaw, W.C., Masumoto, H., 2008. Inactivation of a human kinetochore by specific targeting of chromatin modifiers. Dev. Cell 14, 507–522. https://doi.org/10.1016/j.devcel.2008.02.001
Nakano, M., Okamoto, Y., Ohzeki, J., Masumoto, H., 2003. Epigenetic assembly of centromeric chromatin at ectopic -satellite sites on human chromosomes. J. Cell Sci. 116, 4021–4034. https://doi.org/10.1242/jcs.00697
Nakashima, H., Nakano, M., Ohnishi, R., Hiraoka, Y., Kaneda, Y., Sugino, A., Masumoto, H., 2005. Assembly of additional heterochromatin distinct from centromere-kinetochore chromatin is required for de novo formation of human artificial chromosome. J. Cell Sci. 118, 5885–5898. https://doi.org/10.1242/jcs.02702
Nardi, I.K., Zasadzińska, E., Stellfox, M.E., Knippler, C.M., Foltz, D.R., 2016. Licensing of Centromeric Chromatin Assembly through the Mis18α-Mis18β Heterotetramer. Mol. Cell 61, 774–787. https://doi.org/10.1016/j.molcel.2016.02.014
Nasuda, S., Hudakova, S., Schubert, I., Houben, A., Endo, T.R., 2005. Stable barley chromosomes without centromeric repeats. Proc. Natl. Acad. Sci. 102, 9842–9847. https://doi.org/10.1073/pnas.0504235102
Nava, M.M., Miroshnikova, Y.A., Biggs, L.C., Whitefield, D.B., Metge, F., Boucas, J., Vihinen, H., Jokitalo, E., Li, X., Arcos, J.M.G., Hoffmann, B., Merkel, R., Niessen, C.M., Dahl, K.N., Wickström, S.A., 2020. Heterochromatin-Driven Nuclear Softening Protects the Genome against Mechanical Stress-Induced Damage. Cell 181, 800-817.e22. https://doi.org/10.1016/j.cell.2020.03.052
Nechemia-Arbely, Y., Fachinetti, D., Miga, K.H., Sekulic, N., Soni, G.V., Kim, D.H., Wong, A.K., Lee, A.Y., Nguyen, K., Dekker, C., Ren, B., Black, B.E., Cleveland, D.W., 2017. Human

centromeric CENP-A chromatin is a homotypic, octameric nucleosome at all cell cycle points. J. Cell Biol. 216, 607–621. https://doi.org/10.1083/jcb.201608083
Nechemia-Arbely, Y., Miga, K.H., Shoshani, O., Aslanian, A., McMahon, M.A., Lee, A.Y., Fachinetti, D., Yates, J.R., Ren, B., Cleveland, D.W., 2019. DNA replication acts as an error correction mechanism to maintain centromere identity by restricting CENP-A to centromeres. Nat. Cell Biol. 21, 743–754. https://doi.org/10.1038/s41556-019-0331-4
Ng, R., Carbon, J., 1987. Mutational and in vitro protein-binding studies on centromere DNA from Saccharomyces cerevisiae. Mol. Cell. Biol. 7, 4522–4534. https://doi.org/10.1128/mcb.7.12.4522
Niikura, Y., Kitagawa, R., Fang, L., Kitagawa, K., 2019. CENP-A Ubiquitylation Is Indispensable to Cell Viability. Dev. Cell 50, 683-689.e6. https://doi.org/10.1016/j.devcel.2019.07.015
Niikura, Y., Kitagawa, R., Kitagawa, K., 2016. CENP-A Ubiquitylation Is Inherited through Dimerization between Cell Divisions. Cell Rep. 15, 61–76. https://doi.org/10.1016/j.celrep.2016.03.010
Niikura, Y., Kitagawa, R., Ogi, H., Abdulle, R., Pagala, V., Kitagawa, K., 2015. CENP-A K124 Ubiquitylation Is Required for CENP-A Deposition at the Centromere. Dev. Cell 32, 589–603. https://doi.org/10.1016/j.devcel.2015.01.024
Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., Kanemaki, M., 2009. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat. Methods 6, 917–922. https://doi.org/10.1038/nmeth.1401
Nishino, T., Rago, F., Hori, T., Tomii, K., Cheeseman, I.M., Fukagawa, T., 2013. CENP-T provides a structural platform for outer kinetochore assembly. EMBO J. 32, 424–36. https://doi.org/10.1038/emboj.2012.348
Ohzeki, J., Bergmann, J.H., Kouprina, N., Noskov, V.N., Nakano, M., Kimura, H., Earnshaw, W.C., Larionov, V., Masumoto, H., 2012. Breaking the HAC Barrier: histone H3K9 acetyl/methyl balance regulates CENP-A assembly. EMBO J. 31, 2391–2402. https://doi.org/10.1038/emboj.2012.82
Ohzeki, J., Larionov, V., Earnshaw, W.C., Masumoto, H., 2019. De novo formation and epigenetic maintenance of centromere chromatin. Curr. Opin. Cell Biol. 58, 15–25. https://doi.org/10.1016/j.ceb.2018.12.004
Ohzeki, J., Nakano, M., Okada, T., Masumoto, H., 2002. CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. J. Cell Biol. 159, 765–775. https://doi.org/10.1083/jcb.200207112
Ohzeki, J., Shono, N., Otake, K., Martins, N.M.C., Kugou, K., Kimura, H., Nagase, T., Larionov, V., Earnshaw, W.C., Masumoto, H., 2016. KAT7/HBO1/MYST2 Regulates CENP-A Chromatin Assembly by Antagonizing Suv39h1-Mediated Centromere Inactivation. Dev. Cell 37, 413–427. https://doi.org/10.1016/j.devcel.2016.05.006
Ohzeki, J.-I., Otake, K., Masumoto, H., 2020. Human artificial chromosome: Chromatin assembly mechanisms and CENP-B. Exp. Cell Res. 389, 111900. https://doi.org/10.1016/j.yexcr.2020.111900
Okada, M., Cheeseman, I.M., Hori, T., Okawa, K., McLeod, I.X., Yates, J.R., Desai, A., Fukagawa, T., 2006. The CENP-H–I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. Nat. Cell Biol. 8, 446–457. https://doi.org/10.1038/ncb1396
Okada, T., Ohzeki, J., Nakano, M., Yoda, K., Brinkley, W.R., Larionov, V., Masumoto, H., 2007. CENP-B Controls Centromere Formation Depending on the Chromatin Context. Cell 131, 1287–1300. https://doi.org/10.1016/j.cell.2007.10.045
Okamoto, Y., Nakano, M., Ohzeki, J., Larionov, V., Masumoto, H., 2007. A minimal CENP-A core is required for nucleation and maintenance of a functional human centromere. EMBO J. 26, 1279–1291. https://doi.org/10.1038/sj.emboj.7601584

Olszak, A., Essen, D., Pereira, A., Diehl, S., Manke, T., Maiato, H., Saccani, S., Heun, P., 2011. Heterochromatin boundaries are hotspots for de novo kinetochore formation. Nat. Cell Biol. 13, 799–808. https://doi.org/10.1038/ncb2272

Otake, K., Ohzeki, J., Shono, N., Kugou, K., Okazaki, K., Nagase, T., Yamakawa, H., Kouprina, N., Larionov, V., Kimura, H., Earnshaw, W.C., Masumoto, H., 2020. CENP-B creates alternative epigenetic chromatin states permissive for CENP-A or heterochromatin assembly. J. Cell Sci. https://doi.org/10.1242/jcs.243303

Ou, H.D., Phan, S., Deerinck, T.J., Thor, A., Ellisman, M.H., O'Shea, C.C., 2017. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. Science 357. https://doi.org/10.1126/science.aag0025

Palladino, J., Chavan, A., Sposato, A., Mason, T.D., Mellone, B.G., 2020. Targeted De Novo Centromere Formation in Drosophila Reveals Plasticity and Maintenance Potential of CENP-A Chromatin. Dev. Cell 52, 379-394.e7. https://doi.org/10.1016/j.devcel.2020.01.005

Palmer, D., O'Day, K., Wener, M., Andrews, B., Margolis, R., 1987. A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J. Cell Biol. 104, 805–815. https://doi.org/10.1083/jcb.104.4.805

Palmer, D.K., O'Day, K., Margolis, R.L., 1990. The centromere specific histone CENP-A is selectively retained in discrete foci in mammalian sperm nuclei. Chromosoma 100, 32–36. https://doi.org/10.1007/BF00337600

Pan, D., Klare, K., Petrovic, A., Take, A., Walstein, K., Singh, P., Rondelet, A., Bird, A.W., Musacchio, A., 2017. CDK-regulated dimerization of M18BP1 on a Mis18 hexamer is necessary for CENP-A loading. eLife 6, e23352. https://doi.org/10.7554/elife.23352

Pan, D., Walstein, K., Take, A., Bier, D., Kaiser, N., Musacchio, A., 2019. Mechanism of centromere recruitment of the CENP-A chaperone HJURP and its implications for centromere licensing. Nat. Commun. 10, 4046. https://doi.org/10.1038/s41467-019-12019-6

Pearson, C.G., Yeh, E., Gardner, M., Odde, D., Salmon, E.D., Bloom, K., 2004. Stable kinetochore-microtubule attachment constrains centromere positioning in metaphase. Curr. Biol. CB 14, 1962–1967. https://doi.org/10.1016/j.cub.2004.09.086

Pentakota, S., Zhou, K., Smith, C., Maffini, S., Petrovic, A., Morgan, G.P., Weir, J.R., Vetter, I.R., Musacchio, A., Luger, K., 2017. Decoding the centromeric nucleosome through CENP-N. eLife 6, e33442. https://doi.org/10.7554/elife.33442

Perea-Resa, C., Bury, L., Cheeseman, I.M., Blower, M.D., 2020. Cohesin Removal Reprograms Gene Expression upon Mitotic Entry. Mol. Cell 78, 127-140.e7. https://doi.org/10.1016/j.molcel.2020.01.023

Perez-Castro, A.V., Shamanski, F.L., Meneses, J.J., Lovato, T.L., Vogel, K.G., Moyzis, R.K., Pedersen, R., 1998. Centromeric protein B null mice are viable with no apparent abnormalities. Dev. Biol. 201, 135–143. https://doi.org/10.1006/dbio.1998.9005

Pesenti, M.E., Prumbaum, D., Auckland, P., Smith, C.M., Faesen, A.C., Petrovic, A., Erent, M., Maffini, S., Pentakota, S., Weir, J.R., Lin, Y.-C., Raunser, S., McAinsh, A.D., Musacchio, A., 2018. Reconstitution of a 26-Subunit Human Kinetochore Reveals Cooperative Microtubule Binding by CENP-OPQUR and NDC80. Mol. Cell 71, 923-939.e10. https://doi.org/10.1016/j.molcel.2018.07.038

Peters, J.-M., Tedeschi, A., Schmitz, J., 2008. The cohesin complex and its roles in chromosome biology. Genes Dev. 22, 3089–114. https://doi.org/10.1101/gad.1724308

Petrovic, A., Keller, J., Liu, Y., Overlack, K., John, J., Dimitrova, Y.N., Jenni, S., Gerwen, S. van, Stege, P., Wohlgemuth, S., Rombaut, P., Herzog, F., Harrison, S.C., Vetter, I.R., Musacchio, A., 2016. Structure of the MIS12 Complex and Molecular Basis of Its Interaction with CENP-C at Human Kinetochores. Cell 167, 1028-1040.e15. https://doi.org/10.1016/j.cell.2016.10.005

Piacentini, L., Marchetti, M., Bucciarelli, E., Casale, A.M., Cappucci, U., Bonifazi, P., Renda, F., Fanti, L., 2019. A role of the Trx-G complex in Cid/CENP-A deposition at Drosophila melanogaster centromeres. Chromosoma 128, 503–520. https://doi.org/10.1007/s00412-019-00711-x

Pidoux, A.L., Choi, E.S., Abbott, J.K.R., Liu, X., Kagansky, A., Castillo, A.G., Hamilton, G.L., Richardson, W., Rappsilber, J., He, X., Allshire, R.C., 2009. Fission Yeast Scm3: A CENP-A Receptor Required for Integrity of Subkinetochore Chromatin. Mol. Cell 33, 299–311. https://doi.org/10.1016/j.molcel.2009.01.019

Pierre, R.V., Hoagland, H.C., 1972. Age-associated aneuploidy: loss of Y chromosome from human bone marrow cells with aging. Cancer 30, 889–894. https://doi.org/10.1002/1097-0142(197210)30:4<889::aid-cncr2820300405>3.0.co;2-1

Pluta, A.F., Mackay, A.M., Ainsztein, A.M., Goldberg, I.G., Earnshaw, W.C., 1995. The centromere: hub of chromosomal activities. Science 270, 1591–1594. https://doi.org/10.1126/science.270.5242.1591

Politi, V., Perini, G., Trazzi, S., Pliss, A., Raska, I., Earnshaw, W.C., Della Valle, G., 2002. CENP-C binds the alpha-satellite DNA in vivo at specific centromere domains. J. Cell Sci. 115, 2317–2327.

Prendergast, L., Müller, S., Liu, Y., Huang, H., Dingli, F., Loew, D., Vassias, I., Patel, D.J., Sullivan, K.F., Almouzni, G., 2016. The CENP-T/-W complex is a binding partner of the histone chaperone FACT. Genes Dev. 30, 1313–1326. https://doi.org/10.1101/gad.275073.115

Przewloka, M.R., Venkei, Z., Bolanos-Garcia, V.M., Debski, J., Dadlez, M., Glover, D.M., 2011. CENP-C Is a Structural Platform for Kinetochore Assembly. Curr. Biol. 21, 399–405. https://doi.org/10.1016/j.cub.2011.02.005

Quénet, D., Dalal, Y., 2014. A long non-coding RNA is required for targeting centromeric protein A to the human centromere. eLife 3, e03254. https://doi.org/10.7554/eLife.03254

Rago, F., Gascoigne, K.E., Cheeseman, I.M., 2015. Distinct Organization and Regulation of the Outer Kinetochore KMN Network Downstream of CENP-C and CENP-T. Curr. Biol. 25, 671–677. https://doi.org/10.1016/j.cub.2015.01.059

Ransom, M., Dennehey, B.K., Tyler, J.K., 2010. Chaperoning histones during DNA replication and repair. Cell 140, 183–95. https://doi.org/10.1016/j.cell.2010.01.004

Raychaudhuri, N., Dubruille, R., Orsi, G.A., Bagheri, H.C., Loppin, B., Lehner, C.F., 2012. Transgenerational propagation and quantitative maintenance of paternal centromeres depends on Cid/Cenp-A presence in Drosophila sperm. PLoS Biol. 10, e1001434. https://doi.org/10.1371/journal.pbio.1001434

Régnier, V., Vagnarelli, P., Fukagawa, T., Zerjal, T., Burns, E., Trouche, D., Earnshaw, W., Brown, W., 2005. CENP-A is required for accurate chromosome segregation and sustained kinetochore association of BubR1. Mol. Cell. Biol. 25, 3967–3981. https://doi.org/10.1128/MCB.25.10.3967-3981.2005

Ribeiro, S.A., Vagnarelli, P., Dong, Y., Hori, T., McEwen, B.F., Fukagawa, T., Flors, C., Earnshaw, W.C., 2010. A super-resolution map of the vertebrate kinetochore. Proc. Natl. Acad. Sci. U. S. A. 107, 10484–9. https://doi.org/10.1073/pnas.1002325107

Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y., Allis, C.D., 2003. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol. Cell 12, 1591–1598. https://doi.org/10.1016/s1097-2765(03)00479-9

Richter, M.M., Poznanski, J., Zdziarska, A., Czarnocki-Cieciura, M., Lipinszki, Z., Dadlez, M., Glover, D.M., Przewloka, M.R., 2016. Network of protein interactions within the Drosophila inner kinetochore. Open Biol. 6, 150238. https://doi.org/10.1098/rsob.150238

Rocchi, M., Archidiacono, N., Schempp, W., Capozzi, O., Stanyon, R., 2012. Centromere repositioning in mammals. Heredity 108, 59–67. https://doi.org/10.1038/hdy.2011.101

Rosandić, M., Gluncić, M., Paar, V., Basar, I., 2008. The role of alphoid higher order repeats (HORs) in the centromere folding. J. Theor. Biol. 254, 555–560. https://doi.org/10.1016/j.jtbi.2008.06.012
Rošić, S., Köhler, F., Erhardt, S., 2014. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. J. Cell Biol. 207, 335–349. https://doi.org/10.1083/jcb.201404097
Ross, J.E., Woodlief, K.S., Sullivan, B.A., 2016. Inheritance of the CENP-A chromatin domain is spatially and temporally constrained at human centromeres. Epigenetics Chromatin 9, 20. https://doi.org/10.1186/s13072-016-0071-7
Roulland, Y., Ouararhni, K., Naidenov, M., Ramos, L., Shuaib, M., Syed, S.H., Lone, I.N., Boopathi, R., Fontaine, E., Papai, G., Tachiwana, H., Gautier, T., Skoufias, D., Padmanabhan, K., Bednar, J., Kurumizaka, H., Schultz, P., Angelov, D., Hamiche, A., Dimitrov, S., 2016. The Flexible Ends of CENP-A Nucleosome Are Required for Mitotic Fidelity. Mol. Cell 63, 674–685. https://doi.org/10.1016/j.molcel.2016.06.023
Roure, V., Medina-Pritchard, B., Lazou, V., Rago, L., Anselm, E., Venegas, D., Jeyaprakash, A.A., Heun, P., 2019. Reconstituting Drosophila Centromere Identity in Human Cells. Cell Rep. 29, 464-479.e5. https://doi.org/10.1016/j.celrep.2019.08.067
Saffery, R., Irvine, D.V., Griffiths, B., Kalitsis, P., Wordeman, L., Choo, K.H., 2000. Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins. Hum. Mol. Genet. 9, 175–185. https://doi.org/10.1093/hmg/9.2.175
Saffery, R., Sumer, H., Hassan, S., Wong, L.H., Craig, J.M., Todokoro, K., Anderson, M., Stafford, A., Choo, K.H.A., 2003. Transcription within a functional human centromere. Mol. Cell 12, 509–516. https://doi.org/10.1016/s1097-2765(03)00279-x
Saitoh, H., Tomkiel, J., Cooke, C.A., Ratrie, H., Maurer, M., Rothfield, N.F., Earnshaw, W.C., 1992. CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. Cell 70, 115–125. https://doi.org/10.1016/0092-8674(92)90538-n
Sakuno, T., Watanabe, Y., 2009. Studies of meiosis disclose distinct roles of cohesion in the core centromere and pericentromeric regions. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. 17, 239–249. https://doi.org/10.1007/s10577-008-9013-y
Saldivar, J.C., Hamperl, S., Bocek, M.J., Chung, M., Bass, T.E., Cisneros-Soberanis, F., Samejima, K., Xie, L., Paulson, J.R., Earnshaw, W.C., Cortez, D., Meyer, T., Cimprich, K.A., 2018. An intrinsic S/G2 checkpoint enforced by ATR. Science 361, 806–810. https://doi.org/10.1126/science.aap9346
Samel, A., Cuomo, A., Bonaldi, T., Ehrenhofer-Murray, A.E., 2012. Methylation of CenH3 arginine 37 regulates kinetochore integrity and chromosome segregation. Proc. Natl. Acad. Sci. U. S. A. 109, 9029–9034. https://doi.org/10.1073/pnas.1120968109
Santaguida, S., Amon, A., 2015. Short- and long-term effects of chromosome mis-segregation and aneuploidy. Nat. Rev. Mol. Cell Biol. 16, 473–485. https://doi.org/10.1038/nrm4025
Santaguida, S., Musacchio, A., 2009. The life and miracles of kinetochores. EMBO J. 28, 2511–2531. https://doi.org/10.1038/emboj.2009.173
Sathyan, K.M., Fachinetti, D., Foltz, D.R., 2017. α-amino trimethylation of CENP-A by NRMT is required for full recruitment of the centromere. Nat. Commun. 8, 14678. https://doi.org/10.1038/ncomms14678
Sato, H., Masuda, F., Takayama, Y., Takahashi, K., Saitoh, S., 2012. Epigenetic Inactivation and Subsequent Heterochromatinization of a Centromere Stabilize Dicentric Chromosomes. Curr. Biol. 22, 658–667. https://doi.org/10.1016/j.cub.2012.02.062
Scelfo, A., Fachinetti, D., 2019. Keeping the Centromere under Control: A Promising Role for DNA Methylation. Cells 8. https://doi.org/10.3390/cells8080912

Schittenhelm, R.B., Althoff, F., Heidmann, S., Lehner, C.F., 2010. Detrimental incorporation of excess Cenp-A/Cid and Cenp-C into Drosophila centromeres is prevented by limiting amounts of the bridging factor Cal1. J. Cell Sci. 123, 3768–3779. https://doi.org/10.1242/jcs.067934

Schneider, V.A., Graves-Lindsay, T., Howe, K., Bouk, N., Chen, H.-C., Kitts, P.A., Murphy, T.D., Pruitt, K.D., Thibaud-Nissen, F., Albracht, D., Fulton, R.S., Kremitzki, M., Magrini, V., Markovic, C., McGrath, S., Steinberg, K.M., Auger, K., Chow, W., Collins, J., Harden, G., Hubbard, T., Pelan, S., Simpson, J.T., Threadgold, G., Torrance, J., Wood, J.M., Clarke, L., Koren, S., Boitano, M., Peluso, P., Li, H., Chin, C.-S., Phillippy, A.M., Durbin, R., Wilson, R.K., Flicek, P., Eichler, E.E., Church, D.M., 2017. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. Genome Res. 27, 849–864. https://doi.org/10.1101/gr.213611.116

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., Jenuwein, T., 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev. 18, 1251–1262. https://doi.org/10.1101/gad.300704

Screpanti, E., Antoni, A.D., Alushin, G.M., Petrovic, A., Melis, T., Nogales, E., Musacchio, A., 2011. Direct binding of Cenp-C to the Mis12 complex joins the inner and outer kinetochore. Curr. Biol. CB 21, 391–8. https://doi.org/10.1016/j.cub.2010.12.039

Sekulic, N., Bassett, E.A., Rogers, D.J., Black, B.E., 2010. The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. Nature 467, 347–351. https://doi.org/10.1038/nature09323

Shang, W.-H., Hori, T., Martins, N.M.C., Toyoda, A., Misu, S., Monma, N., Hiratani, I., Maeshima, K., Ikeo, K., Fujiyama, A., Kimura, H., Earnshaw, W.C., Fukagawa, T., 2013. Chromosome engineering allows the efficient isolation of vertebrate neocentromeres. Dev. Cell 24, 635–48. https://doi.org/10.1016/j.devcel.2013.02.009

Shang, W.-H., Hori, T., Toyoda, A., Kato, J., Popendorf, K., Sakakibara, Y., Fujiyama, A., Fukagawa, T., 2010. Chickens possess centromeres with both extended tandem repeats and short non-tandem-repetitive sequences. Genome Res. 20, 1219–28. https://doi.org/10.1101/gr.106245.110

Shang, W.-H., Hori, T., Westhorpe, F.G., Godek, K.M., Toyoda, A., Misu, S., Monma, N., Ikeo, K., Carroll, C.W., Takami, Y., Fujiyama, A., Kimura, H., Straight, A.F., Fukagawa, T., 2016. Acetylation of histone H4 lysine 5 and 12 is required for CENP-A deposition into centromeres. Nat. Commun. 7, 13465. https://doi.org/10.1038/ncomms13465

Shelby, R.D., Monier, K., Sullivan, K.F., 2000. Chromatin Assembly at Kinetochores Is Uncoupled from DNA Replication. J. Cell Biol. 151, 1113–1118. https://doi.org/10.1083/jcb.151.5.1113

Shelby, R.D., Vafa, O., Sullivan, K.F., 1997. Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. J. Cell Biol. 136, 501–513. https://doi.org/10.1083/jcb.136.3.501

Shibahara, K., Stillman, B., 1999. Replication-Dependent Marking of DNA by PCNA Facilitates CAF-1-Coupled Inheritance of Chromatin. Cell 96, 575–585. https://doi.org/10.1016/s0092-8674(00)80661-3

Shivaraju, M., Unruh, J.R., Slaughter, B.D., Mattingly, M., Berman, J., Gerton, J.L., 2012. Cell-cycle-coupled structural oscillation of centromeric nucleosomes in yeast. Cell 150, 304–16. https://doi.org/10.1016/j.cell.2012.05.034

Shono, N., Ohzeki, J., Otake, K., Martins, N.M.C., Nagase, T., Kimura, H., Larionov, V., Earnshaw, W.C., Masumoto, H., 2015. CENP-C and CENP-I are key connecting factors for kinetochore and CENP-A assembly. J Cell Sci 128, 4572–4587. https://doi.org/10.1242/jcs.180786

Shrestha, R.L., Ahn, G.S., Staples, M.I., Sathyan, K.M., Karpova, T.S., Foltz, D.R., Basrai, M.A., 2017. Mislocalization of centromeric histone H3 variant CENP-A contributes to

chromosomal instability (CIN) in human cells. Oncotarget 8, 46781–46800. https://doi.org/10.18632/oncotarget.18108
Shu, X., Liu, M., Lu, Z., Zhu, C., Meng, H., Huang, S., Zhang, X., Yi, C., 2018. Genome-wide mapping reveals that deoxyuridine is enriched in the human centromeric DNA. Nat. Chem. Biol. 14, 680–687. https://doi.org/10.1038/s41589-018-0065-9
Shuaib, M., Ouararhni, K., Dimitrov, S., Hamiche, A., 2010. HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. Proc. Natl. Acad. Sci. U. S. A. 107, 1349–1354. https://doi.org/10.1073/pnas.0913709107
Shukla, M., Tong, P., White, S.A., Singh, P.P., Reid, A.M., Catania, S., Pidoux, A.L., Allshire, R.C., 2018a. Centromere DNA Destabilizes H3 Nucleosomes to Promote CENP-A Deposition during the Cell Cycle. Curr. Biol. 28, 3924-3936.e4. https://doi.org/10.1016/j.cub.2018.10.049
Shukla, M., Tong, P., White, S.A., Singh, P.P., Reid, A.M., Catania, S., Pidoux, A.L., Allshire, R.C., 2018b. Centromere DNA Destabilizes H3 Nucleosomes to Promote CENP-A Deposition during the Cell Cycle. Curr. Biol. 28, 3924-3936.e4. https://doi.org/10.1016/j.cub.2018.10.049
Silkworth, W.T., Nardi, I.K., Paul, R., Mogilner, A., Cimini, D., 2011. Timing of centrosome separation is important for accurate chromosome segregation. Mol. Biol. Cell 23, 401–11. https://doi.org/10.1091/mbc.e11-02-0095
Silva, M.C.C., Bodor, D.L., Stellfox, M.E., Martins, N.M.C., Hochegger, H., Foltz, D.R., Jansen, L.E.T., 2011. Cdk activity couples epigenetic centromere inheritance to cell cycle progression. Dev. Cell 22, 52–63. https://doi.org/10.1016/j.devcel.2011.10.014
Skene, P.J., Henikoff, S., 2017. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. eLife 6. https://doi.org/10.7554/eLife.21856
Smoak, E.M., Stein, P., Schultz, R.M., Lampson, M.A., Black, B.E., 2016. Long-Term Retention of CENP-A Nucleosomes in Mammalian Oocytes Underpins Transgenerational Inheritance of Centromere Identity. Curr. Biol. CB 26, 1110–1116. https://doi.org/10.1016/j.cub.2016.02.061
Sogo, J.M., Stahl, H., Koller, T., Knippers, R., 1986. Structure of replicating simian virus 40 minichromosomes. J. Mol. Biol. 189, 189–204. https://doi.org/10.1016/0022-2836(86)90390-6
Spiller, F., Medina-Pritchard, B., Abad, M.A., Wear, M.A., Molina, O., Earnshaw, W.C., Jeyaprakash, A.A., 2017. Molecular basis for Cdk1-regulated timing of Mis18 complex assembly and CENP-A deposition. EMBO Rep. 18, 894–905. https://doi.org/10.15252/embr.201643564
Srivastava, V., Iglesias, P.A., Robinson, D.N., 2015. Cytokinesis: Robust cell shape regulation. Semin. Cell Dev. Biol. 53, 39–44. https://doi.org/10.1016/j.semcdb.2015.10.023
Stankovic, A., Guo, L.Y., Mata, J.F., Bodor, D.L., Cao, X.-J., Bailey, A.O., Shabanowitz, J., Hunt, D.F., Garcia, B.A., Black, B.E., Jansen, L.E.T., 2017. A Dual Inhibitory Mechanism Sufficient to Maintain Cell-Cycle-Restricted CENP-A Assembly. Mol. Cell 65, 231–246. https://doi.org/10.1016/j.molcel.2016.11.021
Stankovic, A., Jansen, L.E.T., 2017. Progress in Molecular and Subcellular Biology. Prog. Mol. Subcell. Biol. 56, 139–162. https://doi.org/10.1007/978-3-319-58592-5_6
Steensel, B. van, Belmont, A.S., 2017. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell 169, 780–791. https://doi.org/10.1016/j.cell.2017.04.022
Stellfox, M.E., Nardi, I.K., Knippler, C.M., Foltz, D.R., 2016. Differential Binding Partners of the Mis18α/β YIPPEE Domains Regulate Mis18 Complex Recruitment to Centromeres. Cell Rep. 15, 2127–2135. https://doi.org/10.1016/j.celrep.2016.05.004

Stephens, P.J., Greenman, C.D., Fu, B., Yang, F., Bignell, G.R., Mudie, L.J., Pleasance, E.D., Lau, K.W., Beare, D., Stebbings, L.A., McLaren, S., Lin, M.-L., McBride, D.J., Varela, I., Nik-Zainal, S., Leroy, C., Jia, M., Menzies, A., Butler, A.P., Teague, J.W., Quail, M.A., Burton, J., Swerdlow, H., Carter, N.P., Morsberger, L.A., Iacobuzio-Donahue, C., Follows, G.A., Green, A.R., Flanagan, A.M., Stratton, M.R., Futreal, P.A., Campbell, P.J., 2011. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144, 27–40. https://doi.org/10.1016/j.cell.2010.11.055

Stimpson, K.M., Matheny, J.E., Sullivan, B.A., 2012. Dicentric chromosomes: unique models to study centromere function and inactivation. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. 20, 595–605. https://doi.org/10.1007/s10577-012-9302-3

Stoler, S., Keith, K.C., Curnick, K.E., Fitzgerald-Hayes, M., 1995. A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. Genes Dev. 9, 573–586. https://doi.org/10.1101/gad.9.5.573

Stoler, S., Rogers, K., Weitze, S., Morey, L., Fitzgerald-Hayes, M., Baker, R.E., 2007. Scm3, an essential Saccharomyces cerevisiae centromere protein required for G2/M progression and Cse4 localization. Proc. Natl. Acad. Sci. 104, 10571–10576. https://doi.org/10.1073/pnas.0703178104

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., Karpen, G.H., 2017. Phase separation drives heterochromatin domain formation. Nature 547, 241–245. https://doi.org/10.1038/nature22989

Subramanian, L., Medina-Pritchard, B., Barton, R., Spiller, F., Kulasegaran-Shylini, R., Radaviciute, G., Allshire, R.C., Jeyaprakash, A.A., 2016. Centromere localization and function of Mis18 requires Yippee-like domain-mediated oligomerization. EMBO Rep. 17, 496–507. https://doi.org/10.15252/embr.201541520

Sugimoto, K., Himeno, M., 1992. Casein kinase II site of human centromere protein B (CENP-B) is phosphorylated in vitro. Biosci. Biotechnol. Biochem. 56, 1174–1175. https://doi.org/10.1271/bbb.56.1174

Sugimoto, K., Tsutsui, M., AuCoin, D., Vig, B.K., 1999. Visualization of prekinetochore locus on the centromeric region of highly extended chromatin fibers: does kinetochore autoantigen CENP-C constitute a kinetochore organizing center? Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. 7, 9–19. https://doi.org/10.1023/a:1009267010071

Sugimoto, K., Yata, H., Muro, Y., Himeno, M., 1994. Human centromere protein C (CENP-C) is a DNA-binding protein which possesses a novel DNA-binding motif. J. Biochem. (Tokyo) 116, 877–881. https://doi.org/10.1093/oxfordjournals.jbchem.a124610

Sullivan, B.A., Karpen, G.H., 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat. Struct. Mol. Biol. 11, 1076–1083. https://doi.org/10.1038/nsmb845

Sullivan, K.F., Hechenberger, M., Masri, K., 1994. Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J. Cell Biol. 127, 581–592. https://doi.org/10.1083/jcb.127.3.581

Sullivan, L.L., Boivin, C.D., Mravinac, B., Song, I.Y., Sullivan, B.A., 2011. Genomic size of CENP-A domain is proportional to total alpha satellite array size at human centromeres and expands in cancer cells. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. 19, 457–470. https://doi.org/10.1007/s10577-011-9208-5

Suzuki, A., Badger, B.L., Salmon, E.D., 2015. A quantitative description of Ndc80 complex linkage to human kinetochores. Nat. Commun. 6, 8161. https://doi.org/10.1038/ncomms9161

Suzuki, N., Nakano, M., Nozaki, N., Egashira, S., Okazaki, T., Masumoto, H., 2004. CENP-B interacts with CENP-C domains containing Mif2 regions responsible for centromere localization. J. Biol. Chem. 279, 5934–5946. https://doi.org/10.1074/jbc.M306477200

Swartz, S.Z., McKay, L.S., Su, K.-C., Bury, L., Padeganeh, A., Maddox, P.S., Knouse, K.A., Cheeseman, I.M., 2019. Quiescent Cells Actively Replenish CENP-A Nucleosomes to Maintain Centromere Identity and Proliferative Potential. Dev. Cell 51, 35-48.e7. https://doi.org/10.1016/j.devcel.2019.07.016

Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.-Y., Kimura, H., Kurumizaka, H., 2011. Crystal structure of the human centromeric nucleosome containing CENP-A. Nature 476, 232–5. https://doi.org/10.1038/nature10258

Tachiwana, H., Miya, Y., Shono, N., Ohzeki, J., Osakabe, A., Otake, K., Larionov, V., Earnshaw, W.C., Kimura, H., Masumoto, H., Kurumizaka, H., 2013. Nap1 regulates proper CENP-B binding to nucleosomes. Nucleic Acids Res. 41, 2869–2880. https://doi.org/10.1093/nar/gks1464

Tachiwana, H., Müller, S., Blümer, J., Klare, K., Musacchio, A., Almouzni, G., 2015. HJURP Involvement in De Novo CenH3CENP-A and CENP-C Recruitment. Cell Rep. 11, 22–32. https://doi.org/10.1016/j.celrep.2015.03.013

Takayama, Y., Sato, H., Saitoh, S., Ogiyama, Y., Masuda, F., Takahashi, K., 2008. Biphasic Incorporation of Centromeric Histone CENP-A in Fission Yeast. Mol. Biol. Cell 19, 682–690. https://doi.org/10.1091/mbc.e07-05-0504

Takeuchi, K., Nishino, T., Mayanagi, K., Horikoshi, N., Osakabe, A., Tachiwana, H., Hori, T., Kurumizaka, H., Fukagawa, T., 2013. The centromeric nucleosome-like CENP-T-W-S-X complex induces positive supercoils into DNA. Nucleic Acids Res. 42, 1644–55. https://doi.org/10.1093/nar/gkt1124

Takizawa, Y., Ho, C.-H., Tachiwana, H., Matsunami, H., Kobayashi, W., Suzuki, M., Arimura, Y., Hori, T., Fukagawa, T., Ohi, M.D., Wolf, M., Kurumizaka, H., 2020. Cryo-EM Structures of Centromeric Tri-nucleosomes Containing a Central CENP-A Nucleosome. Struct. Lond. Engl. 1993 28, 44-53.e4. https://doi.org/10.1016/j.str.2019.10.016

Talbert, P.B., Henikoff, S., 2010. Histone variants — ancient wrap artists of the epigenome. Nat. Rev. Mol. Cell Biol. 11, 264–275. https://doi.org/10.1038/nrm2861

Tan, T., Chen, Z., Lei, Y., Zhu, Y., Liang, Q., 2014. A regulatory effect of INMAP on centromere proteins: antisense INMAP induces CENP-B variation and centromeric halo. PloS One 9, e91937. https://doi.org/10.1371/journal.pone.0091937

Tanaka, Y., Nureki, O., Kurumizaka, H., Fukai, S., Kawaguchi, S., Ikuta, M., Iwahara, J., Okazaki, T., Yokoyama, S., 2001. Crystal structure of the CENP-B protein-DNA complex: the DNA-binding domains of CENP-B induce kinks in the CENP-B box DNA. EMBO J. 20, 6612–6618. https://doi.org/10.1093/emboj/20.23.6612

Tanaka, Y., Tachiwana, H., Yoda, K., Masumoto, H., Okazaki, T., Kurumizaka, H., Yokoyama, S., 2005. Human centromere protein B induces translational positioning of nucleosomes on alpha-satellite sequences. J. Biol. Chem. 280, 41609–41618. https://doi.org/10.1074/jbc.M509666200

Tawaramoto, M.S., Park, S.-Y., Tanaka, Y., Nureki, O., Kurumizaka, H., Yokoyama, S., 2003. Crystal Structure of the Human Centromere Protein B (CENP-B) Dimerization Domain at 1.65-Å Resolution. J. Biol. Chem. 278, 51454–51461. https://doi.org/10.1074/jbc.m310388200

Thongchum, R., Nishihara, H., Srikulnath, K., Hirai, H., Koga, A., 2020. The CENP-B box, a nucleotide motif involved in centromere formation, has multiple origins in New World monkeys. Genes Genet. Syst. 94, 301–306. https://doi.org/10.1266/ggs.19-00042

Tian, T., Li, X., Liu, Y., Wang, C., Liu, X., Bi, G., Zhang, X., Yao, X., Zhou, Z.H., Zang, J., 2018. Molecular basis for CENP-N recognition of CENP-A nucleosome on the human kinetochore. Cell Res. 28, 374–378. https://doi.org/10.1038/cr.2018.13

Tomonaga, T., Matsushita, K., Yamaguchi, S., Oohashi, T., Shimada, H., Ochiai, T., Yoda, K., Nomura, F., 2003. Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. Cancer Res. 63, 3511–3516.

Uhlmann, F., 2000. Chromosome cohesion: A polymerase for chromosome bridges. Curr. Biol. 10, R698–R700. https://doi.org/10.1016/s0960-9822(00)00709-0

Umbreit, N.T., Zhang, C.-Z., Lynch, L.D., Blaine, L.J., Cheng, A.M., Tourdot, R., Sun, L., Almubarak, H.F., Judge, K., Mitchell, T.J., Spektor, A., Pellman, D., 2020. Mechanisms generating cancer genome complexity from a single cell division error. Science 368, eaba0712. https://doi.org/10.1126/science.aba0712

Voullaire, L.E., Slater, H.R., Petrovic, V., Choo, K.H., 1993. A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? Am. J. Hum. Genet. 52, 1153–1163.

Waddington, C.H., 1954. Reviews. Heredity 8, 279–282. https://doi.org/10.1038/hdy.1954.25

Wade, C.M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerre, S., Imsland, F., Lear, T.L., Adelson, D.L., Bailey, E., Bellone, R.R., Blöcker, H., Distl, O., Edgar, R.C., Garber, M., Leeb, T., Mauceli, E., MacLeod, J.N., Penedo, M.C.T., Raison, J.M., Sharpe, T., Vogel, J., Andersson, L., Antczak, D.F., Biagi, T., Binns, M.M., Chowdhary, B.P., Coleman, S.J., Della Valle, G., Fryc, S., Guérin, G., Hasegawa, T., Hill, E.W., Jurka, J., Kiialainen, A., Lindgren, G., Liu, J., Magnani, E., Mickelson, J.R., Murray, J., Nergadze, S.G., Onofrio, R., Pedroni, S., Piras, M.F., Raudsepp, T., Rocchi, M., Røed, K.H., Ryder, O.A., Searle, S., Skow, L., Swinburne, J.E., Syvänen, A.C., Tozaki, T., Valberg, S.J., Vaudin, M., White, J.R., Zody, M.C., Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team, Lander, E.S., Lindblad-Toh, K., 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. Science 326, 865–867. https://doi.org/10.1126/science.1178158

Wang, F., Dai, J., Daum, J.R., Niedzialkowska, E., Banerjee, B., Stukenberg, P.T., Gorbsky, G.J., Higgins, J.M.G., 2010. Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. Science 330, 231–235. https://doi.org/10.1126/science.1189435

Wang, J., Liu, X., Dou, Z., Chen, L., Jiang, H., Fu, C., Fu, G., Liu, D., Zhang, J., Zhu, T., Fang, J., Zang, J., Cheng, J., Teng, M., Ding, X., Yao, X., 2014. Mitotic regulator Mis18β interacts with and specifies the centromeric assembly of molecular chaperone holliday junction recognition protein (HJURP). J. Biol. Chem. 289, 8326–36. https://doi.org/10.1074/jbc.m113.529958

Warburton, P.E., Cooke, C.A., Bourassa, S., Vafa, O., Sullivan, B.A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K.F., Poirier, G.G., Earnshaw, W.C., 1997. Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Curr. Biol. CB 7, 901–904. https://doi.org/10.1016/s0960-9822(06)00382-4

Watanabe, R., Hara, M., Okumura, E., Hervé, S., Fachinetti, D., Ariyoshi, M., Fukagawa, T., 2019. CDK1-mediated CENP-C phosphorylation modulates CENP-A binding and mitotic kinetochore localization. J. Cell Biol. 218, 4042–4062. https://doi.org/10.1083/jcb.201907006

Watanabe, Y., 2005. Shugoshin: guardian spirit at the centromere. Curr. Opin. Cell Biol. 17, 590–595. https://doi.org/10.1016/j.ceb.2005.10.003

Waye, J.S., Willard, H.F., 1987. Nucleotide sequence heterogeneity of alpha satellite repetitive DNA: a survey of alphoid sequences from different human chromosomes. Nucleic Acids Res. 15, 7549–7569. https://doi.org/10.1093/nar/15.18.7549

Weber, C.M., Henikoff, S., 2014. Histone variants: dynamic punctuation in transcription. Genes Dev. 28, 672–82. https://doi.org/10.1101/gad.238873.114

Weir, J.R., Faesen, A.C., Klare, K., Petrovic, A., Basilico, F., Fischböck, J., Pentakota, S., Keller, J., Pesenti, M.E., Pan, D., Vogt, D., Wohlgemuth, S., Herzog, F., Musacchio, A., 2016. Insights

from biochemical reconstitution into the architecture of human kinetochores. Nature 537, 249–253. https://doi.org/10.1038/nature19333
Westhorpe, F.G., Fuller, C.J., Straight, A.F., 2015. A cell-free CENP-A assembly system defines the chromatin requirements for centromere maintenance. J. Cell Biol. 209, 789–801. https://doi.org/10.1083/jcb.201503132
Westhorpe, F.G., Straight, A.F., 2015. The Centromere: Epigenetic Control of Chromosome Segregation during Mitosis. Cold Spring Harb. Perspect. Biol. 7, a015818. https://doi.org/10.1101/cshperspect.a015818
Willard, H.F., Wevrick, R., Warburton, P.E., 1989. Human centromere structure: organization and potential role of alpha satellite DNA. Prog. Clin. Biol. Res. 318, 9–18.
Williams, B.C., Murphy, T.D., Goldberg, M.L., Karpen, G.H., 1998. Neocentromere activity of structurally acentric mini-chromosomes in Drosophila. Nat. Genet. 18, 30–38. https://doi.org/10.1038/ng0198-30
Williams, J.S., Hayashi, T., Yanagida, M., Russell, P., 2009. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. Mol. Cell 33, 287–298. https://doi.org/10.1016/j.molcel.2009.01.017
Wisniewski, J., Hajj, B., Chen, J., Mizuguchi, G., Xiao, H., Wei, D., Dahan, M., Wu, C., 2014. Imaging the fate of histone Cse4 reveals de novo replacement in S phase and subsequent stable residence at centromeres. eLife 3, e02203. https://doi.org/10.7554/eLife.02203
Wong, L.H., Brettingham-Moore, K.H., Chan, L., Quach, J.M., Anderson, M.A., Northrop, E.L., Hannan, R., Saffery, R., Shaw, M.L., Williams, E., Choo, K.H.A., 2007. Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. Genome Res. 17, 1146–1160. https://doi.org/10.1101/gr.6022807
Xiao, H., Mizuguchi, G., Wisniewski, J., Huang, Y., Wei, D., Wu, C., 2011. Nonhistone Scm3 binds to AT-rich DNA to organize atypical centromeric nucleosome of budding yeast. Mol. Cell 43, 369–380. https://doi.org/10.1016/j.molcel.2011.07.009
Xiao, H., Wang, F., Wisniewski, J., Shaytan, A.K., Ghirlando, R., FitzGerald, P.C., Huang, Y., Wei, D., Li, S., Landsman, D., Panchenko, A.R., Wu, C., 2017. Molecular basis of CENP-C association with the CENP-A nucleosome at yeast centromeres. Genes Dev. 31, 1958–1972. https://doi.org/10.1101/gad.304782.117
Xu, T., Park, S.K., Venable, J.D., Wohlschlegel, J.A., Diedrich, J.K., Cociorva, D., Lu, B., Liao, L., Hewel, J., Han, X., Wong, C.C.L., Fonslow, B., Delahunty, C., Gao, Y., Shah, H., Yates, J.R., 2015. ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. J. Proteomics 129, 16–24. https://doi.org/10.1016/j.jprot.2015.07.001
Yamagishi, Y., Honda, T., Tanno, Y., Watanabe, Y., 2010. Two histone marks establish the inner centromere and chromosome bi-orientation. Science 330, 239–243. https://doi.org/10.1126/science.1194498
Yang, C.H., Tomkiel, J., Saitoh, H., Johnson, D.H., Earnshaw, W.C., 1996. Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. Mol. Cell. Biol. 16, 3576–3586. https://doi.org/10.1128/mcb.16.7.3576
Yoda, K., Ando, S., Okuda, A., Kikuchi, A., Okazaki, T., 1998. In vitro assembly of the CENP-B/α-satellite DNA/core histone complex: CENP-B causes nucleosome positioning. Genes Cells 3, 533–548. https://doi.org/10.1046/j.1365-2443.1998.00210.x
Yoda, K., Kitagawa, K., Masumoto, H., Muro, Y., Okazaki, T., 1992. A human centromere protein, CENP-B, has a DNA binding domain containing four potential alpha helices at the NH2 terminus, which is separable from dimerizing activity. J. Cell Biol. 119, 1413–1427. https://doi.org/10.1083/jcb.119.6.1413
Yu, Z., Zhou, X., Wang, W., Deng, W., Fang, J., Hu, H., Wang, Z., Li, S., Cui, L., Shen, J., Zhai, L., Peng, S., Wong, J., Dong, S., Yuan, Z., Ou, G., Zhang, X., Xu, P., Lou, J., Yang, N., Chen, P., Xu, R.-M., Li, G., 2015. Dynamic phosphorylation of CENP-A at Ser68 orchestrates its

cell-cycle-dependent deposition at centromeres. Dev. Cell 32, 68–81. https://doi.org/10.1016/j.devcel.2014.11.030

Zaratiegui, M., Vaughn, M.W., Irvine, D.V., Goto, D., Watt, S., Bähler, J., Arcangioli, B., Martienssen, R.A., 2011. CENP-B preserves genome integrity at replication forks paused by retrotransposon LTR. Nature 469, 112–115. https://doi.org/10.1038/nature09608

Zasadzińska, E., Barnhart-Dailey, M.C., Kuich, P.H.J.L., Foltz, D.R., 2013. Dimerization of the CENP-A assembly factor HJURP is required for centromeric nucleosome deposition. EMBO J. 32, 2113–24. https://doi.org/10.1038/emboj.2013.142

Zasadzińska, E., Foltz, D.R., 2017. Progress in Molecular and Subcellular Biology. Prog. Mol. Subcell. Biol. 56, 165–192. https://doi.org/10.1007/978-3-319-58592-5_7

Zasadzińska, E., Huang, J., Bailey, A.O., Guo, L.Y., Lee, N.S., Srivastava, S., Wong, K.A., French, B.T., Black, B.E., Foltz, D.R., 2018. Inheritance of CENP-A Nucleosomes during DNA Replication Requires HJURP. Dev. Cell 47, 348-362.e7. https://doi.org/10.1016/j.devcel.2018.09.003

Zhang, W., Friebe, B., Gill, B.S., Jiang, J., 2010. Centromere inactivation and epigenetic modifications of a plant chromosome with three functional centromeres. Chromosoma 119, 553–63. https://doi.org/10.1007/s00412-010-0278-5

Zhao, H., Winogradoff, D., Bui, M., Dalal, Y., Papoian, G.A., 2016. Promiscuous Histone Mis-Assembly Is Actively Prevented by Chaperones. J. Am. Chem. Soc. 138, 13207–13218. https://doi.org/10.1021/jacs.6b05355

Zhu, J., Cheng, K.C.L., Yuen, K.W.Y., 2018. Histone H3K9 and H4 Acetylations and Transcription Facilitate the Initial CENP-AHCP−3 Deposition and De Novo Centromere Establishment in Caenorhabditis elegans Artificial Chromosomes. Epigenetics Chromatin 11, 16. https://doi.org/10.1186/s13072-018-0185-1

12. ABBREVIATION REGISTER

ACA	Anti-Centromere-Autoantibody
AFG	African Green Monkey
AID	Auxin Inducible Degron
APC/C	Anaphase Promoting Complex/Cyclosome
AS	Analog Sensitive
ASF1	Anti Silencing Function
ASH1L,	Achaete-Scute Homologue-1-Like histone-lysine N-methyltransferase
ATRX	Alpha thalassemia/mental retardation syndrome X-linked
BAC	Bacterial Artificial Chromosome
BF	Bright Field
BFB	Breakage-Fusion-Bridge
bp	Base Pair
BSA	Bovine Serum Albumin
Bub1/3	Budding Uninhibited by Benzimidazoles 1/3
CAF-1	Chromatin Assembly Factor 1
CATD	CENP-A Targeting Domain
CCAN	Constitutive Centromere Associated Network
CD4	Cluster of Differentiation 4
Cdc20/42	Cell Division Cycle 20/42
CDE	Centromere DNA Element
CDK1	Cyclin Dependent Kinase 1
CenDNA	CENtromeric DNA
CENP	CENtromere Protein
CENP-A^{EA}	CENP-A$^{EYFP\text{-}AID}$
CFSE	Carboxyfluorescein succinimidyl ester
ChIP-Seq	Chromatin ImmunoPrecipitation in combination with Sequencing
CREST	Calcinosis/Raynaud's phenomenon/Esophageal dysmotility/Sclerodactyly/ Telangiectasia
CRISPR/Cas9	Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated protein 9

CUT&RUN	Cleavage Under Targets and Release Using Nuclease
DAPI	4′,6-DiAmidino-2-PhenylIndole
DAXX	Death Domain Assoicated Protein
DBD	DNA Binding Domain
DHFR	DiHydroFolate Reductase
DLD-1	cell line derived from human colorectal adenocarcinoma
DNA	DesoxyriboNucleic Acid
DOX	Doxycycline
DSN1	Dosage Suppressor of NNF1
DTT	DiThioThreitol
eCENP-A/C	exogenous CENP-A/C
Ect2	epithelial cell transforming 2
ENC	Evolutionary New Centromeres
et al.	et alii, et aliae (and others)
EYFP	Enhanced Yellow Fluorescent Protein
FACS	Fluorescence Activated Cell Sorting
FACT	FAcilitates Chromatin Transcription
FBS	Fetal Bovine Serum
FISH	Fluorescence *In Situ* Hybridization
FL	Full Length
GAP	GTPase activating protein
GFP	Green Fluorescent Protein
GST	Glutathione S-Transferase
GTP	GuAnosine-5'-Triphosphate
h (prefix)	human
HAC	Human Artificial Chromosomes HAC
HAT1	Histone AcetylTransferase 1 HAT1
HCTD	HJURP Carboxy Terminal Domains
HILO	Highly Inclined and Laminated Optical sheet
HJURP	Holliday Junction Recognizing Protein
HOR	Higher Order Repeat
HP1α	Heterochromatin Protein 1 α

HU protein	Histone U93 protein
IAA	Indole-3-acetic acid
ICF	Immunodeficiency, Centromeric region instability, Facial anomalies
IF	ImmunoFluorescence Microscopy
IgG	Immunoglobulin G
K-fiber	Kinetochore Fiber
KAT7	Lysine Acetyltransferase 7
kb	kilo-bases
KMN	KNL1, Mis12, Ndc80 complex
KNL1	Kinetochore Null Protein 1
M18BP1	Mis18 Binding Protein 1
Mad2/3	Mitotic arrest deficient 2/3
mb	mega-bases
MCAK	Mitotic Centromere Associated Kinesin
MCC	Mitotic Checkpoint Complex
MCM2-7	Mini Chromosome Maintenance
MEF	MEF
Mis12	MIS-segregation 12
MIS12/18	MISsegregation 12/18
MNase	Micrococcal Nuclease
MPS1	MonoPolar Spindle 1
mRFP	monomeric Red Fluorescent Protein
Ms	mouse
NAP1	Nucleosome Assembly Protein 1
Ndc80	Nuclear division cycle 80
NMR	Nuclear magnetic resonance
NSL1	NonSpecific Lethal 1
NT	Non-Treated
NuRD	Nucleosome Remodeling and Deactylase
OD	Opitical Density
OE	Over-Expression
os-TIR1	*Oryza sativa*-Transport Inhibitor Response 1

PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate Buffered Saline
PCNA	Proliferating Cell Nuclear Antigen
PD-NC4	PseudoDicentric NeoCentromere 4
Plk1	Polo-like kinase 1
PMF1	Polyamine Modulated Factor 1
PMSF	PhenylMethylSulfonyl Fluoride
PTM	Post Translational Modifications
qPCR	quantitative Polymerase Chain Reaction
Rb	rabbit
RbAp46/48	Retinoblastoma protein Associated protein 46/48
RC	Replication Coupled
RI	Replication Independent
RNA	RiboNucleic Acid
RNAi	RNA interference,
RPA	Replication Protein A
RPE-1	Retinal Pigment Epithelia-1
RPS4Y	Ribosomal Protein S4, Y-linked
SAC	Spindle Assembly Checkpoint
Sap1	Switch-activating protein 1
Scm3	Suppressor of chromosome mis-segregation
SD	Standard Deviation
SDS-PAGE	Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis
SEM	Standard Error of the Mean
SENP	Sumoylation-Specific Proteases
Sf9	*Spodoptera frugiperda*
siRNA	small interfering RNA
SMM	single molecule microscopy
Suv39h	Suppressor of variegation 3–9 homolog
TCR	T Cell Receptor
TMP	TriMethoPrim
TRE	Tetracycline-Responsive Element

TSA	Trichostatin A
U-2 OS	Human Bone Osteosarcoma Epithelial Cells
WAPL	Wing APart-Like
WO	Wash-Out
Y2H	Yeast-two-hybrid

13. CURRICULUM VITAE

- **PERSONAL INFORMATION**

Family name, First name: **HOFFMANN**, Sebastian Gert Albrecht
Date of birth: 01 January 1990
Nationality: German
Address: 72 Avenue de Flandre, 75019 Paris, France
In Germany: Heimstättenweg 8, 12163 Berlin, DE

- **EDUCATION**

2016 PhD Cell Biology, Institut Curie, UMR144, FRANCE (Advisor: Dr. *Daniele Fachinetti*)

2015 Master of Science in Biochemistry and Molecular Biology, University of Potsdam, GERMANY.

2012 Bachelor of Science in Life Science, University of Potsdam, GERMANY

- **CURRENT POSITION**

2016 – present PhD-student (PSL Research University) at Institut Curie, FRANCE.

- **PREVIOUS POSITIONS**

2015 – 2016 Research Engineer (full time job), Cell Biology, Institut Curie, UMR144, FRANCE. Advisor: Dr. *Daniele Fachinetti*

2011 – 2014 Student Research Assistant (part time job), Plant Biology, Dahlem Centre of Plant Sciences (FU-Berlin), GERMANY.

- **FELLOWSHIPS AND AWARDS**

2019 Funding for the 4th year of PhD in France awarded by the French foundation for medical research (FRM)

2019 Poster Price at EMBO workshop on Chromosome Segregation and Aneuploidy in Lisbon, PT

2016 international PhD fellowship (IC3i) co-founded by Institut Curie and the European Unions's Horizon 2020 research and innovation program

2015 best student of the year in Biochemistry and Molecular Biology of the University of Potsdam (Germany)

2014 PROMOS scholarship (Program to promote the mobility of German students) awarded by the German Academic Exchange Service (DAAD) for the performance of the master thesis at the Ludwig Institute for Cancer Research (San Diego, USA)

- **DOCTORAL MISSION AND ORGANISATION OF MEETINGS**

- Organization of LIBRA events at Institut Curie (LIBRA - Unifying innovative efforts of European research centres to achieve gender equality in academia. https://www.eu-libra.eu/)
- Support of the training unit at Institut Curie
- Member of the organization team for the joint retreat of young researchers of Institut Curie/Pasteur and CEITEC in Brno (Czech Republic)

14. PUBLICATIONS

S. HOFFMANN, H.M. Izquierdo, R. Gamba, F. Chardon, M. Dumont, V. Keizer, S. Hervé, S.M. McNulty, B.A. Sullivan, N. Manel and D. Fachinetti (2020), A genetic memory initiates the epigenetic loop necessary to preserve centromere position. **EMBO Journal** 39: e105505

V. Barra, G. A. Logsdon, A. Scelfo, **S. HOFFMANN**, S. Hervé, A. Aslanian, Y. Nechemia-Arbely, D.W. Cleveland, B. E. Black and Daniele Fachinetti (2019), Phosphorylation of CENP-A on serine 7 does not control centromere function**. Nature Communications *10(1)*:175**

S. HOFFMANN, D. Fachinetti (2018), book chapter: Real Time *de novo* Deposition of Centromeric Histone-associated Proteins Using the Auxin Inducible Degradation system in *Histone Variants: Methods and Protocols* Methods In Molecular Biology **Springer Science+Business Media, LLC, New York**

S. HOFFMANN, D. Fachinetti (2017). A time-out for CENP-A. **Molecular & Cellular Oncology**

S. HOFFMANN*, M. Dumont*, V. Barra, P. Ly, Y. Nechemia-Arbely, M. A. McMahon, S. Hervé, D.W. Cleveland, D. Fachinetti (2016) (*Equal Contribution). CENP-A Is Dispensable for Mitotic Centromere Function after Initial Centromere/Kinetochore Assembly. **Cell Reports** *17*, 2394–2404

RÉSUMÉ

Au cours de la mitose, les cellules de mammifères doivent conserver un centromère unique afin d'assurer une ségrégation correcte des chromatides sœurs dans chaque cellule. Malgré la présence de séquences d'ADN répétitives dans la plupart des organismes, les centromères sont marqués épigénétiquement par une histone H3 spécifique appelée CENP-A (CENtromeric Protein A) via un mécanisme d'autoréplication. Il n'est pourtant pas certain que CENP-A soit l'unique marqueur de la position du centromère chez l'Homme ou si d'autres facteurs, tels que la séquence d'ADN, y contribuent également.

Afin d'évaluer cette hypothèse, j'ai récemment développé le système CENP-A$^{OFF/ON}$ qui permet d'éliminer l'identité du centromère dans les cellules humaines en quelques minutes. En utilisant ce système unique, je peux identifier le(s) mécanisme(s) qui permet(tent) la réintégration *de novo* de CENP-A à sa position initiale et la formation du centromère qui en résulte, en temps réel, sur des chromosomes humains. Il est intéressant de noter que c'est la présence de la protéine CENP-B liée à l'ADN répétitif qui va permettre la réintégration de CENP-A, de nouveau exprimée, à la position initiale. En revanche, l'absence de CENP-B et d'autres marqueurs épigénétiques CENP-A-dépendants au centromère, empêchent toute nouvelle déposition de CENP-A ce qui entraîne la formation d'un néocentromère. La spécification du centromère liée à CENP-B a lieu, en partie, indépendamment de CENP-C, composant clé du centromère connu pour son rôle dans le recrutement de CENP-A. Cet ensemble de résultats démontrent que le mécanisme bien établi d'auto-assemblage de CENP-A n'est pas essentiel pour le maintien de l'identité du centromère et que CENP-B joue un rôle clé dans le maintien de la position du centromère. Ainsi, ceci permet d'accroître nos connaissances, actuellement restreintes, concernant les procédés (épi)génétiques qui contrôlent l'identité du centromère, essentiel pour la transmission correcte du matériel génétique. Enfin, dans le cadre de cette thèse, nous avons identifié une population de cellules T CD4+ au repos CENP-A-négatives, CENP-B/C-positives, capables de ré-exprimer et de réassembler le CENP-A à l'entrée du cycle cellulaire, ce qui démontre l'importance physiologique de la mémoire génétique.

MOTS CLÉS

CENP-B, centromère identité, CENP-A, CENP-C, l'ADN alpha satellite, la stabilité du génome

ABSTRACT

Centromeres are built on repetitive DNA sequences (CenDNA) and a specific chromatin enriched with the histone H3 variant CENP-A, the epigenetic mark that identifies centromere position. During my thesis project, I interrogated the importance of CenDNA in centromere specification by developing a system to rapidly remove and re-activate CENP-A (CENP-A$^{OFF/ON}$). Using this system, I define the temporal cascade of events necessary to maintain centromere position. I unveil that CENP-B bound to CenDNA provides memory for maintenance on human centromeres by promoting *de novo* CENP-A deposition. Indeed, the lack of CENP-B favors neocentromere formation under selective pressure. Occasionally, CENP-B triggers centromere re-activation initiated by CENP-C, but not CENP-A, recruitment at both ectopic and native centromeres. This is then sufficient to initiate the CENP-A-based epigenetic loop. Finally, in the frame of this thesis, we identified a population of CENP-A-negative, CENP-B/C-positive resting CD4^{+} T cells capable to re-express and reassembles CENP-A upon cell cycle entry, demonstrating the physiological importance of the genetic memory.

KEYWORDS

CENP-B, centromere identity, CENP-A, CENP-C, alpha satellite DNA, genome stability