Intricate structure of the interphase chromocenter revealed by proximity-based biotinylation

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Betreuer: Prof. Dr. rer. nat. Axel Imhof

Zweitgutachter: Univ. Prof. Dr. Jovica Ninkovic

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 09.10.2020

Eidesstattliche Versicherung

Kochanova, Natalia Yurievna

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Thema

"Intricate structure of the interphase chromocenter revealed by the analysis of factor involved in species formation"

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

____Edinburgh, 14.10.2020__   ___Natalia Y. Kochanova_____

Ort, Datum   Unterschrift Doktorandin/Doktorand

Publications resulting from this study:

Preprint:

<u>Kochanova, N.Y.,</u> Schauer, T., Mathias, G.P., Lukacs, A., Schmidt, A., Flatley, A., Schepers, A., Thomae, A.W., and Imhof, A. (2018). Intricate structure of the interphase chromocenter revealed by the analysis of a factor involved in species formation. bioRxiv.

The citations from the publication (p. 109) are indicated in italic.

Collaborations:

Dr. Andreas Thomae – STED microscopy, pioneer experiment with dCenpC knockdown, cloning of 4 constructs;

Dr. Tamas Schauer – bioinformatics analysis;

Dr. Andreas Schmidt and Dr. Ignasi Forne – mass spectrometry;

Angelika Zabel – ChIP libraries' preparation.

Dr. Stefan Krebs – sequencing

Table of Contents.

Summary.

Centromeres and adjacent pericentric heterochromatin are sometimes defined as a chromocenter, due to their structural similarity with fused regions traditionally spotted on *Drosophila* polytene chromosomes. The architecture of the chromocenter has been extensively studied in mammals, *Drosophila* and other organisms. In this thesis, we contribute to this field using recently developed state of art methods: high-resolution STED microscopy and APEX2 proximity labeling. We dissect the structures of individual centromeric protein domains, and find that they are separated on STED. Moreover, we perform the proximity labeling using APEX2 fusions with proteins that could be visually separated using high-resolution microscopy. We generate a molecular map of HMR, dCenpA and HP1a bound chromatin and suggest new players in centromere biology. Furthermore, our microscopy and proximity labeling results suggest that HMR forms a boundary between dCenpA and HP1a chromatin.

We also address the question of speciation, where the centromeric protein HMR has a role. Using proximity labeling, we find condensin and cohesin complexes in proximity to HMR and by the analysis of the ChIP-sequencing data point to the molecular signs that HMR's function in pure species as well as hybrids might be connected with chromosome cohesion and condensation. Furthermore, simulating the hybrid situation, we perform ChIP-sequencing of the condensin subunit CAPH2 upon HMR+LHR overexpression, and find that the condensin binding to chromatin in these conditions is reduced.

Zusammenfassung.

Chromosomale Centromere und das benachbarte perizentrische Heterochromatin werden aufgrund ihrer strukturellen Ähnlichkeit mit genomischen Bereichen, die ursprünglich auf Drosophila-Polytän-Chromosomen entdeckt wurden, manchmal als Chromozentrum definiert. Die Architektur des Chromozentrums wurde sowohl in Säugetieren als auch in Insekten und Pflanzen ausführlich untersucht. In dieser Arbeit wurden neuartige Methoden wie hochauflösende STED-Mikroskopie und APEX2-Proximity-Labeling zur Aufklärung der Struktur des Chromozentrums verwendet. Wir haben die Strukturen einzelner sehr nahe zusammenliegender zentromerischer Proteindomänen untersucht, die wir mit Hilfe der STED Mikroskopie voneinander trennen konnten. Darüber hinaus führen wir eine sogenannte Proximity-Markierung mithilfe von APEX2-Fusionsproteinen um dadurch eine molekulare Karte von HMR-, dCenpA- und HP1a-gebundenem Chromatin zu erstellen. Dies führte zur Identifizierung neuer zentromerischer oder perizentromerischer Proteine. Darüber hinaus legen unsere Ergebnisse im Bezug auf Mikroskopie und Proximity-Markierung nahe, dass HMR eine Grenze zwischen dCenpA und HP1a-Chromatin bildet.

Wir beschäftigen uns auch mit der Frage der Artbildung, bei der das das Zentromer-bindende Protein HMR eine wichtige Rolle spielt. Mithilfe der Proximity-Markierung finden wir Condensin- und Cohesin-Komplexe in der Nähe von HMR. Durch Analyse von ChIP-Sequenzierungsdaten fanden wir Hinweise dafür, dass die Funktion von HMR in reinen Spezies und Hybriden mit der Chromosomenkohäsion und -kondensation verbunden sein könnten. Weiterhin stellten wir durch Simulation der Hybridsituation durch Überexpression von HMR + LHR einer verringerte Chromatinbindung einer Condensinuntereinheit fest, was den Phänotyp in Hybriden Fliegen erklären könnte.

Acknowledgements and gratitudes.

Last but not least, I am grateful to numerous cups of coffee and all who bought the Moka pots for our kitchen. Without it I wouldn't have been able to stay awake in difficult times with a lot of work going on.

1 Introduction.

1.1 Chromatin.

DNA molecules, packed together with different RNAs and proteins in the nucleus and mitochondria, are called chromatin. In higher eukaryotes the prominent structural feature of nuclear chromatin are the nucleosomes, around which DNA is packed as a string around the beads. Nucleosomes are octamers, consisting of histone proteins: two molecules of histone H3, two of H4, two of H2A and two of H2B. The nucleosome is organized in a way, that an H3-H4 tetramer is held between two dimers of H2A-H2B. 146 base pairs of DNA are wrapped around one nucleosome (Luger et al., 1997). Histone tails can be modified, and mark different chromatin states (Filion et al., 2010). Also histones in nucleosomes can be substituted by the respective histone variants, which are encoded by the separate genes (Franklin and Zweidler, 1977).

Lots of proteins with different functions are a part of chromatin. "Reader" proteins can recognize histone marks, "writers" can deposit them and "erasers" can remove them (reviewed in (Cosgrove, 2012)). Chromatin remodelers move nucleosomes back and forth on DNA. Transcription, DNA damage and replication proteins are responsible for respective processes in the nucleus.

Chromatin also contains a lot of small and large non-coding RNAs. Those can play a structural role (Schubert et al., 2012) or be a part of active chromatin complexes (Akhtar et al., 2000; Franke and Baker, 1999).

2 meters of DNA are tightly packed in a several µm nucleus. Apart from compaction on the nucleosome level, DNA is organized into the higher-order structure of loops kilobases-megabases in size, packed into globules. The most known type of these globules are called topologically associated domains, or TADs (Dixon et al., 2012; Nora et al., 2012; Rao et al., 2014).

1.2 Nuclear domains and their features.

Linear chromatin domains are stretches of specific DNA composition, DNA modification or/and protein composition along the DNA (reviewed in (Bickmore and van Steensel, 2013)). Active and inactive protein linear domains

tend to cluster with domains of the same activity in 3D (reviewed in (Bickmore and van Steensel, 2013)), and some major proteins, which shape a subset of these domains, are predominantly focused together in one or several places in the nucleus, which can be visualised by immunofluorescence staining. Also, a considerable number of protein linear domains 10-500 kb long were shown to make separate clusters in the nucleus by FISH in *Drosophila* Kc167 cells (Boettiger et al., 2016). Different linear domains were shown to correlate with separate domains in 3D (Sexton et al., 2012).

Chromosomal domains are compartments within one chromosome, while nuclear domains are compartments within the nucleus, which comprise one or several chromosomal domains. For example, chromosomal territories, including the X-chromosome territory (Straub et al., 2005) and the 4th chromosome territory in *Drosophila melanogaster* (Larsson et al., 2001; Riddle et al., 2012), contain only one chromosome, while pericentromeric chromatin from all chromosomes is clustered into one or several constitutive heterochromatin domains in the interphase nucleus (Brown, 1996; Chiolo et al., 2011; Eissenberg and Reuter, 2009). Similarly, the centromeres from different chromosomes are clustered together into several foci near the nucleoli (Imakaev et al., 2012; Padeken et al., 2013; Wiblin et al., 2005). Together, centromeres and pericentric heterochromatin often coalesce into one or several intranuclear domains, named chromocenters (Fig. 1.1).
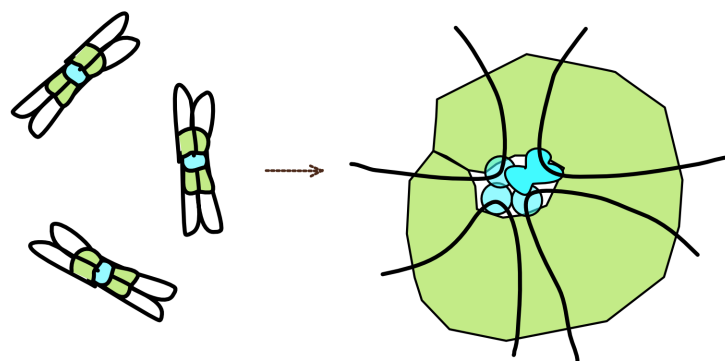
Fig. 1.1. Centromeres and pericentric heterochromatin often coalesce into chromocenter in interphase cells. Figure concept from (Jagannathan et al., 2018).

The nucleus is structurally and functionally divided. Compartments in the nucleus, which could be distinguished by immunofluorescence staining of specific proteins, perform distinct, still sometimes overlapping functions. Nucleoli serve several activities, including being a place for ribosomal genes' transcription and ribosome assembly (reviewed in (Pederson, 2011)). Heterochromatin mainly represses transcription of some genes and transposable elements, promotes the expression of a subset of genes within the compartment and contributes to the centromere stability (reviewed in (Allshire and Madhani, 2018)). Centromeres serve as a site for kinetochore assembly and subsequent attachment of microtubules to chromosomes during mitosis (reviewed in (Muller and Almouzni, 2017)). Inactive female X-chromosome in mammals, male X-chromosome in *Drosophila* and both X-s in *Caenorhabditis elegans* hermaphrodites are regulated chromosome-wide to promote the accurate expression level of X-chromosomal genes (reviewed in (Disteche, 2012; Straub and Becker, 2007)). Nuclear speckles in mammalian cells ensure correct splicing of the processed mRNA (reviewed in (Spector and Lamond, 2011)). Cajal bodies contribute to snRNP maturation (reviewed in (Morris, 2008)). The prominent function of the promyelocytic leukaemia (PML) bodies in mammalian cells is still unclear (reviewed in (Lallemand-Breitenbach and de The, 2010)). Polycomb proteins, which repress transcription of developmental genes, are clustered in the nucleus in so-called Polycomb bodies (reviewed in (Pirrotta and Li, 2012)). RNA polymerase II (PolII) clusters in foci called transcription factories, which by number are far behind the number of PolII occupied genes (reviewed in (Eskiw et al., 2010)).

The domain structure is hierarchical. For example, interphase chromosome territories contain smaller units of DNA organisation – DNA domains of different sizes, including well-known topologically associated domains (TADs) (Dixon et al., 2012; Nora et al., 2012; Rao et al., 2014). TADs, in turn, contain sub-TADs, loops and insulation neighbourhoods (reviewed in (Dixon et al., 2016)). One of the models, which our genome folding on kb-mb scale follows, is a model of the fractal globule, where smaller domains gradually interact to fold into higher and higher order structures (Lieberman-Aiden et al., 2009). Heterochromatic protein 1a (HP1a) heterochromatin contains a network of sub-domains (Swenson et al., 2016). All linear domains in *Drosophila* Kc167

cells can be classified into 5 chromatin states according to their protein composition (Filion et al., 2010), while another model in S2 cells, which is based on histone modification patterns, proposes 9 states (Kharchenko et al., 2011). Interestingly, 8 of these states are "sub-states" of the previous 5-state model (reviewed in (Bickmore and van Steensel, 2013)).

The architecture of at least some domains is dynamic. HP1a heterochromatin in *Drosophila* S2 cells forms from one to three domains, depending on the cell cycle stage (Chiolo et al., 2011). Sub-domains within HP1a heterochromatin in the same cells display different staining pattern throughout the cell cycle (Swenson et al., 2016). TADs disappear during mitosis and form again in the interphase (Gibcus et al., 2018; Naumova et al., 2013). One more example comes from the fact, that linear lamina-associated domains and nucleoli-associated domains in *Drosophila* cells overlap, and the same linear domains contact either the nucleoli or the nuclear lamina throughout cell divisions (Kind et al., 2013).

1.3 Centromeres in interphase and mitosis.

The centromere was initially identified as a place on the chromosome in mitosis where microtubules attach to promote chromosomal segregation. With the discovery of the centromere-specific histone variant Cenp-A (CID/dCenpA (centromere identifier, *Drosophila* centromere protein A) in *Drosophila,* CenH3 in yeast and HCP-3 in *C. elegans*) and its immunofluorescence it turned out that the same centromeric domain is also preserved in the interphase (Earnshaw et al., 1986; Palmer et al., 1987). Later on 16 centromeric proteins in mammalian cells, which associate with centromeres throughout the cell cycle, were discovered by Cenp-A affinity purification and named the constitutive centromere-associated network (CCAN) (Foltz et al., 2006; Izuta et al., 2006; Obuse et al., 2004) (Fig. 1.2). Affinity purification of *Drosophila* CID chromatin did not reveal proteins specifically associated with centromeres (except the known centromeric proteins Cenp-C (centromere protein C) and Cal1 (chromosome alignment defect 1)), but 10 proteins were associated both with centromeres (overlapping with centromeric domains recognisable by immunofluorescence) and the other parts of the nucleus in either interphase or

mitosis (Barth et al., 2014). The almost absence of CCAN-like CID interactors in *Drosophila* implies that many of the fly centromeric proteins might share centromeric and non-centromeric functions and belong to several chromatin domains. Not mutually exclusively, CCAN of *Drosophila* could have, if any, weak or transient interaction with CID.

Centromeric chromatin can occupy a different portion of the chromosome depending on the organism. In budding yeast, the centromere is represented by a single nucleosome containing CenH3 positioned on a particular DNA sequence. In mammals, fission yeast and *Drosophila* CenH3 nucleosomes are positioned within a restricted region of DNA, which contains a lot of repetitive sequences (reviewed in (Muller and Almouzni, 2017; Steiner and Henikoff, 2015; Talbert et al., 2018)). In *C. elegans*, some insects and some plants cenH3 nucleosomes are distributed throughout the chromosomes, forming so-called holocentromeres. Interestingly, despite differences in the architecture of centromeric domain, on transmission electron microscopy, principles of kinetochore architecture are similar between mono- and holocentric organisms (reviewed in (Maddox et al., 2004)).

Centromeres contain a lot of proteins, and, to our knowledge, one major complex protein pathway-network has so far been studied to be important for centromeric architecture. It includes Cenp-A histone variant, which was shown to be essential for the recruitment of several other centromeric proteins and eventually for the formation of kinetochore. Cenp-A is incorporated into nucleosomes by the dedicated chaperone (Holiday junction recognition protein (HJURP) in mammals and Cal1 in *Drosophila*). Over 30 known proteins and their modifications are involved in maintaining the proper Cenp-A level at centromeres and/or supporting the kinetochore architecture, and some of the proteins share non-centromeric functions (reviewed in (Muller and Almouzni, 2017)). In *Drosophila*, recruitment of CID to an ectopic site on the chromosome leads to the formation of the ectopic centromere and the functional kinetochore, which contains at least some of canonical kinetochore proteins (Heun et al., 2006; Mendiburo et al., 2011). In mammalian cells the Cenp-A induction is not sufficient for the ectopic centromere and kinetochore formation (Lacoste et al., 2014). A pathway important not for centromere architecture, but organisation in *Drosophila,* includes Nlp (nucleoplasmin-like protein), Modulo and CTCF

(CCCTC-binding factor) proteins, which are necessary for centromere clustering. All three proteins interact with CID (Padeken et al., 2013).

1.4 Pericentric heterochromatin in interphase and mitosis.

Pericentric chromatin is marked by a network of heterochromatic proteins, including HP1a, and their interactors. Around 50 proteins have been found to localise to pericentromeres in mammals (reviewed in (Fodor et al., 2010), (Saksouk et al., 2015)), and around 390 gene loci have been found to be involved in position effect variegation and heterochromatin stability in *Drosophila* genetic screens (considering that one of the five known heterochromatic pathways-networks refers to pericentric heterochromatin) (reviewed in (Fodor et al., 2010)). This suggests that pericentric heterochromatin is a more complex structure than a centromere or that it is studied much better.

In *Drosophila* HP3 (LHR, lethal hybrid rescue), HP4, HP5 and HP6 (heterochromatic proteins 3, 4, 5, 6) are targeted to heterochromatin by HP1a (Greil et al., 2007). HP1a anchors on H3K9me3 nucleosomes with its chromodomain and dimerizes with its chromoshadow domain, bringing the nucleosomes closer together (Canzio et al., 2011), reviewed in (Eissenberg and Elgin, 2014) (Fig. 1.2). H3K9me3, in turn, is deposited by a heterochromatic methyltransferase Su(var)3-9 (suppressor of variegation 3-9), and in *Drosophila* HP1a and Su(var)3-9 localisation to the chromocenter is mutually dependent (reviewed in (Schotta et al., 2002)). This is an essential pathway-network for heterochromatin architecture. Tethering of HP1a to an ectopic locus in *Drosophila* can create a silenced locus, although independent on Su(var)3-9 dosage (Li et al., 2003). Except the essential H3K9me3 other histone modifications (e.g. H4K20me3 (Schotta et al., 2004)), as well as DNA methylation, are found at heterochromatin and some of these hallmarks were shown to be molecularly linked to the H3K9me3 pathway-network (see below). H3K9me3 remains at heterochromatin during the cell cycle, while HP1a (and HP1β) is ejected from nucleosomes in mitosis because H3 in addition to methylation becomes phosphorylated (Fischle et al., 2005).

Notably, pericentric chromatin is not flanking centromeres in holocentric organisms (Garrigues et al., 2015). In budding yeast, heterochromatin lacks H3K9 methylation and is defined by hypoacetylated histones (reviewed in (Saksouk et al., 2015)).
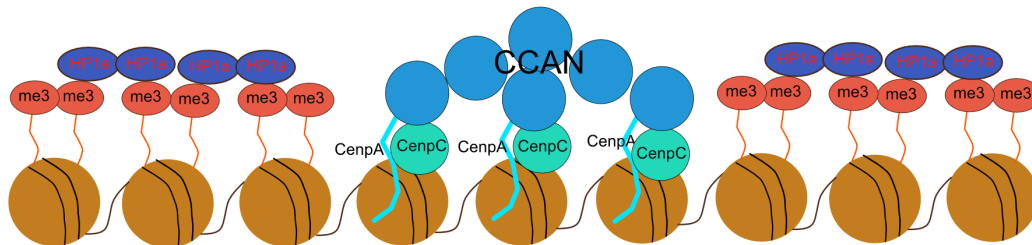
Fig. 1.2. Basic molecular signs of centromeres and pericentric heterochromatin. CCAN – centromere associated protein network – is anchored on Cenp-A and Cenp-C rich chromatin at centromeres. HP1a compacts pericentric heterochromatin via H3K9me3.

1.5 Differences between centromeric and pericentromeric chromatins.

Centromeres and pericentromeric heterochromatin contain different types of chromatin, and (see above) different pathways contribute to the integrity of the domains.

Core centromeres contain "mixed" type of chromatin, in contrast to silent heterochromatin. In the pioneering study of human cells' and *Drosophila* mitotic spreads, centromeric chromatin turned out to be devoid of classical heterochromatic histone marks H3K9me2/3, of classical active marks (histones at centromeres were hypoacetylated), but contained the "poised" histone mark H3K4me2, typically found at promoters and transcribed genomic regions (Sullivan and Karpen, 2004). Following studies in chicken and human cells were sometimes controversial and difficult to interpret, but overall more active/"poised" than silent marks were detected. H4K5Ac, H4K12Ac, H3K36me2/3, H4K20me1, H3K4me1/2, H3K9me3, H3K27me1/2/3 were found at centromeres, although the first two marks only at pre-nucleosomal Cenp-A-H4-HJURP complex (Bailey et al., 2016; Bergmann et al., 2011; Hori et al.,

2014; Shang et al., 2016). Also H2AZ histone variant, a hallmark of active chromatin, was found at human centromeres (Greaves et al., 2006). Interestingly, in *Drosophila* cells proteins, which interact with CID, are often excluded from DAPI-dense regions. Moreover, there are no proteins that would interact with CID and localize exclusively to DAPI-dense regions or DAPI-dense regions and centromeres (Barth et al., 2014). HP1a pericentric chromatin, in contrast, contains classical heterochromatic histone marks H3K9me1/2/3 and H4K20me2/3, as well as less studied in context of heterochromatin H3K27me1 and H3K64me3 (reviewed in (Saksouk et al., 2015)). In addition, heterochromatic histones are hypoacetylated (reviewed in (Saksouk et al., 2015)). *Drosophila* HP1a has no centromeric interactors typical for the interphase core centromeres (Alekseyenko et al., 2014; Swenson et al., 2016).

Centromeres are transcribed: in *Drosophila* and human cells RNA PolII localises to centromeres, and transcription is necessary for Cenp-A incorporation (Bobkov et al., 2018; Chan et al., 2012; Quenet and Dalal, 2014). In *Xenopus*, centromeric transcripts are required for localization and activation of a kinetochore part, called chromosomal passenger complex (CPC) (Blower, 2004). Heterochromatin is transcribed only partially, with large repetitive regions being deliberately silenced, and RNA PolII on immunofluorescence of *Drosophila* cells is excluded from DAPI-dense regions (Bobkov et al., 2018).

1.6 Crosstalk and similarities between centromeric and pericentromeric chromatins.

In humans, mouse and *Drosophila*, centromeres and heterochromatin are assembled on different, although sometimes overlapping, DNA sequences. Centromeres need heterochromatin, but only at certain positioning and/or time point. Heterochromatin bordering is necessary for centromere formation: centromeres are embedded into heterochromatin and when overexpressed, *Drosophila* CID localises to heterochromatin boundaries (Olszak et al., 2011). Moreover, a part of heterochromatin (HP1a and HP1b proteins) is a part of centromeres in mitosis (Hayakawa et al., 2003; Minc et al., 1999). Accordingly, *Drosophila* and human HP1a have mitotic centromeric interactors: inner

centromeric protein Incenp and/or borealin (Ainsztein et al., 1998; Alekseyenko et al., 2014). In fission yeast, HP1a (Swi6) together with RNAi system that establishes heterochromatin is necessary for Cenp-A incorporation and kinetochore assembly (Folco et al., 2008). Functionally, both centromeres and heterochromatin are essential for proper mitosis (see below).

In *Drosophila*, speciation proteins HMR (hybrid male rescue) and LHR in cells and imaginal discs localise close to centromeres, and in cells HMR localises to a border between centromeres and heterochromatin (Thomae et al., 2013), (this study). This near-centromeric localisation in cells is dependent on Cenp-C (this study). However, in larval brains HMR and LHR localise to heterochromatin (Blum et al., 2017). Both proteins interact with HP1a (Satyaki et al., 2014; Thomae et al., 2013). Thus, the same proteins can localise mostly either to the centromeres or heterochromatin depending on the tissue or the cell cycle stage.

On the other hand, centromeres and heterochromatin contain different types of chromatin (see above) and spreading of H3K9me3 heterochromatin into the centromere is deleterious for centromere architecture and function (reviewed in (Ohzeki et al., 2016)). Human centromere contains acetyltransferase KAT7/HBO1/MYST2, which prevents it (Ohzeki et al., 2016). Reversibly, overexpression of Cenp-A leads to decreased H3K9me2 levels at heterochromatin (Lam et al., 2006).

1.7 Centromeric and pericentromeric chromatins are important in cell division.

Cell division (mitosis or meiosis) is an important stage of the cell cycle, when the chromosomes condense and the bulks of sister chromatids segregate to different poles of the dividing cell, followed by the cytokinesis (division of the cytoplasm). Mitosis consists of prophase, metaphase, anaphase and telophase.

Since the centromere is important for the kinetochore formation, and heterochromatin is important for the centromere architecture and the chromosome condensation, a fraction of centromeric, heterochromatic or bordering centromere (this study) proteins are required for proper cell division.

Changing levels of these proteins or their modifications can result in different types of mitotic defects, such as lagging and broken chromosomes, anaphase bridges, formation of micronuclei, or disorganised anaphases (Barth et al., 2014; Blum et al., 2017; Goutte-Gattat et al., 2013; Heun et al., 2006; Regnier et al., 2005; Thomae et al., 2013).

1.8 Centromeric and pericentromeric chromatins in speciation.

Speciation results from reproductive isolation of species. This reproductive isolation can be a result of prezygotic (for example, geographical) or postzygotic barriers. In case of the latter hybrids from sibling species cannot develop due to genetic incompatibility of maternal and paternal genomes. Several genes in flies and mice responsible for reproductive isolation have been described. One of the best studied cases of reproductive isolation are the crosses between *D. melanogaster* mothers flies and *Drosophila simulans* fathers flies*,* which result in infertile female and lethal male progeny. The progeny becomes viable and partially fertile with crosses after knockout of *Hmr* gene in *D. melanogaster* (Hutter and Ashburner, 1987), *Lhr* gene in *D. simulans* (Brideau et al., 2006), or GST-containing FLYWCH zinc-finger protein (*gfzf*) gene in *D. simulans* (Phadnis et al., 2015). The current model of this case of hybrid incompatibility comes from evidence, that *Hmr* and *Lhr* genes encode for two centromeric proteins, which expression level differs between two species. HMR is expressed higher in *D. melanogaster*, LHR – in *D. simulans*, and hybrids have elevated expression of both proteins. The overexpressed protein complex mislocalizes from centromeres, and gains additional binding sites all over the genome (Thomae et al., 2013).

The exact function of the HMR/LHR complex is still unknown, but the phenotypes observed in dying hybrid males resemble defects in cell cycle checkpoints and possibly chromosome condensation (Blum et al., 2017; Bolkan et al., 2007; Orr et al., 1997). The phenotypes observed in *Hmr* mutant flies are transposon deregulation and problems with sister chromatid cohesion (Blum et al., 2017).

Many gene products responsible for speciation bind to chromatin, and many of them localize to centromeres or pericentric heterochromatin. These include HMR, LHR and GFZF proteins (Barth et al., 2014; Thomae et al., 2013), responsible for hybrid lethality/sterility in the cross between *D. melanogaster* males and *D. simulans* females. Interestingly, the reverse cross between *D. melanogaster* and *D. simulans* depends on so-called *zhr* locus, which is basically a pericentric block of 359 repeats (Ferree and Barbash, 2009).

In crosses between *D. simulans* and *Drosophila mauritiana*, OdsH protein from *D. mauritiana* binds to *D. simulans* Y-chromosome and acts as a male sterilizing factor (Bayes and Malik, 2009). Since chromocenters are subject to rapid evolution (reviewed in (Sawamura, 2012)), most probably many other so far unknown speciation proteins bind to centromeres or pericentric heterochromatin. Thus, this evolutionarily dynamic intranuclear domain plays an important role in the formation of species.

1.9 Methods to study proteomic composition of nuclear domains.

To gain insights into the biology and function of chromatin domains, we need to study their proteomic composition. Immunoprecipitation (IP) coupled to mass spectrometry has been long the only method to do it. Cenp-A, CID and HP1a immunoprecipitation has gained significant insights into the proteomic composition of centromeric and pericentromeric heterochromatin in humans, mouse and/or *Drosophila* (Alekseyenko et al., 2014; Barth et al., 2014; Obuse et al., 2004; Swenson et al., 2016; Zaidan et al., 2018).

In recent years proximity labelling methods have appeared, which use *Escherichia coli* biotin ligase BirA (Branon et al., 2017; Roux et al., 2012), soybean ascorbate peroxidase APEX or its improved version APEX2 (Lam et al., 2015; Rhee et al., 2013) to biotinylate proteins around the bait in the cell. Proximity proteomics with BirA-dCas9 (Schmidtmann et al., 2016), APEX-dCas9 (Gao et al., 2018) and APEX fused to chromatin proteins (this study) has been successfully applied to capturing proteomes of neighbouring domains of

centromeric and pericentromeric chromatin, or domains of centromeric and telomeric chromatin.

Both approaches of proximity labelling and immunoprecipitation have positive sides and drawbacks, and both provide valuable information about the domain proteomic composition, although seek to answer different questions: affinity purification targets strong affinity interactions of the protein of interest, while proximity labelling captures both strong and weak/transient interactions, as well as proteins not interacting but residing in proximity to the bait (Fig. 1.3). Also, proximity labelling can reveal a proteomic snapshot of the compartments, difficult to purify with fractionation (for example, chromocenter).
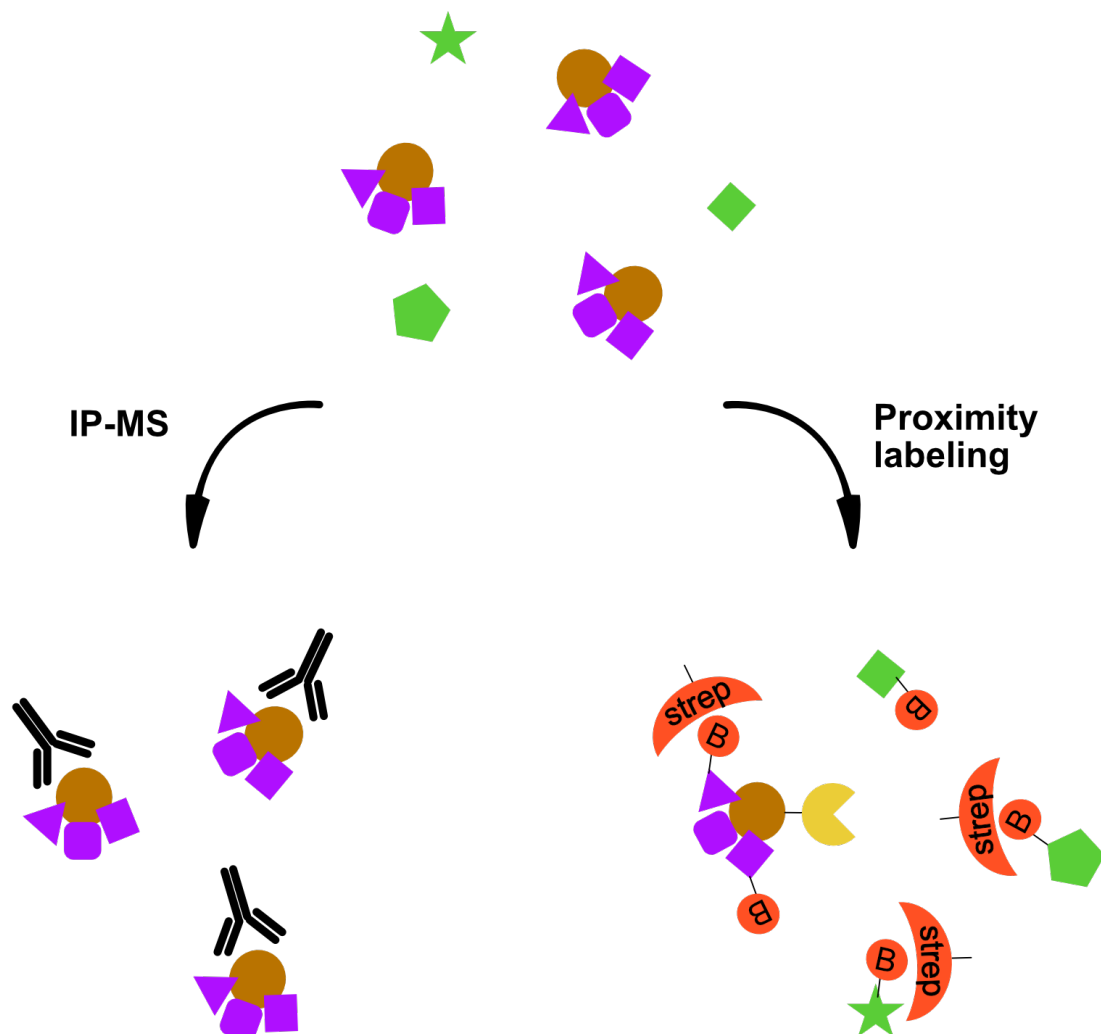
Fig. 1.3. Affinity IP-MS captures strong protein interactions, while pulldown after proximity biotinylation captures proteins in proximity to the bait. The bait is depicted in brown, interactors of the bait – in violet, proteins in proximity – in green, biotin tags (B) – in red.

1.10 Thesis aims.

Two main aims of the thesis were approached. First, we wanted to gain additional insights into the structure of the chromocenter taking advantage of modern state-of-art methods. The proteomic composition of the chromocenter in *Drosophila* has so far only been described by conventional IP-mass spectrometry (IP-MS) (Alekseyenko et al., 2014; Barth et al., 2014; Swenson et al., 2016; Thomae et al., 2013), and APEX2 proximity labeling, which was shown to be efficient in capturing proteomes of different organelles (see Chapter 2.2), could gain new insights into the biology of this intranuclear compartment. Apart from this, chromocenter has only been studied by low-resolution microscopy, and our STED findings together with (Anselm et al., 2018) show the intricate structure of the domain, unknown before.

Second, one of the centromeric factors investigated in this work, HMR, has a role in separating *D. melanogaster* and *D. simulans* species. Despite its well-known role in speciation, the function of HMR and the molecular links to the phenotypes in hybrids and pure species remained enigmatic. We aimed to use proximity labeling to gain insights into HMR biology, which could potentially point to its exact molecular function in speciation.

Together, the findings of this thesis should be of interest both for scientists interested in the chromocenter architecture and for scientists interested in the formation of species.

2 Results.

2.1 Confocal and STED microscopy reveal the complex architecture of the chromocenter.

As mentioned before, HMR is a speciation protein in crosses between close species of *Drosophila*. *D. melanogaster Hmr* mutants can be crossed to *D. simulans* and these crosses, otherwise producing lethal/unfertile progeny, result in a viable and fertile offspring (Brideau et al., 2006). HMR was reported to colocalize in cells and imaginal discs with the centromere-specific histone variant dCenpA (Thomae et al., 2013), with telomeres on polytene chromosomes (Thomae et al., 2013) and with heterochromatin in larval brains (Blum et al., 2017). To confirm HMR localization in cell culture, we used an HMR-FLAG CRISPR cell line (Gerland et al., 2017), which we costained with anti-FLAG, anti-HMR and anti-dCenpA antibodies.
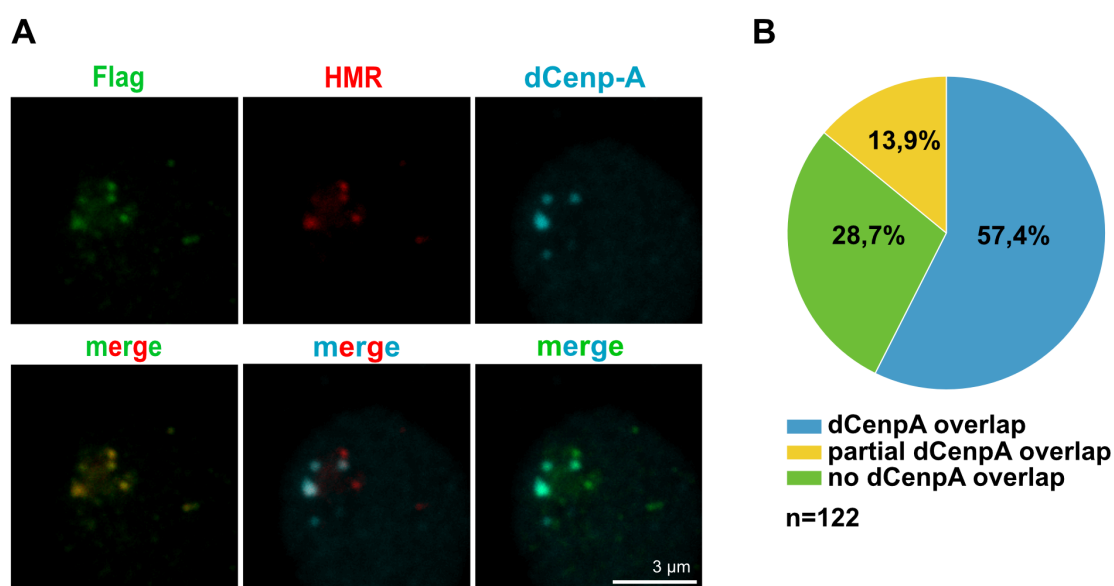
Fig. 2.1. A - Mouse anti-FLAG M2, rat anti-HMR 2C10 and rabbit anti-dCenpA (Actif Motif) staining of FLAG-CRISPRed HMR cell line (Gerland et al., 2017). B - Quantification of fractions of centromeres that overlap or do not overlap with HMR. For quantification dCenpA-GFP cell line (Heun et al., 2006), stained against HMR, was used. The data was acquired by Grusha Primal Mathias.

Consistent with what was observed before, we detected FLAG and HMR centromere staining (Fig. 2.1 A) Interestingly, among 122 centromeres quantified, HMR domains colocalized or partially colocalized only with a subset

of centromeres (71,3%) (Fig. 2.1 B). This suggests that either HMR localization to centromere is cell cycle regulated, or HMR specifically localizes to a subset of chromosomes.

In *Drosophila*, the area of the chromosome, which includes centromeres and nearby pericentric heterochromatin, is defined as the chromocenter. Several proteins, known to play a role in the chromocenter architecture, such as NLP and D1, upon depletion demonstrate the phenotypes of deregulation of transposable elements, defects in mitosis, and formation of macronuclei (Jagannathan et al., 2018; Padeken et al., 2013). Mutation or knockdowns of HMR show similar phenotypes, suggesting that HMR might play a role in the architecture of the chromocenter. To investigate the structure of the chromocenter and HMR's role in it, we performed high-resolution STED microscopy using anti-HMR, anti-dCenpA and anti-dCenpC antibodies.
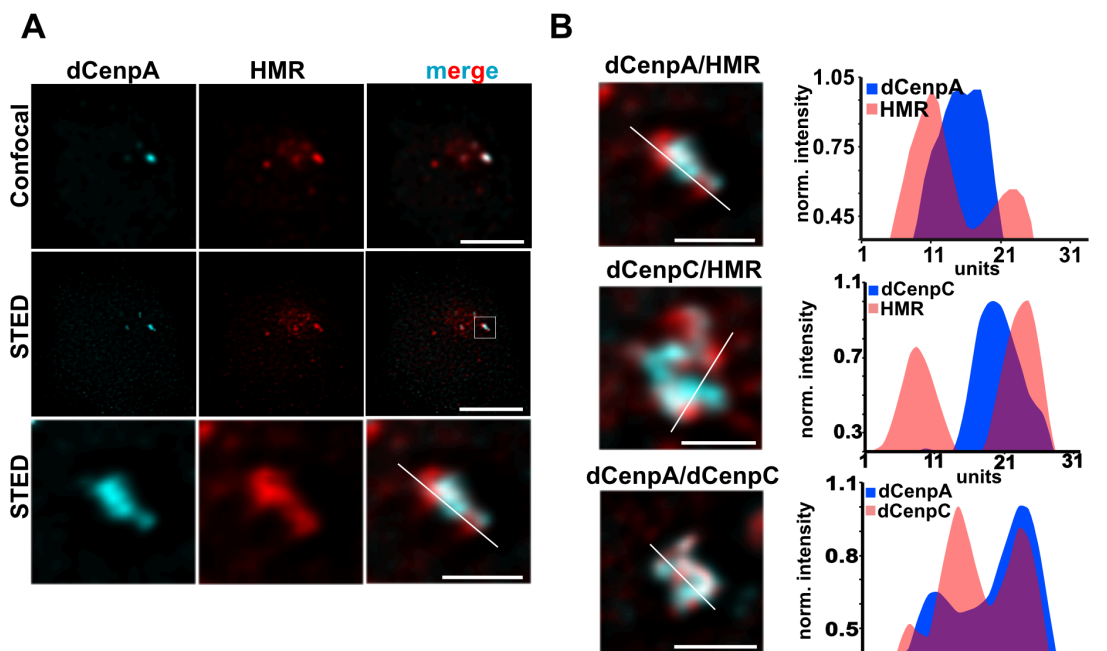
Fig. 2.2. A - Confocal and STED high-resolution microscopy of dCenpA and HMR. Scale bars represent 3 uM and 0.8 uM for zooms. B - High-resolution STED microscopy co-stainings of dCenpA/HMR, dCenpC/HMR and dCenpA/dCenpC. The antibodies used are rat anti-HMR 2C10, rabbit anti-dCenpC, rabbit anti-dCenpA (Actif Motif) and rat anti-dCenpA 7A2. Scale bars represent 0.8 uM. The intensity profiles were built in ImageJ and normalized to one of the maximum peaks. Data from panel B was partially produced and recorded by Dr. Andreas W. Thomae. All stainings were recorded by Dr. Andreas W. Thomae.

According to previous observations (Thomae et al., 2013), HMR colocalized with dCenpA on the images taken using confocal microscopy (Fig. 2.2 A). However, STED microscopy revealed interdigitating domains of dCenpA

and HMR, separated only in their highest intensities (Fig. 2.2 A and B). The same interdigitation was observed for dCenpA/dCenpC and dCenpC/HMR protein pairs (Fig. 2.2 B). Furthermore, we plotted the distributions of Spearman correlations, each correlation between two channels on the same image. Those distributions revealed that dCenpC/dCenpA correlate more, than HMR/dCenpA and HMR/dCenpC (Fig. 2.3).
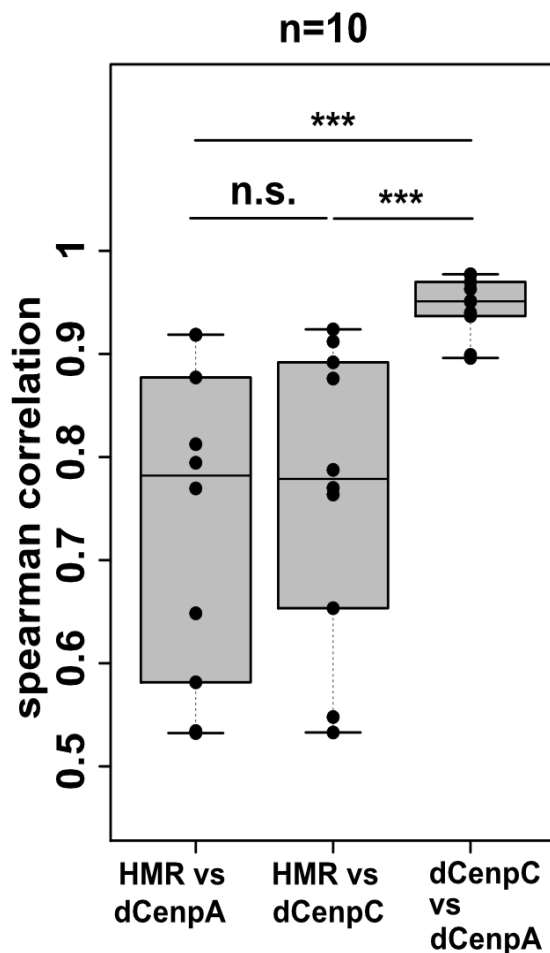
Fig. 2.3. Distributions of Spearman correlations between pairwisely recorded images of mentioned proteins (from 2 independent experiments). Wilcoxon rank sum test was used for statistical analysis. N.s. – non-significant, *** - p-value<0.001.

These experiments revealed that a chromocenter is a much more complex structure than thought before. We suggest that the chromocentric chromatin consists of a meshwork of interdigitating HMR, dCenpA and dCenpC domains. HMR domains are separated from dCenpA and dCenpC in their highest intensities, while dCenpA and dCenpC show some co-localization.
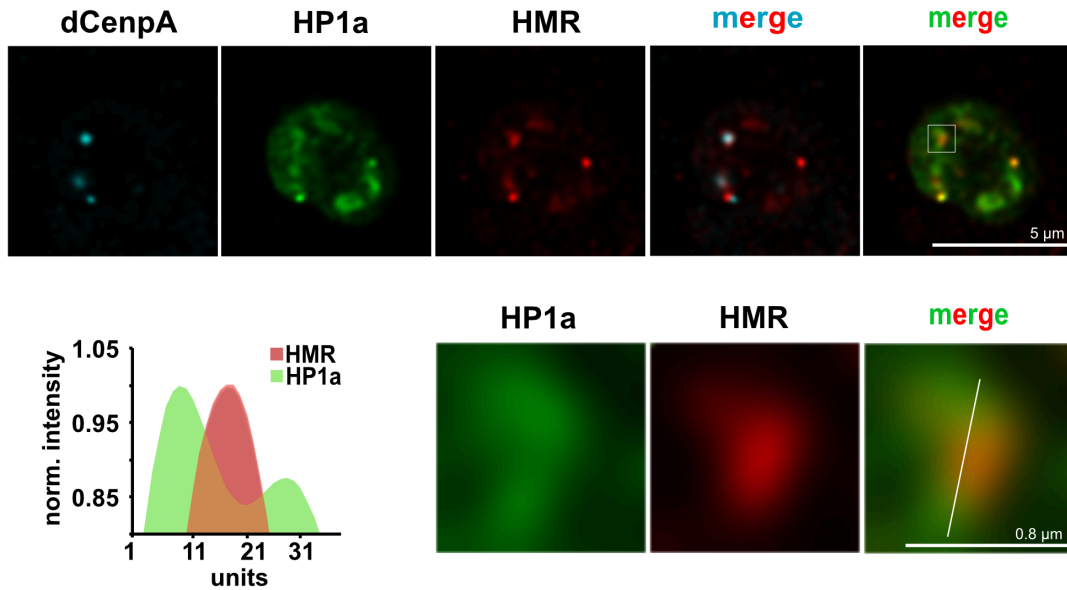
Fig. 2.4. Confocal microscopy of dCenpA, HMR and HP1a. The antibodies used are rat anti-HMR 2C10, mouse anti-HP1a C1A9 and rabbit anti-dCenpA (Actif Motif). The intensity profile was built in ImageJ and normalized to one of the maximum peaks. Data were recorded by Dr. Andreas W. Thomae.

Fig. 2.5. Distribution of Spearman correlations between confocal images (from 2 independent experiments).
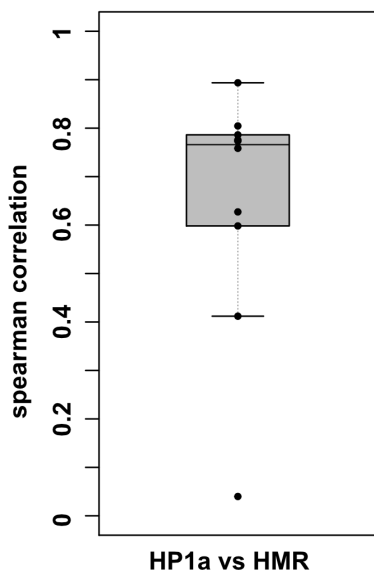
Consistent with what was reported before, the centromeric region is embedded in HP1a chromatin (Fig. 2.4). The wide distribution of Spearman correlations between HMR and HP1a suggests little overlap of those regions (Fig. 2.5). Interestingly, using confocal microscopy of HP1a, as well as high-resolution microscopy of dCenpA and HMR, we observed that HMR chromatin often borders dCenpA chromatin from HP1a chromatin (Fig. 2.6).

Fig. 2.6. A screenshot from the movie, which combines 3D models of HMR and dCenpA staining (STED microscopy) and HP1a staining (confocal microscopy). Data were recorded and the movie was processed by Dr. Andreas W. Thomae.

2.2 Establishing cell lines for proximity labeling.

To further investigate the structure of the chromocenter, we fused dCenpA, HMR and HP1a to the ascorbate peroxidase from soybean (APEX2 (Lam et al., 2015; Rhee et al., 2013)) (Fig. 2.7) and expressed proteins under inducible metallothionine promoter. As a control, we fused APEX2 to two nuclear localization signals (NLS).

Fig. 2.7. Schemes of the cloned constructs. The list of primers for cloning is available in Table 2.

kDa

170
130
100
70
55 — ◄ **GST-APEX**
40
35
25
15
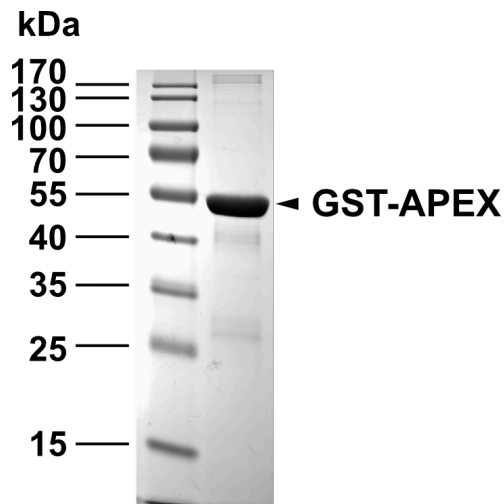
Fig. 2.8. The coomassie staining of the purified recombinant GST-APEX protein.

Upon half an hour treatment with 0.5 mM biotin-phenol and subsequent 1-2 minutes treatment with hydrogen peroxide, APEX2 produces biotin-phenoxyl radicals, which fuse to tyrosines and other electron rich aminoacids on the surfaces of the nearby proteins in a radius which has been estimated to be less than 20 nm (reviewed in (Bendayan, 2001)). Both APEX and APEX2 have been successfully used for different proximity labeling approaches in the last years, including capturing the proteomes of mitochondria (Hung et al., 2014; Rhee et al., 2013), stress granules (Markmiller et al., 2018), lipid droplets (Bersuker et al., 2018), cilia (Mick et al., 2015), ribosomal proteins (Zuzow et al., 2018), proteins involved in DNA damage response (Gupta et al., 2018), as well as separate chromatin loci including centromeres and telomeres in human cells, using APEX2 fused to nuclease deficient Cas9 (Gao et al., 2018; Schmidtmann et al., 2016). To confirm the correct expression of the fusions, we did Western blotting of whole cell extracts as well as the immunostaining of the stable cell lines. We raised an anti-APEX antibody against the recombinant GST-APEX protein (Fig. 2.8). Antibody generation was performed by injection of recombinant proteins into Wistar rats, and fusing their splenic B-cells with myeloma cell line. More than 50 antibody clones were generated, and clone 20H10 was chosen as the one exhibiting highest sensitivity.

As seen from western blotting (Fig. 2.9), induced HP1a$_{AP}$, uninduced HMR$_{AP}$ and uninduced dCenpA$_{AP}$, which we chose for the subsequent biotinylation, are expressed approximately at the levels of the endogenous proteins or lower. Moreover, uninduced HMR$_{AP}$ and dCenpA$_{AP}$ localize mostly to centromeres, while induced proteins localize to other parts of the nucleus and only optionally to the centromeres (Fig. 2.10). Induced HP1a$_{AP}$ occupies a domain in the nucleus, which co-stains with endogenous HP1a. APEX$_{NLS}$ localizes to the nucleus, defined by the DAPI staining (Fig. 2.11). We decided

to induce APEX$_{NLS}$ for the stronger biotinylation and therefore the stronger stringency of the control.
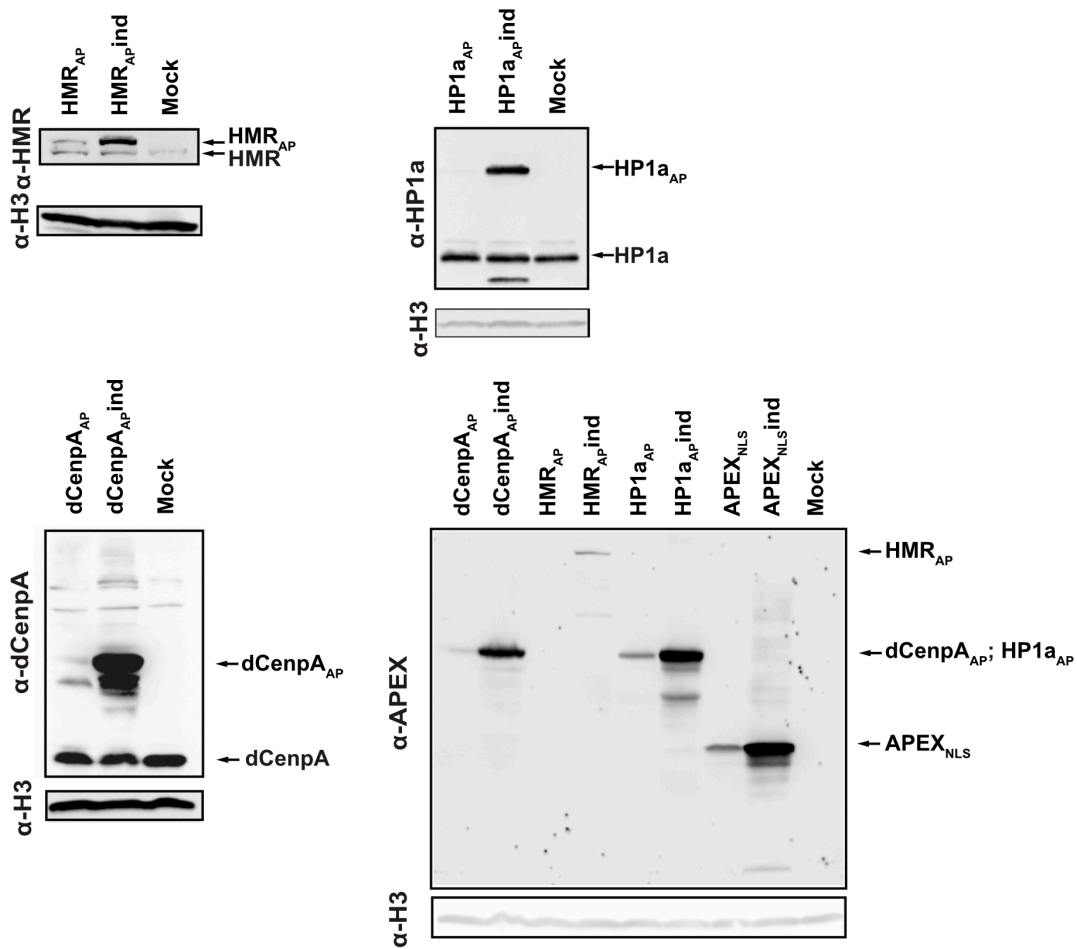
Fig. 2.9. Western blotting of whole cell lysates of respective stable cell lines. The antibodies used are rat anti-HMR 2C10, mouse anti-HP1a C1A9, rabbit anti-dCenpA (Actif Motif) and rat anti-APEX 20H10.

Importantly, dCenpA$_{AP}$, HP1a$_{AP}$ (Fig. 2.10) and HMR$_{AP}$ (Fig. 2.12) at chosen biotinylation conditions exhibit a staining pattern similar to endogenous proteins. This suggests that centromeric and heterochromatic domains retain their structures even with increased dosages of respective proteins.

DAPI **APEX** **dCenpA** **MERGE**

L2-4 untransfected

DAPI **APEX** **dCenpA** **MERGE**

uninduced

induced

dCenpAAP

DAPI **APEX** **dCenpA** **MERGE**

uninduced
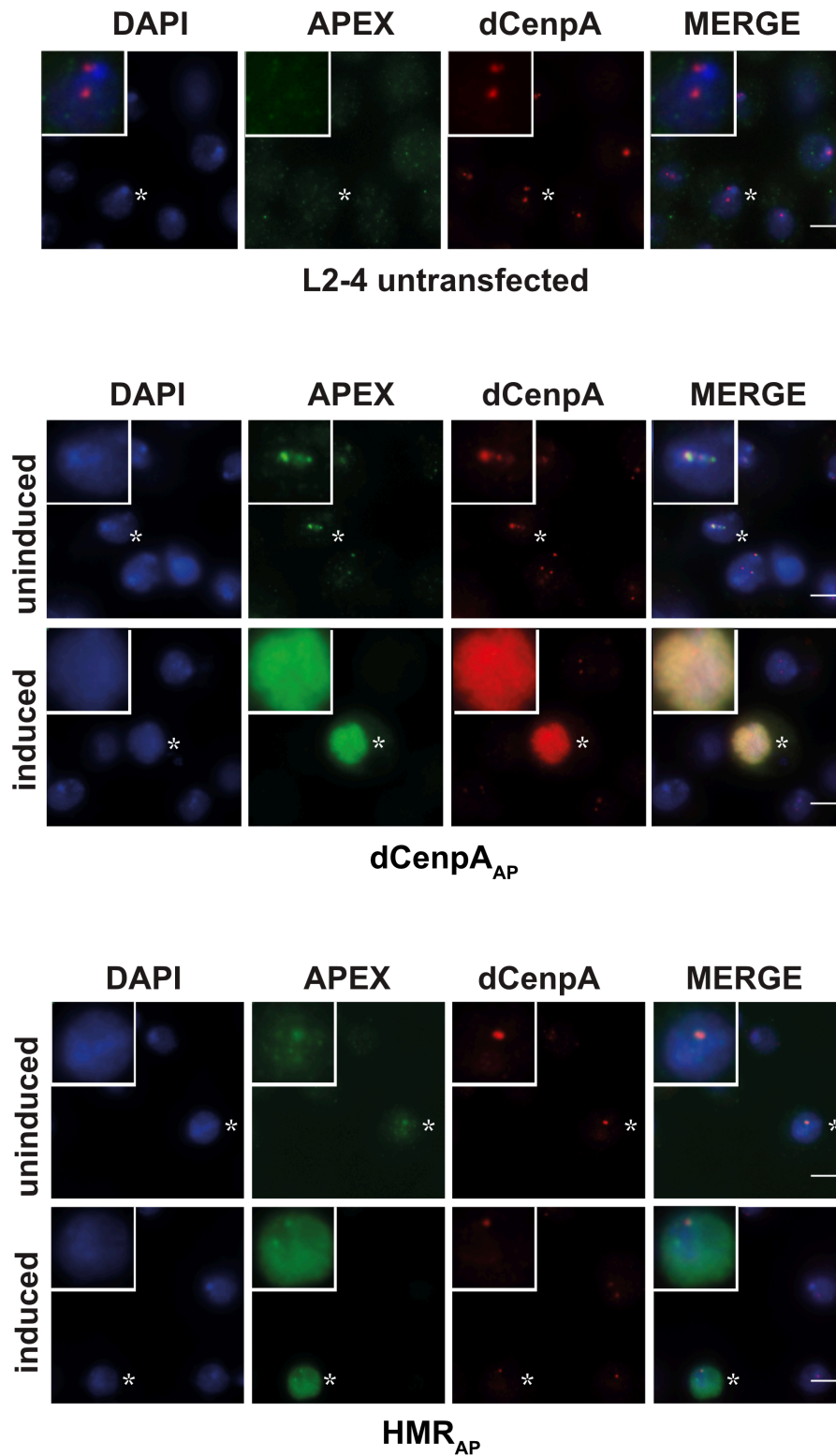
induced

HMRAP

Fig. 2.10. Immunofluorescence of uninduced and induced dCenpA$_{AP}$ and HMR$_{AP}$ cell lines used for the work. The antibodies used are rabbit anti-dCenpA (Actif Motif) and rat anti-APEX 20H10. A chosen cell (marked by an asterisk) is displayed in the inlet with approximately 2.3-fold zoom. Scale bars represent 5 μm.
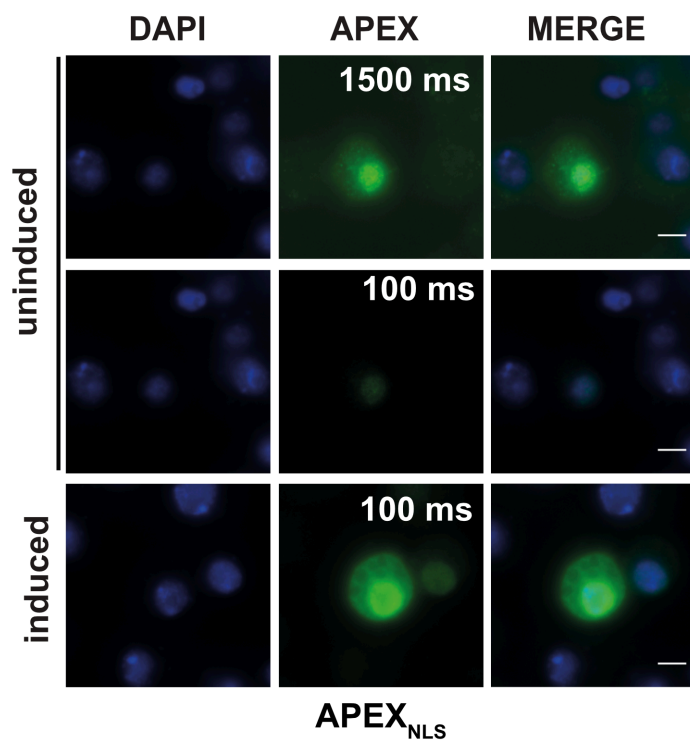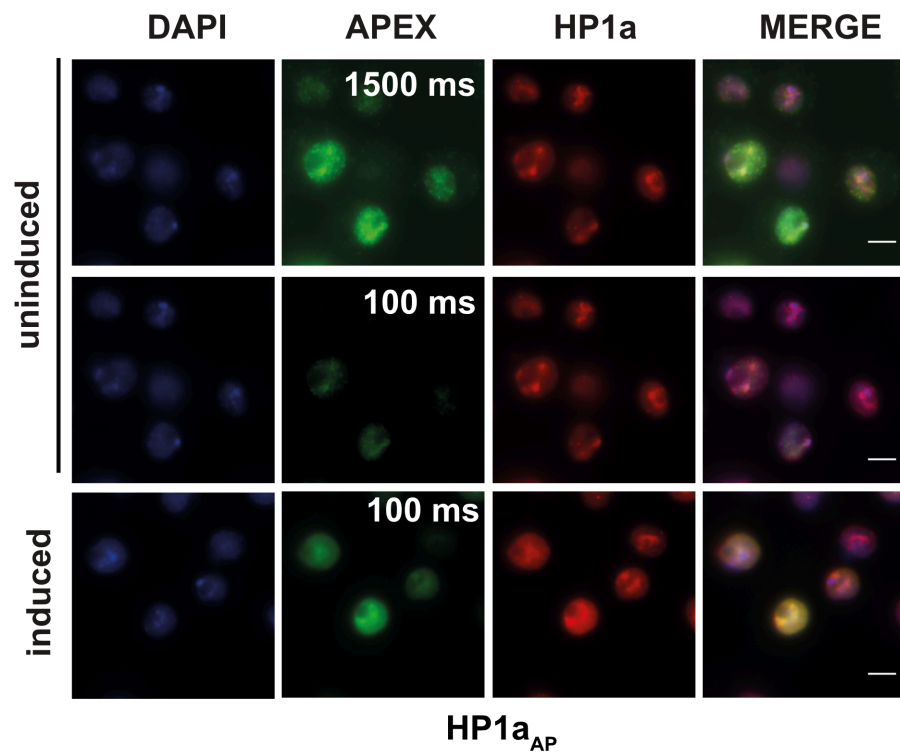
Fig. 2.11. Immunofluorescence of uninduced and induced HP1a$_{AP}$ and APEX$_{NLS}$ cell lines used for the work. The antibodies used are mouse anti-HP1a C1A9 and rat anti-APEX 20H10. The exposure is indicated in white. Scale bars represent 5 μm.
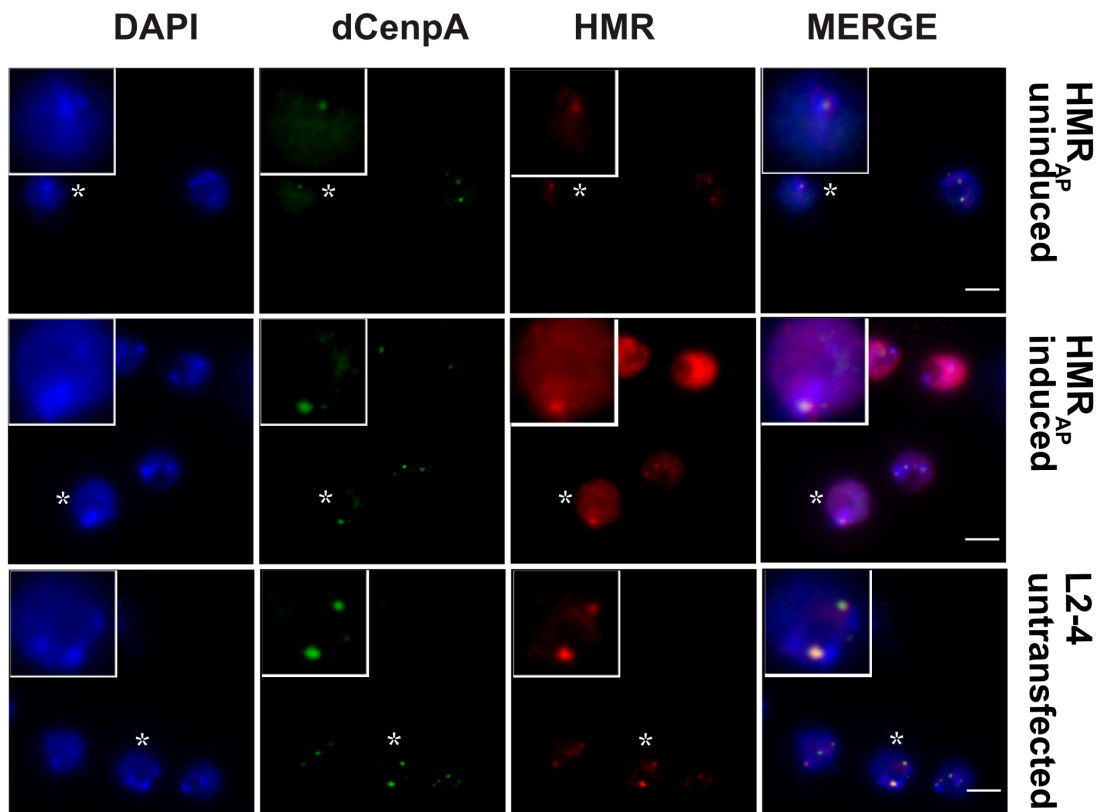
Fig. 2.12. Immunofluorescence of the uninduced and induced HMR$_{AP}$ cell line used for the work. Rabbit anti-dCenpA (Actif Motif) and rat anti-HMR 2C10 antibodies were used for staining. A chosen cell (marked by an asterisk) is displayed in the inlet with approximately 2.3-fold zoom. Scale bars represent 5 μm.

It is noteworthy to mention, that dCenpA$_{AP}$ and HP1a$_{AP}$ cell lines, which were initially established, were diluted to a density of 10^3 cells/ml in 20% conditioned medium, and colonies originating from several cells were selected, grown separately and further screened for the proper localization of the APEX fusion by immunofluorescence. Colonies, which showed the lowest fractions of hugely overexpressing cells (colony 8 for dCenpA$_{AP}$ and colony 29 for HP1a$_{AP}$), were chosen.

2.3 Establishing the conditions for proximity labeling.

2.3.1 1-minute biotinylation reveals focused biotin signal upon different biotin-phenol concentrations for different proteins.

The default conditions used in the first proximity labeling experiments included half an hour incubation with 0.5 mM biotin-phenol and treatment with 1 mM hydrogen peroxide for 1 minute (Rhee et al., 2013) (Fig. 2.13). Using these default conditions, we observed a focused biotin signal after in situ reaction for HP1a$_{AP}$ (at heterochromatin) and APEX$_{NLS}$ (in the nucleus) (Fig. 2.14). HMR$_{AP}$ biotinylation at centromeres was however very weak. dCenpA$_{AP}$ biotinylation at centromeres was not detectable. We thus used 5 mM biotin-phenol treatment for HMR$_{AP}$ and dCenpA$_{AP}$ (Fig. 2.15).
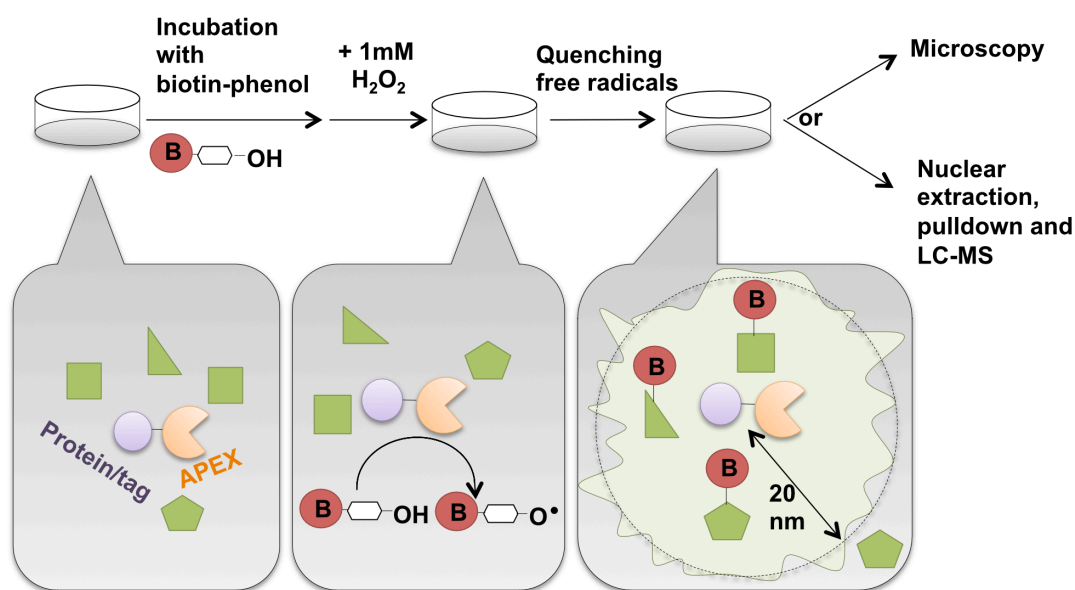
Fig. 2.13. The scheme of the biotinylation experiment with subsequent application to microscopy or proteomics.

2.3.2 Adjusting biotinylation time for different proteins.

We next decided to prolong the labeling to 25 minutes to a) maximize the biotinylation efficiency, b) for the simplicity of the experimental setup, since
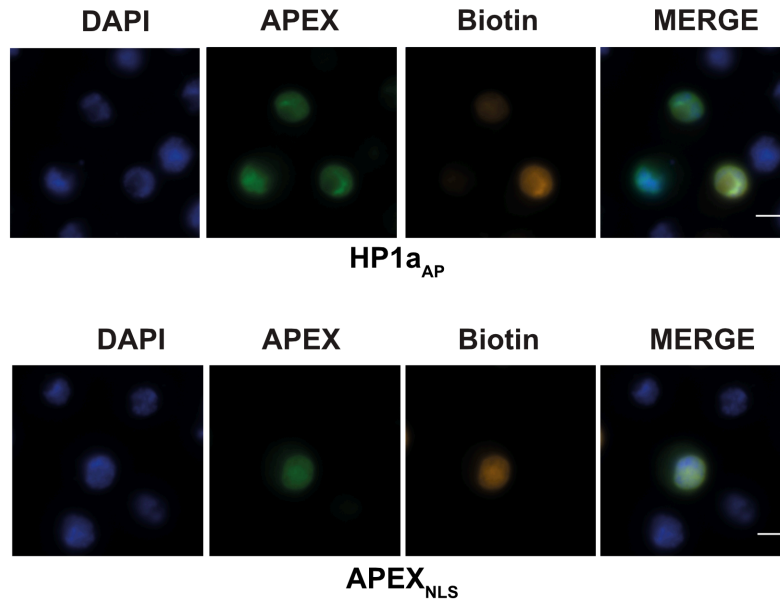
Fig. 2.14. In situ biotinylation reaction for HP1a$_{AP}$ and APEX$_{NLS}$ (with 0.5 mM biotin-phenol). The antibodies used for staining are rat anti-APEX 20H10 and Streptavidin-Alexa-555 (Thermo Fisher Scientific). Scale bars represent 5 µm. The experiment was done in parallel with experiment in Fig. 2.15.

Drosophila cells are not very adherent but could be spun down at large volumes during 20 minutes. Upon 25 minutes' labeling, the centromeric signal in HMR$_{AP}$ and dCenpA$_{AP}$ cells was still focused, while HP1a$_{AP}$ cells showed less constrained signal all over the nucleus (Fig. 2.16, Fig. 2.17). We thus did a time course of HP1a$_{AP}$ labeling and found out that the biotin signal at heterochromatin is focused until 5 minutes of biotinylation but becomes diffuse at longer labeling times (Fig. 2.18). During this time course 1 minute and 5 minutes biotinylation was performed on coverslips, while 10 and more minutes biotinylation was performed in solution.
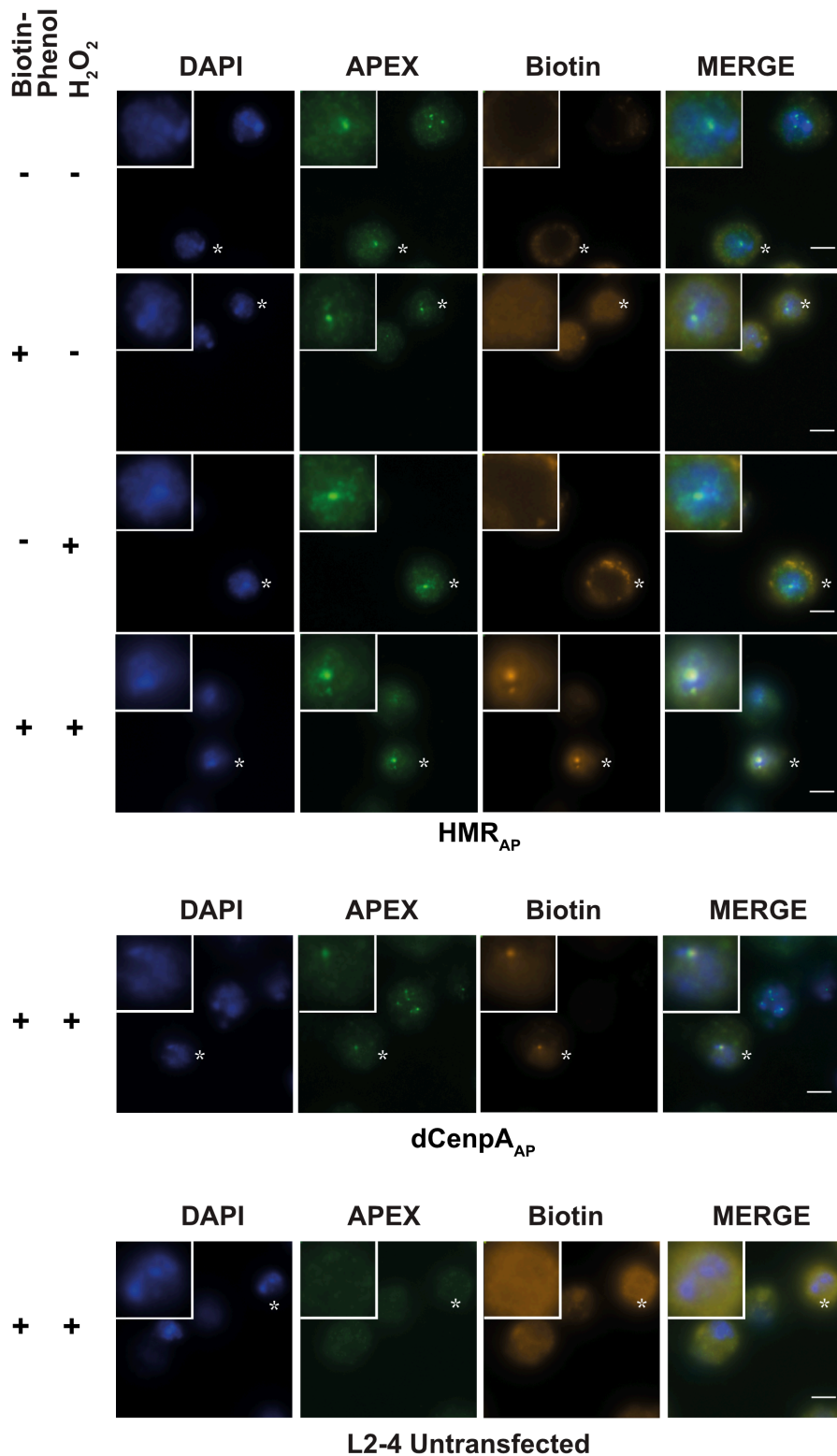
Fig. 2.15. In situ biotinylation reaction for HMR_{AP} and $dCenpA_{AP}$ (with 5 mM biotin-phenol). The antibodies used for staining are rat anti-APEX 20H10 and Streptavidin-Alexa-555 (Thermo Fisher Scientific). A chosen cell (marked by an asterisk) is displayed in the inlet with approximately 2.3-fold zoom. Scale bars represent 5 μm. The experiment was done in parallel with experiment in Fig. 2.14.
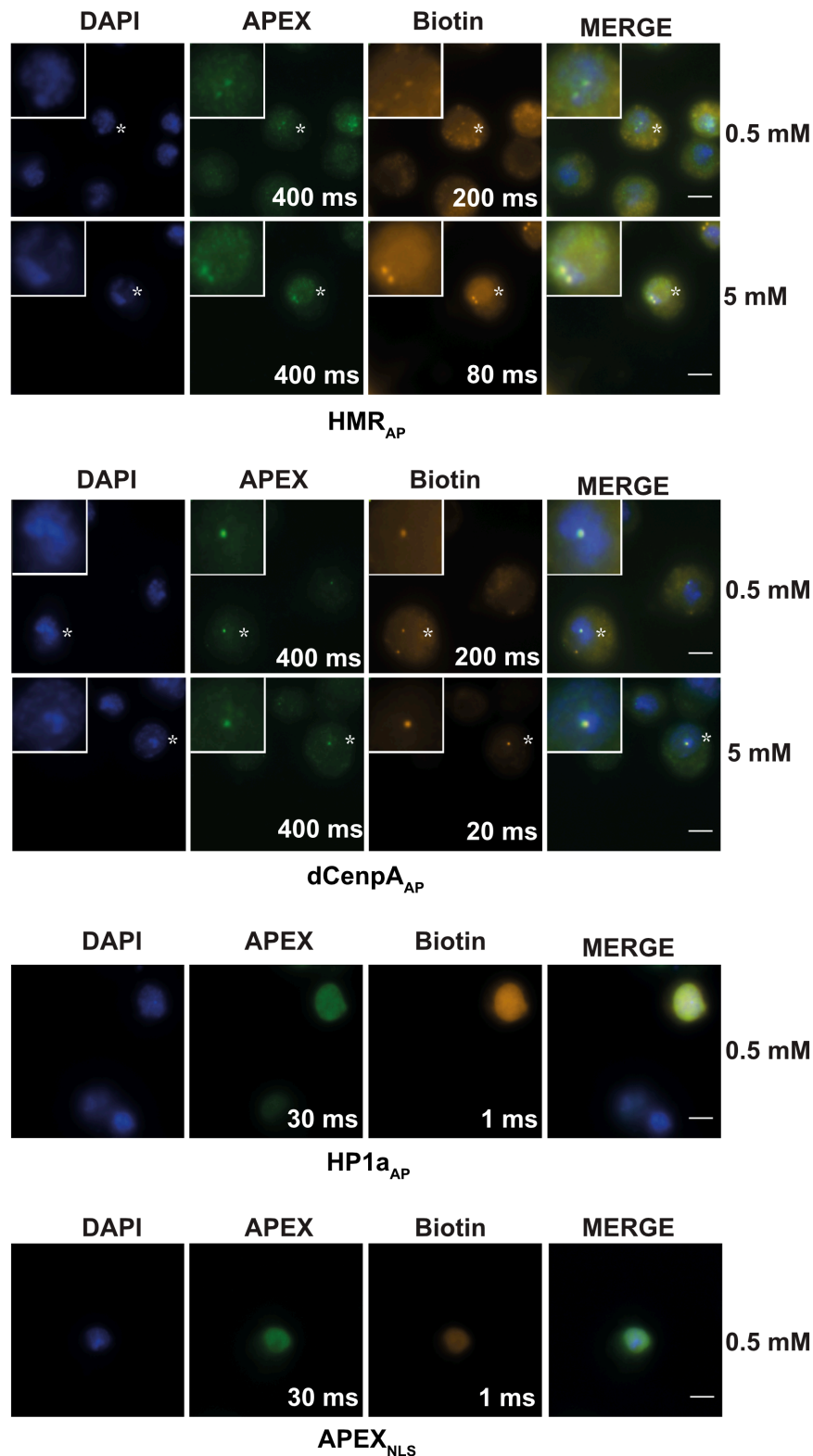
Fig. 2.16. 25 minutes biotinylation reaction for HMR$_{AP}$, dCenpA$_{AP}$, HP1a$_{AP}$ and APEX$_{NLS}$. The antibodies used for staining are rat anti-APEX 20H10 and Streptavidin-Alexa-555 (Thermo Fisher Scientific). In cells with the centromeric staining a chosen cell (marked by an asterisk) is displayed in the inlet with approximately 2.3-fold zoom. For APEX and biotin the exposure is indicated in white. Scale bars represent 5 µm. The experiment was done in parallel with experiment in Fig. 2.17.
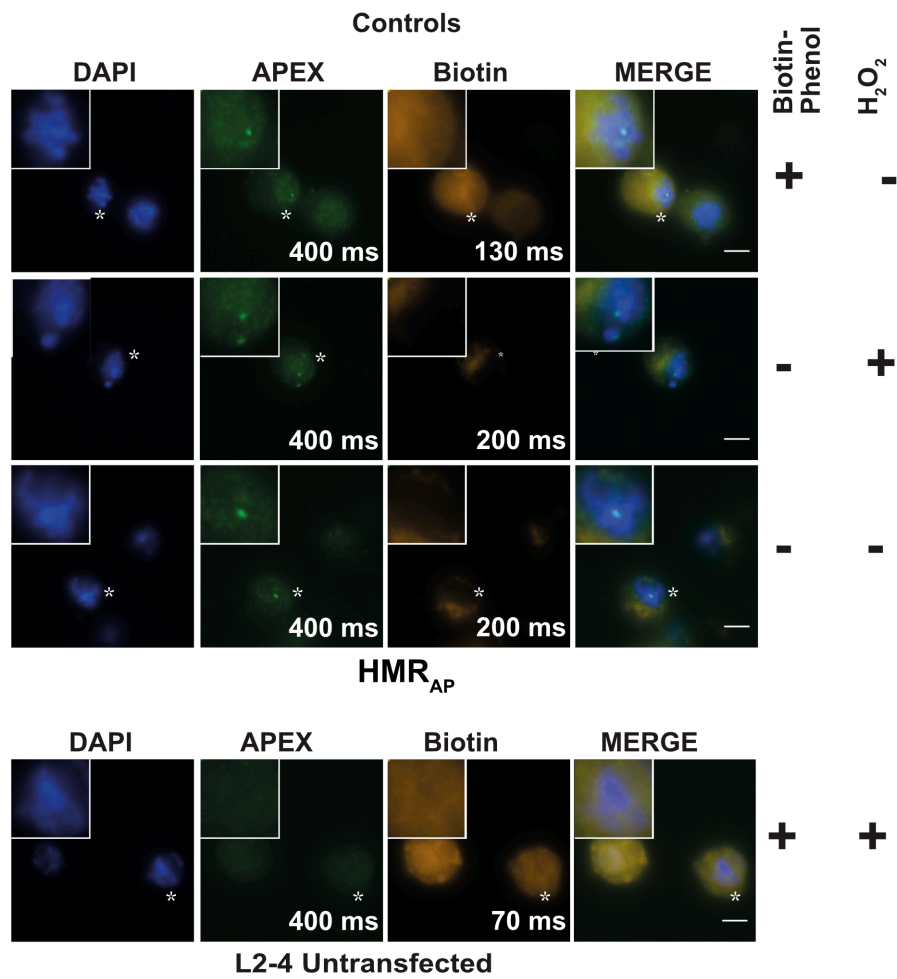
Fig. 2.17. Controls for 25 minutes biotinylation reaction. The antibodies used for staining are rat anti-APEX 20H10 and Streptavidin-Alexa-555 (Thermo Fisher Scientific). A chosen cell (marked by an asterisk) is displayed in the inlet with approximately 2.3-fold zoom. For APEX and biotin the exposure is indicated in white. Scale bars represent 5 μm. The experiment was done in parallel with experiment in Fig. 2.16.

For further proteomic experiments we decided to apply: 1) 25 minutes biotinylation for HMR_{AP} and $dCenpA_{AP}$; 2) 1.5, 5 and 25 minutes biotinylation for $HP1a_{AP}$, as well as 3) 1.5, 5 and 25 minutes biotinylation for $APEX_{NLS}$ as a control.
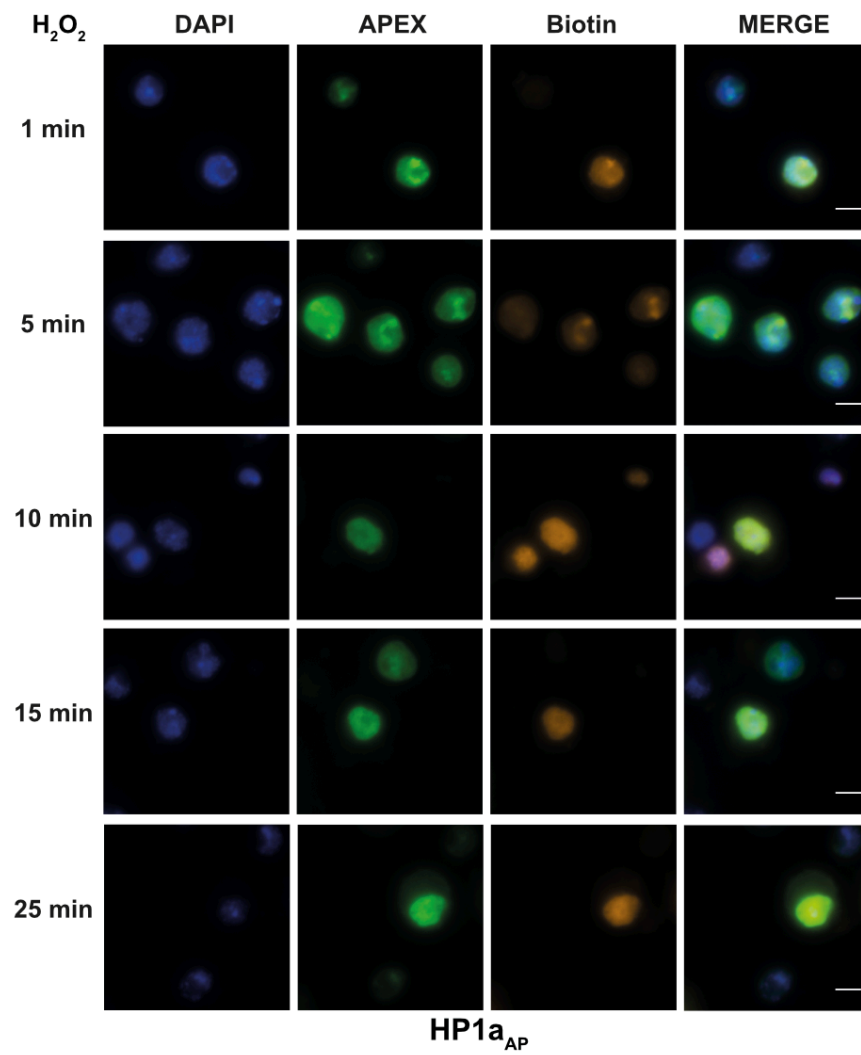
Fig. 2.18. Time course biotinylation reaction for HP1a_AP. The antibodies used for staining are rat anti-APEX 20H10 and Streptavidin-Alexa-555 (Thermo Fisher Scientific). Scale bars represent 5 µM.

2.4 Proximity labeling coupled to proteomics confirms the intricate structure of the chromocenter.

2.4.1 Biotinylation and purification of biotinylated proteins.

We aimed to establish a strategy, with which proximity labeling could be easily applied to *Drosophila* L2-4 different transgenic cell lines. Since *Drosophila* cells are not adherent and since large amounts of cells are needed for nuclear extraction, we grew cells in roller bottles to the density of approximately $5*10^6$/milliliter. Next, cells were counted, and 1 billion cells was used for biotinylation in solution, when cells were first incubated with biotin-phenol, and then hydrogen peroxide was added and cells were spun down during biotinylation procedure. Alternatively, for shorter biotinylation cells couldn't be spun down. Thus, 2 bottles of cells (further used per 1 sample) were adhered on 20 15-cm plates usually used for mammalian cells, and the treatment was performed as described previously for mammalian cells (Rhee et al., 2013), with the difference that cells were washed only once in quenching solution and scraped off.

For nuclear extraction, the nuclei were separated from the cytoplasm, and nuclear architecture was disrupted using 3 methods: MNase/benzonase digestion with large amounts of enzymes to disrupt biotinylated chromatin, douncing using tight-fitting pestle and addition of large amounts of salt and detergents for chromatin extraction. The final salt concentration used (600 mM) was reported to be enough to get 90% of nucleosomes into solution (Henikoff et al., 2009). Detergents were applied since protein-protein interactions did not need to be preserved, and since detergents were reported to be efficient in solubilizing large protein complexes (Henikoff et al., 2009).

Since biotin-streptavidin interaction is very strong (the dissociation constant was reported to be approximately 10^{-14} mol/L (reviewed in (Green, 1975))), after anti-biotin immunoprecipitation (IP) the beads were washed not only with IP buffer, but also 3 times with 4M urea, so that protein-protein interactions, but not biotin-streptavidin interactions, would be disrupted and only biotinylated proteins, but not their interactors, would remain on streptavidin-coated beads (Freire et al., 2013; Kurzban et al., 1991). The

washes were followed by on beads digestion, and the samples were measured in minimum 2 technical replicates on the Q-Exactive mass spectrometer (Thermo Fisher Scientific). The raw data as well as the MaxQuant output .txt files were deposited on ProteomeXchange (http://proteomecentral.proteomexchange.org/dataset/PXD012551).

2.4.2 GO-term analysis of proximities of HMR$_{AP}$ and HP1a$_{AP}$ reveals known categories.

We performed a 25 minute dCenpA$_{AP}$, HMR$_{AP}$, HP1a$_{AP}$ and APEX$_{NLS}$ biotinylation, and defined proteins enriched in all four pulldowns compared to pulldowns from DMSO treated controls (Fig. 2.19). 325, 314, 259 and 273 proteins were biotinylated respectively in dCenpA$_{AP}$, HMR$_{AP}$, HP1a$_{AP}$ and APEX$_{NLS}$ samples. When defining an overlap between three proximity proteomes with APEX$_{NLS}$ proximity proteome, we found dCenpA$_{AP}$ and HMR$_{AP}$ proteomes to be more specific (only 49.9% and 36.6% overlap with APEX$_{NLS}$ proximity proteome), than HP1a$_{AP}$ proximity proteome (72.2% overlap) (Fig. 2.20). This might be the consequence of proteins' biology, since HP1a was reported to be very mobile (reviewed in (Straub, 2003)) and its fusion with APEX could potentially biotinylate a larger fraction of the nucleus. Because 25 minutes' HP1a$_{AP}$ biotinylation was delocalized, we also preformed HP1a$_{AP}$ biotinylation for 1.5 and 5 minutes.

We further defined the proteins enriched in dCenpA$_{AP}$, HMR$_{AP}$ and HP1a$_{AP}$, but not or to a lesser degree in APEX$_{NLS}$ (Fig. 2.21). From these proteins we filtered only nuclear ones, using a Gene Consortium tool (http://www.geneontology.org). With the same tool we further performed the GO-term analysis of all three proximity proteomes, and filtered only specific GO-terms found for each of three pulldowns. For all times points of HP1a$_{AP}$ pulldowns we found such GO-terms as "chromatin organization" and "RNA and/or DNA metabolic processes", which is consistent with previously reported HP1a functions and known aspects of heterochromatin biology. For dCenpA$_{AP}$ pulldowns we found such GO-terms as "mitotic sister chromatid segregation"

and also terms connected to transcription by RNA-polymerase II. This is consistent with known dCenpA function in centromeric biology, as well as with the findings of transcription at the centromeres (Bobkov et al., 2018; Rosic et al., 2014). For HMR$_{AP}$ pulldowns we found GO-terms "cell cycle checkpoint", "mitotic sister chromatid segregation" and both negative and positive regulation of transcription by RNA polymerase II. This is consistent with previously reported phenotypes upon HMR knockdown/knockout in cells/flies (Blum et al., 2017; Bolkan et al., 2007; Satyaki et al., 2014; Thomae et al., 2013).
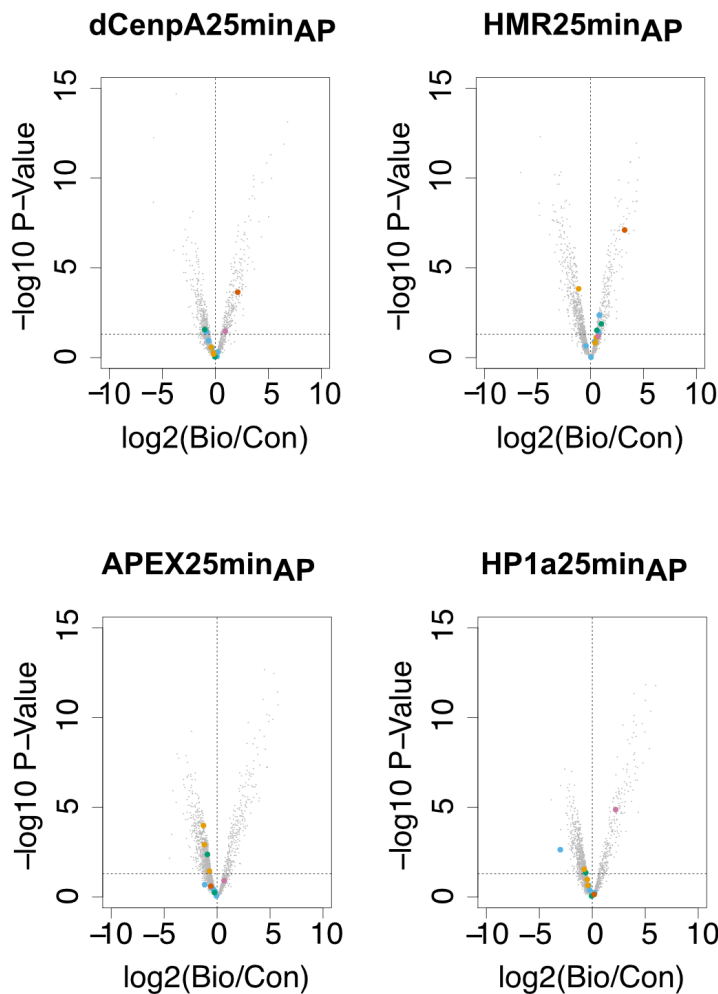
Fig. 2.19. Volcano plots of pulldowns from biotinylated cells vs pulldowns of DMSO-treated cells.

44

Interestingly, specifically for HMR$_{AP}$ pulldowns we found both GO-terms "regulation of histone methylation" and "histone acetylation". Moreover, in proximity to HMR$_{AP}$ we found both previously reported HP1a-interacting and dCenpA-interacting proteins. This strengthens the hypotheses that HMR often borders dCenpA-containing domains from HP1a-containing.
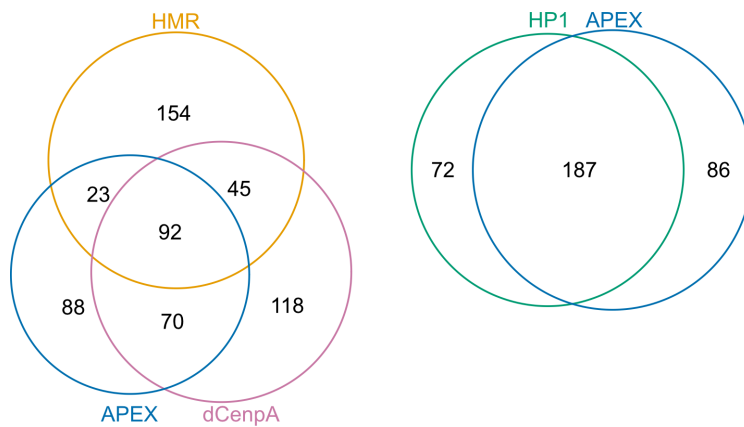
Fig. 2.20. Venn diagrams of proximity proteomes of HMR$_{AP}$, dCenpA$_{AP}$ and APEX$_{NLS}$, or HP1a$_{AP}$ and APEX$_{NLS}$.
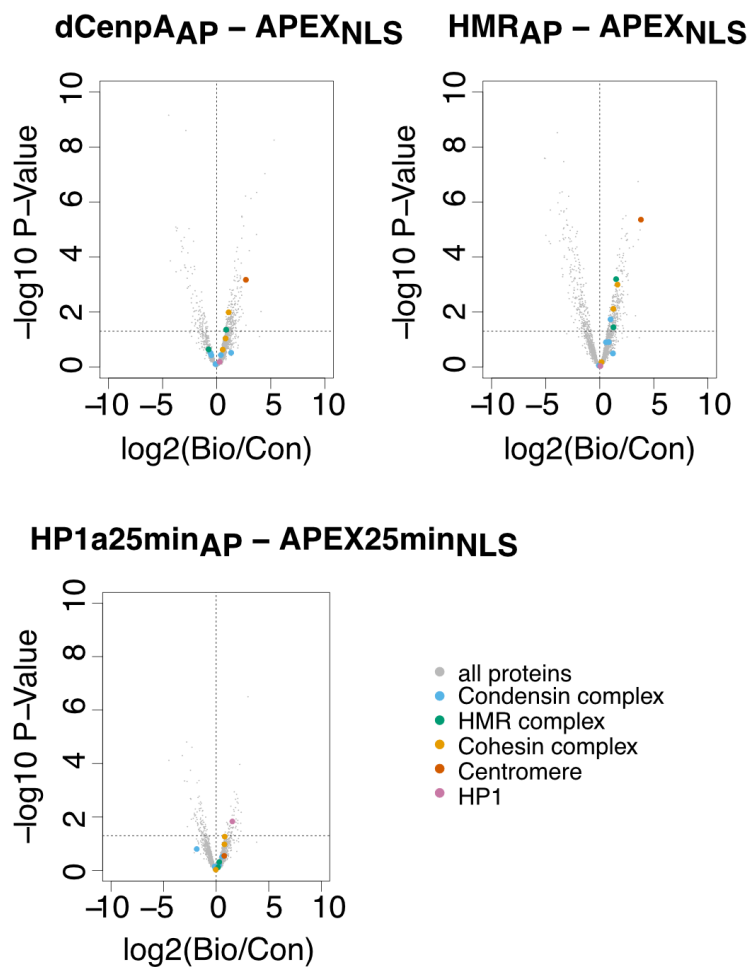
Fig. 2.21. Volcano plots of proteins enriched from biotinylated pulldowns vs APEX$_{NLS}$ pulldowns.

2.4.3 dCenpA_{AP} and HMR_{AP} proteomes reveal overlapping as well as different clusters.

For further analysis we decided to display three nuclear proximity proteomes as network graphs. Analysis of specific HP1a_{AP} proteome using the STRING database did not reveal many defined protein clusters, which might be due to a low number of proteins enriched in specifically HP1a_{AP} pulldowns. However, these clusters were visible in the STRING network graphs of dCenpA_{AP} and HMR_{AP} proteomes, which were built in Cytoscape using both STRING and manual Flybase connections (dotted lines) (Fig. 2.22).

Interestingly, we found that some clusters in the STRING networks were similar for dCenpA_{AP} and HMR_{AP}, while some were different. In particular, in both networks there were protein clusters corresponding to proteins involved in transcription, replication, components of the centromere, nucleolus and nuclear pore complex. Nucleosome remodelers, boundary factors, Polycomb complex and cohesin/condensin complex were found exclusively in HMR_{AP} pulldowns. Since HMR not always localizes to centromeres, it is expectable, that HMR_{AP} proximity network has several additional clusters compared to the dCenpA_{AP} network.
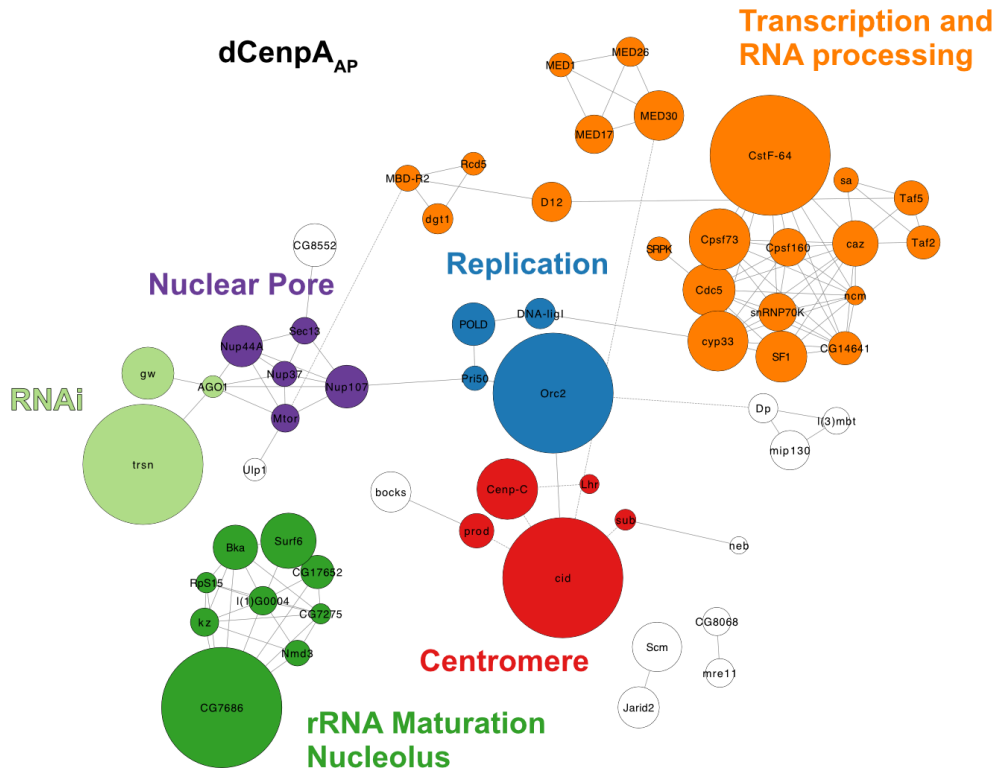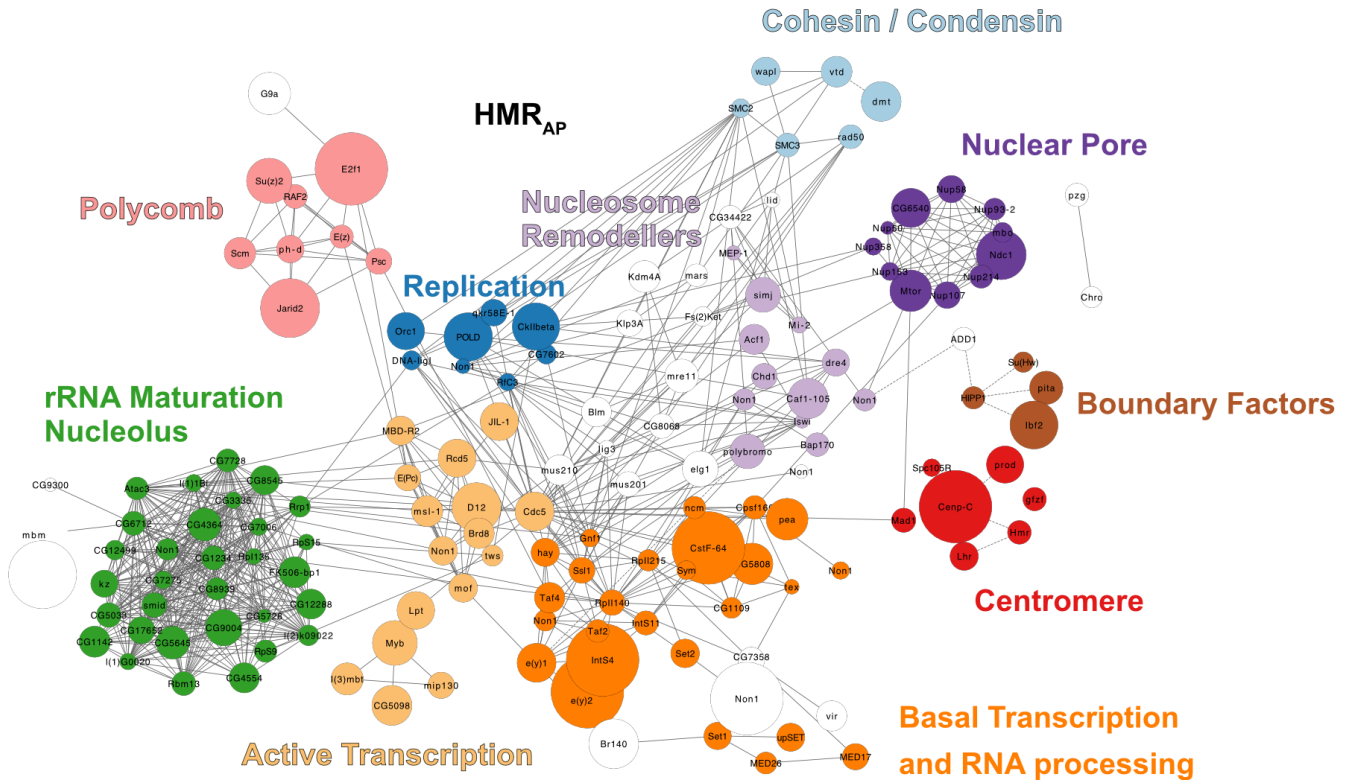
Fig. 2.22. STRING networks of dCenpA$_{AP}$ (A) and HMR$_{AP}$ (B) proximity proteomes. Ungrouped proteins are depicted in white. The size of each node depicts the enrichment in the pulldowns.

2.4.4 HP1a_{AP} proximity proteome upon different conditions.

We performed HP1a_{AP} biotinylation for 1.5 and 5 minutes on plates as well as for 25 minutes in solution. As a control we did not treat cells with biotin-phenol and hydrogen peroxide, but with DMSO. In addition, we performed 1.5, 5 and 25 minutes' labeling for APEX_{NLS}. We built a volcano plot of proteins enriched in HP1a_{AP} pulldowns compared to APEX_{NLS} pulldowns and highlighted the proteins which previously were reported to interact with HP1a (Alekseyenko et al., 2014; Swenson et al., 2016) (Fig. 2.23). We found, that HP1a_{AP} pulldowns after 1.5 minutes of biotinylation had more previously reported HP1a interactors, suggesting that 1.5 minutes' biotinylation is most specific.
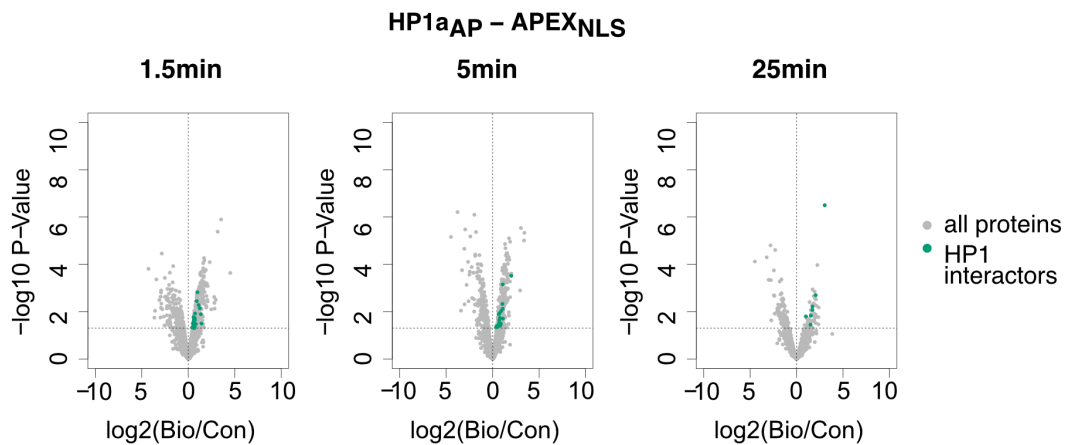
Fig. 2.23. Volcano plots of proteins enriched in HP1a_{AP} but not APEX_{NLS} pulldowns with highlighted HP1a interactors.

2.4.5 Proximity proteome reveals only minor overlap with affinity proteome.

We further compared the results of APEX proteomic technique with previously reported affinity pulldowns for dCenpA (Barth et al., 2014), HMR (Thomae et al., 2013) and HP1a (Alekseyenko et al., 2014; Swenson et al., 2016). Consistent with previous reports for BirA biotin ligase (Lambert et al., 2015), we detected only minor overlap between affinity purification and proximity labeling approaches (Fig. 2.24).

However, if we consider only those proteins, which were shown to localize to centromeres in *Drosophila* (Anselm et al., 2018; Barth et al., 2014; Chen et al., 2015; Erhardt et al., 2008; Heun et al., 2006; Jankovics et al., 2018; Padeken et al., 2013; Swenson et al., 2016; Thomae et al., 2013; Török et al., 1997), the overlap between proximity and IP of HMR and dCenpA is solid (Fig. 2.25). We thus hypothesize, that APEX centromeric labeling gives a strong biotinylation burst at centromeres and less efficient biotinylation at other parts of the nucleus.

HP1a liquid droplet domain might be difficult to purify using conventional IP-MS. This is supported by the fact that IP lists from different labs show minor overlap (Fig. 2.24). Thus, proximity labeling might be a method of choice to describe the composition of the fragile HP1a domain.

dCenpAAP.vs.APEXNLS.proximity

dCenpA.IP

265   9   77

HMRAP.vs.APEXNLS.proximity

HMR.IP

325   11   68

dCenpC

cohesins/condensins

HP1aAP.vs.APEXNLS.proximity.25.min

57

3   2

2

6   24

168   HP1a.IP..Alekseyenko.et.al..2014.

HP1a.IP..Swenson.et.al..2016.

HP1aAP.vs.APEXNLS.proximity.5.min

209

15   0   3

8   23

156   HP1a.IP..Alekseyenko.et.al..2014.

HP1a.IP..Swenson.et.al..2016.

HP1aAP.vs.APEXNLS.proximity.1.5.min

245

19   2   2

6   24

152   HP1a.IP..Alekseyenko.et.al..2014.

HP1a.IP..Swenson.et.al..2016.

Fig. 2.24 Venn diagrams of proteins enriched in bait vs APEX$_{NLS}$ pulldowns and in bait affinity pulldowns.
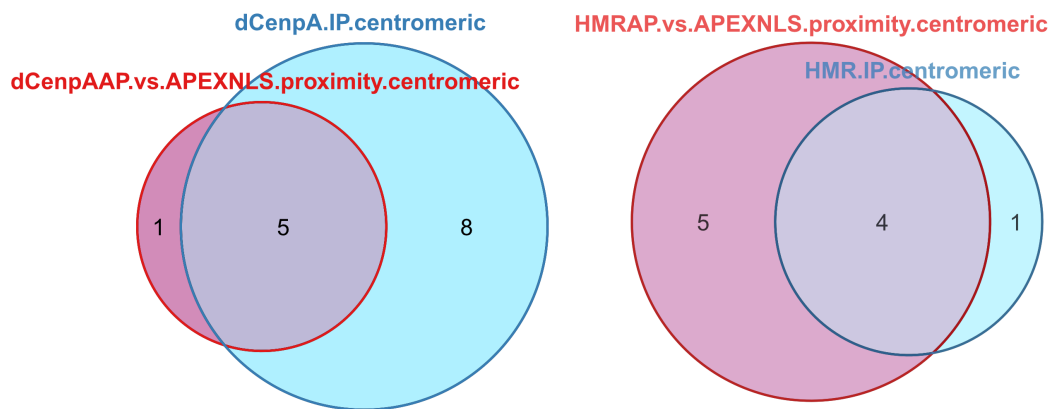
Fig. 2.25. Venn diagrams of proteins, enriched in bait vs APEX$_{NLS}$ pulldowns and in bait affinity pulldowns, which were shown to localize to centromere in *Drosophila*.

2.4.6 All known speciation proteins are found in proximity to HMR.

As mentioned before, hybrid offspring from crosses between *D. melanogaster* females and *D. simulans* males is either lethal or sterile. However, a knockout of *Hmr* gene in *D. melanogaster*, or *Lhr* or *gfzf* in *D. simulans* saves the hybrids. Interestingly, we find both GFZF and LHR proteins in proximity to HMR (Fig 2.26). Possibly, new undiscovered speciation proteins were also found in our proximity pulldowns. It will thus be interesting to establish and perform LHR$_{AP}$ and GFZF$_{AP}$ proximity biotinylation to determine the overlap of proximities of all known three speciation proteins and test the corresponding genes for their role in speciation.
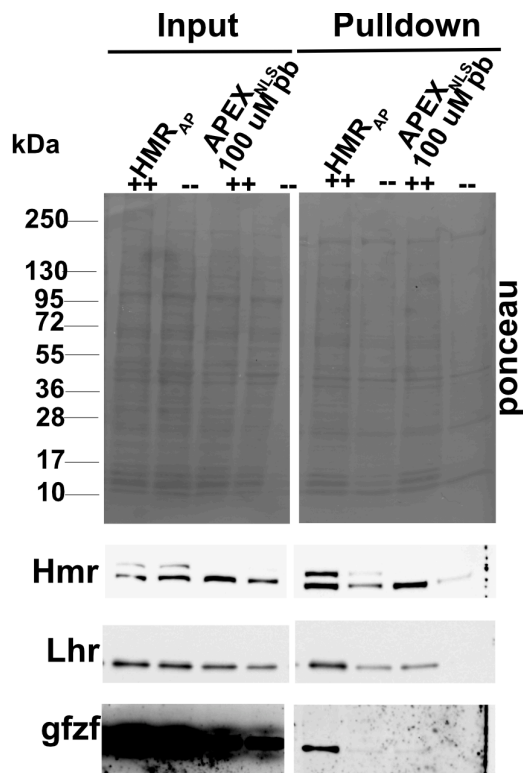
Fig. 2.26. Western blot of HMR_{AP} and $APEX_{NLS}$ pulldowns against HMR, LHR and GFZF. Rat anti-HMR 2C10, rat anti-LHR 12F4 and rabbit anti-GFZF antibodies were used.

2.5 Validation of proximity labeling.

We decided to validate the usability of proximity labeling approach with two alternative methods: immunofluorescence and analysis based on ChIP-sequencing profiles.

We chose 11 proteins found in proximity to the baits, antibodies/tagged proteins' cell lines of which we had in the laboratory and performed immunofluorescent staining. We confirmed previously reported heterochromatic proteins HP5 and ADD1 to localize to heterochromatin (Fig. 2.27, 2.28). We also discovered two new HMR-colocalizing proteins: XNP and CG8108 (Fig. 2.27). Interestingly, we found both proteins in proximity to HMR_{AP}, but not $dCenpA_{AP}$, and according to our finding the staining pattern of those proteins correlated more with HMR than with dCenpA (Fig. 2.28).
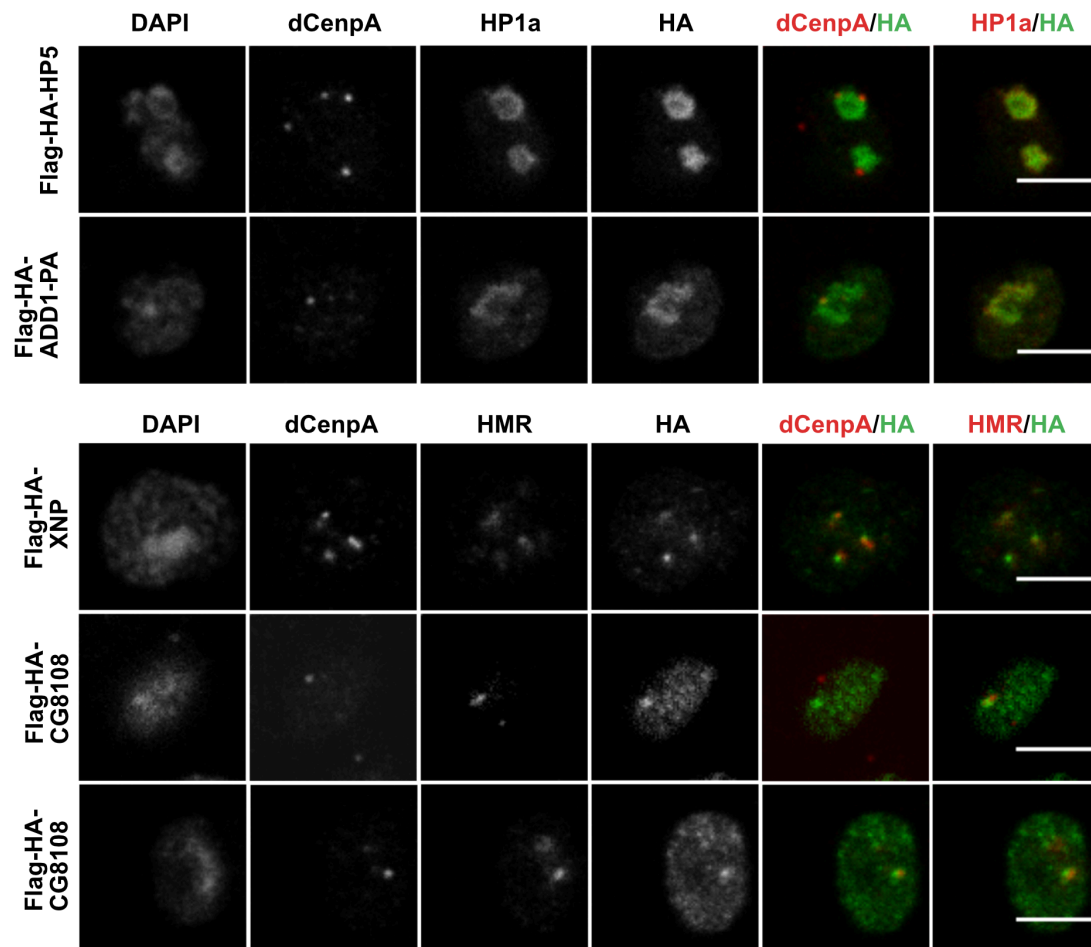
Fig. 2.27. Staining of stable cells lines with rat anti-HA 3F10, rabbit anti-dCenpA (Actif Motif), mouse anti-HP1a C1A9 (upper panel) and mouse anti-HA 12CA5, rabbit anti-dCenpA (Actif Motif), rat anti-HMR 2C10 (lower panel) antibodies. dCenpA/HP1a (upper panel) and dCenpA/HA (lower panel) channels were recorded simultaneously. The percentage of Flag-HA-CG8108 cells among all cells with moderate levels of Flag-HA-CG8108 overexpression which look similar to the panel is 15-30% (calculated from single stack). Scale bars represent 5 uM.
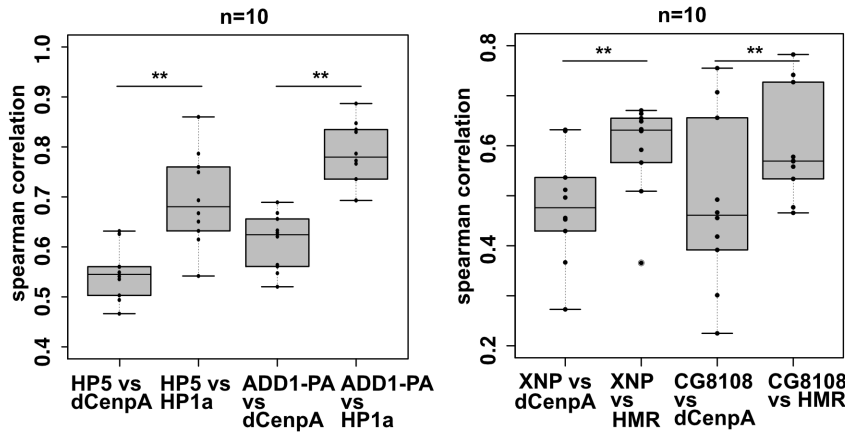
Fig. 2.28. Distributions of Spearman correlations between pairwise stainings of different proteins. 10 cells from 2-3 independent experiments were counted. Wilcoxon signed rank test was used for statistical analysis. ** - p-value<0.01.

We also validated the utility of proximity labeling using high-quality ChIP-sequencing profiles of HMR, HP1a (Gerland et al., 2017), proteins in proximity to those baits and proteins, anti-enriched in the bait pulldowns. For each protein in proximity (anti-proximity) to the bait, we calculated the median distance from the HMR/HP1a peak to the nearest protein peak and compared the distributions of median distances (Fig. 2.29). A statistical comparison of those distributions revealed that the median distance to the nearest peak is 2.5-4 kb for the proteins in proximity to HMR, and 10-12.5 kb for proteins anti-enriched in HMR_{AP} pulldowns. This is statistically significant for all 3 available HMR ChIP-sequencing profiles tested (Cooper et al., 2019; Gerland et al., 2017). For HP1a we compared median distances from the HP1a peak to the nearest peaks of proteins, enriched in 2 out of 3 HP1a biotinylation time points, and proteins anti-enriched in 2 out of 3 HP1a biotinylation time points with enrichment cutoff -0.3. This yielded median distances of 2.8 kb for proteins in proximity to HP1a and 126 kb for proteins anti-enriched in $HP1a_{AP}$ pulldowns. Statistical comparison revealed a trend with p-value=0.067 (Fig. 2.29).
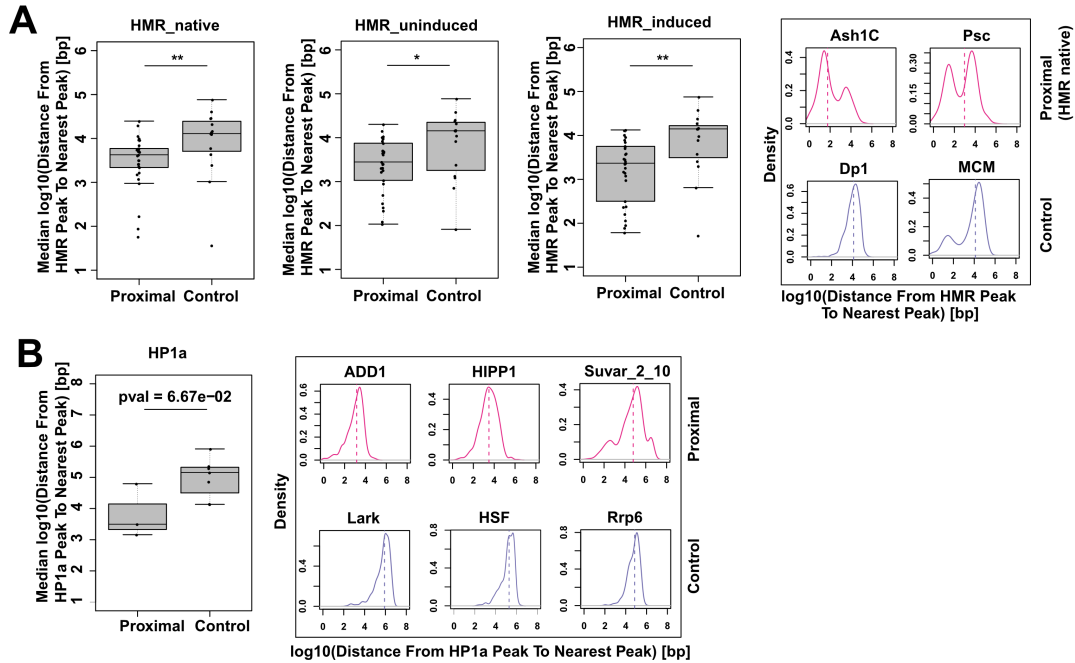
Fig. 2.29. A) Distributions of median distances from the HMR peak to the nearest protein peak. Wilcoxon rank sum test was used for statistical analysis. * - p-value<0.05, ** - p-value<0.01. Examples of distances' distributions for selected proteins are given on the right. B) Distributions of median distances from the HP1a peak to the nearest protein peak. Wilcoxon rank sum test was used for statistical analysis. Examples of distances' distributions for selected proteins are given on the right.

2.6 Spindle assembly checkpoint protein Mad1 is in proximity to HMR$_{AP}$.

We were interested by the identification of spindle assembly checkpoint (SAC) protein Mad1 in proximity to HMR.

Spindle assembly checkpoint is a mechanism that prevents the transition from metaphase to anaphase onset until all the microtubules are properly attached to kinetochores. This prevents from unequal distribution of chromosomes between the daughter cells and further aneuploidy (reviewed in (Musacchio, 2015)). The SAC proteins are recruited to kinetochore in prometaphase. Some of them (Mad1, Mad2, BubR1, Bub3 and Cdc20) form sequential complexes which eventually inhibit the activity of the anaphase

promoting complex (APC), also called the cyclosome, which in turn prevents the degradation of Cyclin B and Securin (reviewed in (Conde et al., 2013)).

The striking logical discrepancy in our results is that SAC components persist at centromeres in mitosis, while HMR is removed from mitotic centromeres (Thomae et al., 2013). We hypothesized, however, that HMR already in interphase creates a certain chromatin state that might favor or disfavor the recruitment of SAC components in mitosis.

2.6.1 Co expression of HMR and LHR in interphase slightly increases BubR1 amounts in prometaphase/metaphase.

Before actually accessing HMR role in SAC, we tried to stain the cells with antibodies against several SAC components, and of this BubR1 antibody showed a localized centromere staining in mitosis (Fig. 2.30, 2.31). We thus decided to decipher the influence of HMR levels on BubR1 levels using confocal microscopy.

BubR1 short exposure
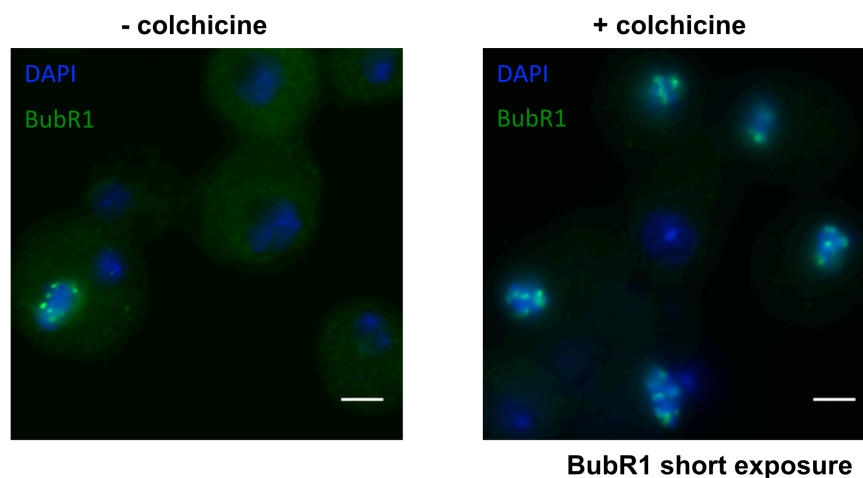
Fig. 2.30. Staining of cells with rabbit anti-BubR1 antibody upon colchicine treatment. Scale bars represent 5 uM.

Cells were treated with colchicine, and HMR was either knocked down or overexpressed (for this, induced cell line with Flag-HA-HMR + Myc-LHR under copper inducible promoter (Thomae et al., 2013) was used). The 0.025 mM colchicine treatment lasted for 16 hours (Godinho and Tavares, 2008) to depolymerize microtubules, arrest cells in prometaphase (so that enough

prometaphase cells could be collected for microscopy experiment) and induce accumulation of SAC proteins on prometaphase/metaphase chromosomes (to obtain enough bright signal at the centromeres for quantification) (Fig. 2.31). The cells were stained with DAPI, as well as with anti-HMR and anti-BubR1 antibodies. The BubR1 signal was scanned through Z-stacks, and sum intensities' projections were obtained in ImageJ. The signal was manually quantified in ImageJ for 5 random cells in each technical replicate, and 2 technical replicates were taken for each condition in each biological replicate. At least 5 biological replicates of the experiment were performed.
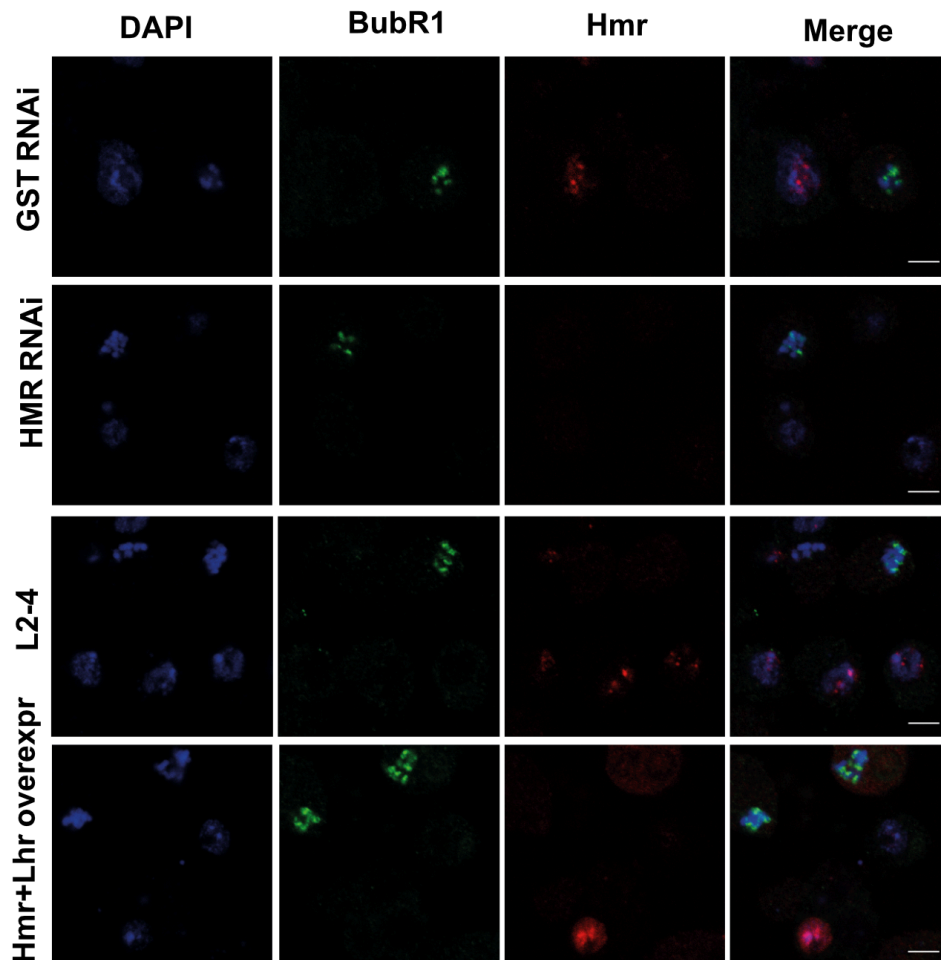
Fig. 2.31. Representative immunofluorescence images of BubR1 and HMR stainings upon HMR RNAi and HMR+LHR overexpression. Antibodies used for staining are rat anti-HMR 2C10 and rabbit anti-BubR1. Scale bars represent 5 uM.

We found, that upon HMR knockdown BubR1 intensity was not significantly affected, however, upon HMR and LHR overexpression we detected a slight increase in BubR1 intensity (Fig. 2.32). Thus, HMR+LHR overexpression slightly affects the SAC, and it would be tempting to speculate that SAC is slightly affected in hybrids from *D. melanogaster* mothers and *D. simulans* fathers, where HMR and LHR are known to be overexpressed.
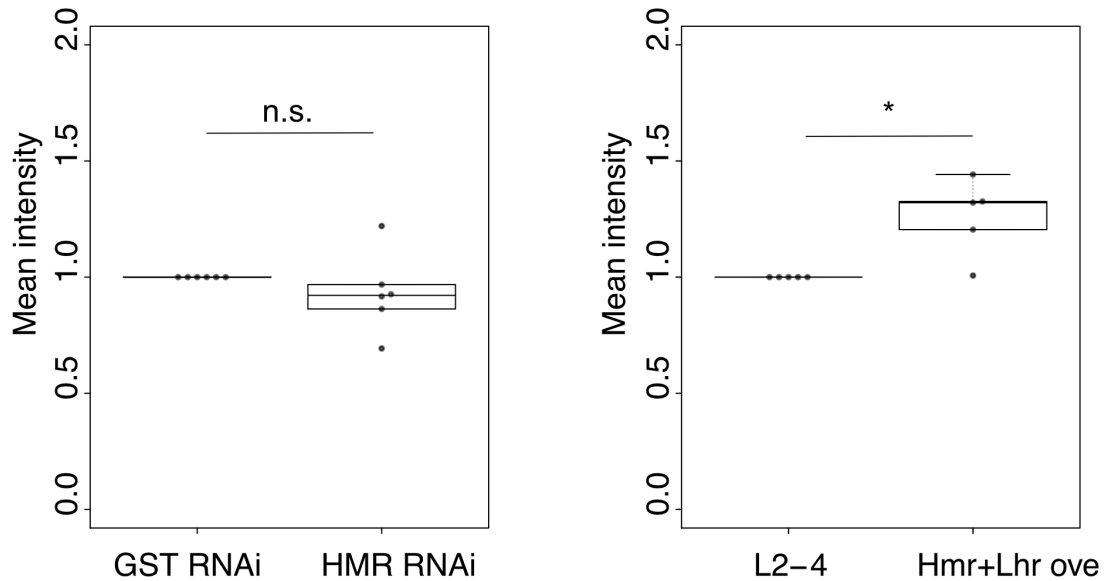
Fig. 2.32. Quantification of BuBR1 intensity upon HMR knockdown and HMR+LHR overexpression. Paired t-test was used for statistical analysis.

2.7 dCenpC RNAi brings centromeres and HMR domains together.

2.7.1 dCenpC RNAi results in centromere and HMR, but not pericentric heterochromatin declustering.

dCenpC was previously reported to be dCenpA and HMR interactor (Barth et al., 2014; Erhardt et al., 2008; Thomae et al., 2013). Interestingly, we found dCenpC in proximity to both dCenpA and HMR, but not HP1a. The role of dCenpC in dCenpA incorporation has been investigated (Erhardt et al.,

2008), however, its role in centromere and pericentric chromatin architecture has been less described. We thus wondered how dCenpC deletion would influence the chromocenter architecture. We removed dCenpC by RNAi, confirmed the knockdown by western blotting and stained the cells against HP1a, dCenpA and HMR. We observed that upon dCenpC RNAi HMR was no more localized to centromeres, but co-stained with HP1a (Fig. 2.33 A). Interestingly, ChIP-sequencing revealed that upon dCenpC RNAi HMR-binding sites did not change much (Fig. 2.34 A and B). Also, HMR levels were not changed (Fig. 2.33 B). We therefore conclude that HMR-binding sites decluster from near-centromeric border and diffuse into heterochromatin.
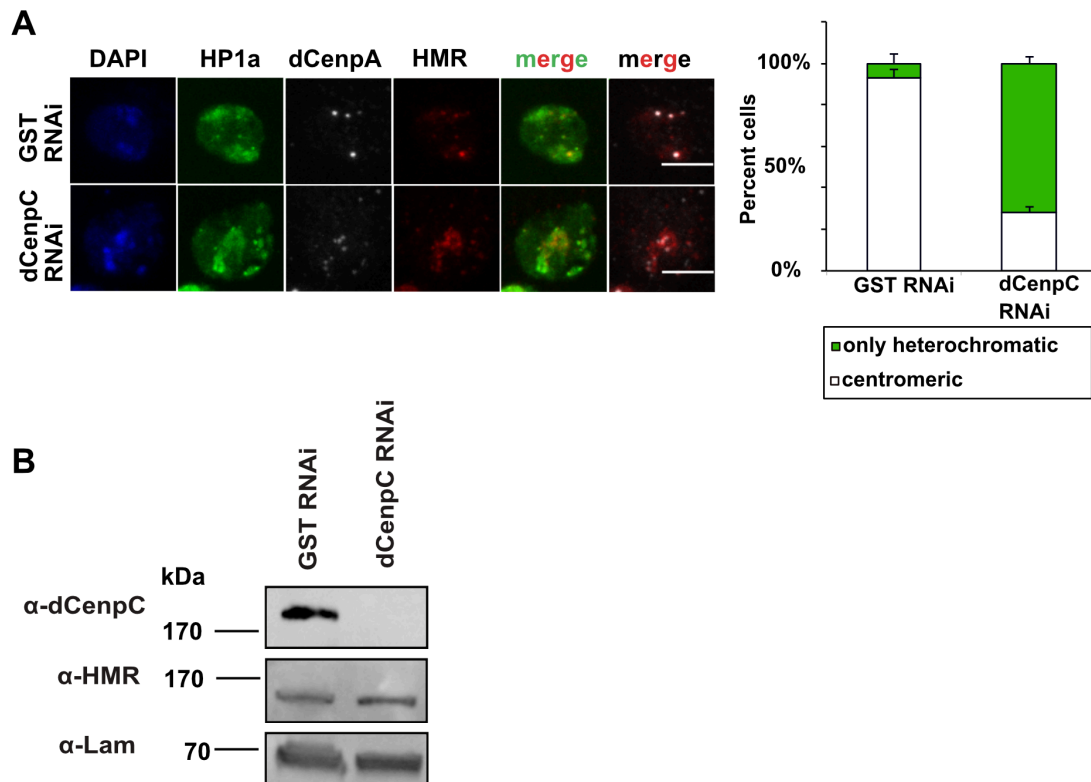
Fig. 2.33. dCenpC RNAi results in HMR mislocalization from centromere. Pioneer experiment was performed by Dr. Andreas W. Thomae. A) Right panel: representative immunofluorescent images upon GST and dCenpC knockdown. Cells were stained with mouse anti-HP1a C1A9, rat anti-HMR 2C10 and rabbit anti-dCenpA (Actif Motif). Scale bars represent 5 uM. Left panel: quantification of centromeric and only heterochromatic localization of HMR. Localization was considered heterochromatic if less than 20% of centromeres in a cell colocalized with HMR. Error bars represent standard deviation. 2 independent experiments were performed, 50 cells quantified in each. B) Western blotting against dCenpC, HMR and Lamin upon GST and dCenpC RNAi. Rabbit anti-dCenpC, rat anti-HMR 2C10 and mouse anti-Lamin antibodies were used.
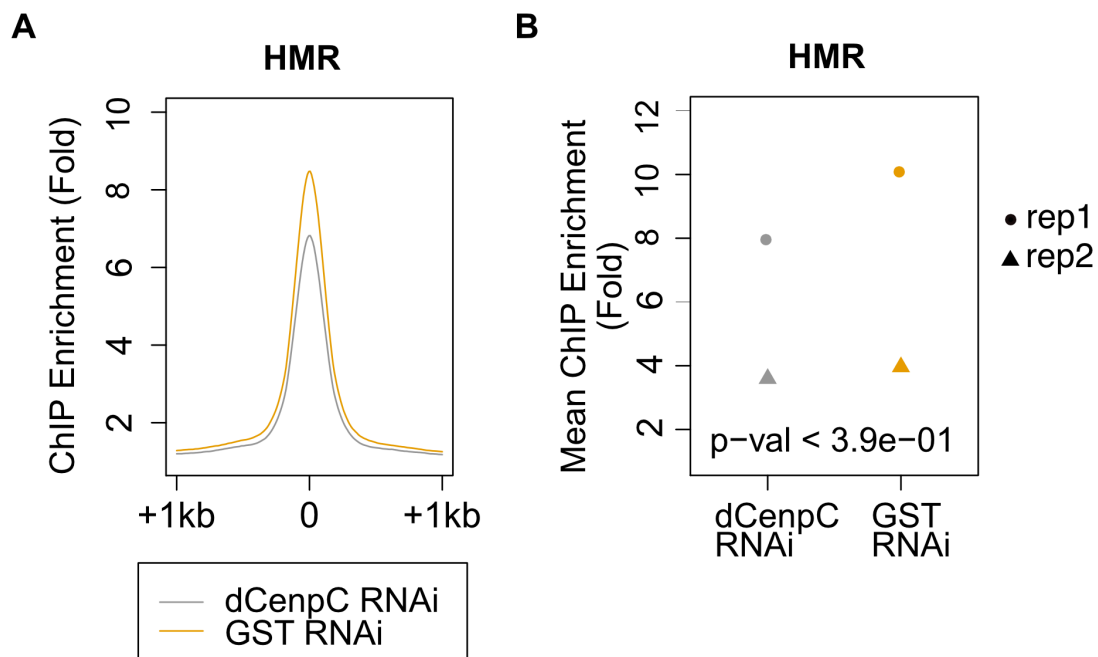
Fig. 2.34. A) ChIP enrichment of HMR upon GST and dCenpC RNAi. B) ChIP enrichment (fold) of HMR upon GST and dCenpC RNAi in different replicates. Paired t-test was used for statistical analysis. Rat anti-HMR 2C10 antibody was used for ChIP-sequencing.

By quantifying the number of dCenpA foci upon GST and dCenpC RNAi, we concluded that centromeres also declustered (Fig. 2.35 A). In contrast, pericentric heterochromatin, as judged by HP1a staining (Fig. 2.33 A) and

number of D1 protein domains, which are located specifically at pericentromeres (Blattes et al., 2006) (Fig. 2.35 B), did not exhibit declustering.
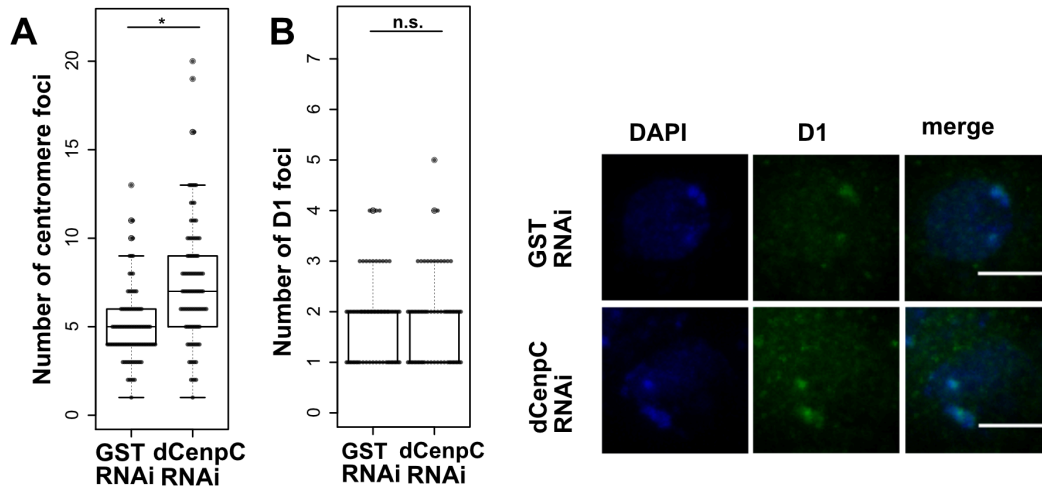
Fig. 2.35. A) Distributions of numbers of centromeric foci upon GST and dCenpC RNAi. * - p-value<0.05. Rat anti-dCenpA 7A2 antibody was used for staining. Since dCenpA staining is reduced upon dCenpC RNAi, different exposures were used for recording images upon GST and dCenpC RNAi. B) Left panel: distributions of D1 foci number upon GST and dCenpC RNAi. N.s. – non-significant. Right panel: representative images of D1 staining. Rabbit anti-*D. simulans* D1 antibody was used. Scale bars represent 5 uM. For statistical analysis in A and B a common linear model including means of foci number for each replicate was fitted (formula: Value=Replicate+Condition*Domain). 2 independent experiments were performed, 30-50 cells quantified in each.

2.7.2 dCenpC N-terminus is not necessary for HMR and centromere clustering.

In order to rescue the phenotype of HMR and dCenpA declustering, we generated a construct of Flag-HA-tagged dCenpC, resistant to 2 dsRNAs. We were unable to clone the construct without mutations because of bacterial anti-selection, so for our study we chose a plasmid with 2 mutations in RNAi-resistant dCenpC (Fig. 2.36 A). We also generated 2 mutant deletion constructs, one lacking C-terminus, the other – N-terminus. All full-length and

deletion constructs showed correct expression pattern on western blotting upon dCenpC RNAi and transient transfection (Fig. 2.36 B). As described previously (Heeger et al., 2005), N-terminal deletion mutant localized to centromeres, whereas C-terminal deletion mutant not (Fig. 2.37).
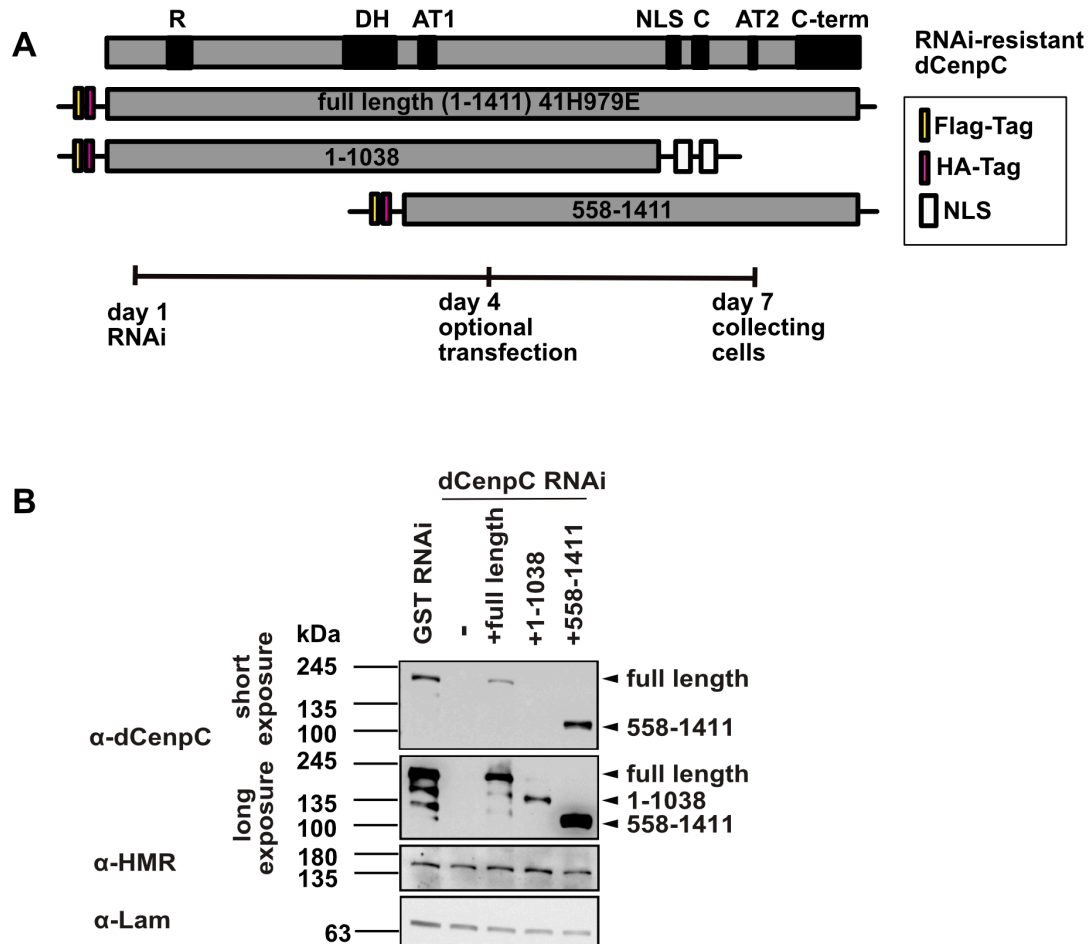
Fig. 2.36. A) Upper panel: scheme of dCenpC (was taken from (Heeger et al., 2005)) and RNAi-resistant dCenpC constructs used for rescue experiments. R – arginine-rich domain; DH – a conserved *Drosophila* dCenpC region, AT1/2 – predicted AT-hooks; NLS – nuclear localization signal; C – CenpC motif; C-term – metazoan-like C-terminus. Bottom panel: Scheme of the rescue experiment. B) Western blotting against dCenpC, HMR and Lamin upon GST and dCenpC RNAi, as well as transient transfection of RNAi-resistant dCenpC proteins. Rabbit anti-dCenpC, rat anti-HMR 2C10 and mouse anti-Lamin antibodies were used.

Upon dCenpC RNAi and transient transfection, full-length and N-terminal deletion mutant rescued HMR centromeric localization and partially rescued centromere declustering, whereas C-terminal deletion mutant did not (Fig. 2.37, 2.38 A and B).
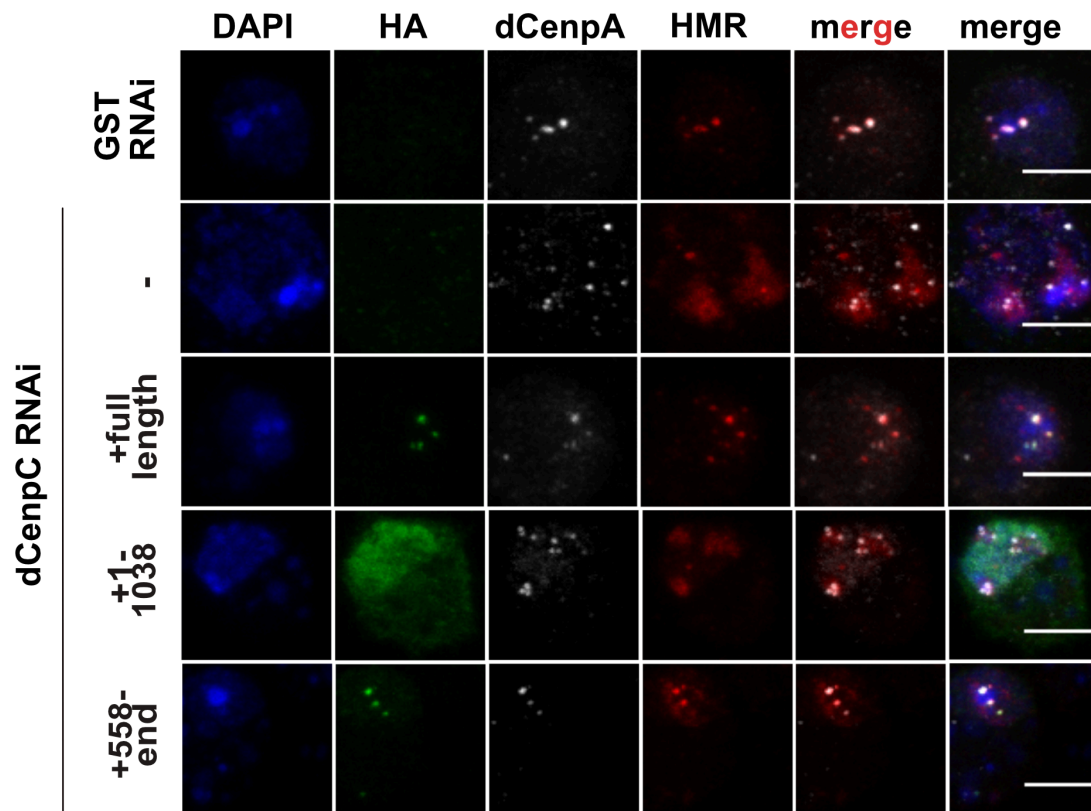
Fig. 2.37. Representative immunofluorescence images upon GST and dCenpC RNAi, as well as transient transfection of different RNAi-resistant dCenpC constructs. Mouse anti-HA 12CA5, rabbit anti-dCenpA (Actif Motif) and rat anti-HMR 2C10 antibodies were used for staining.
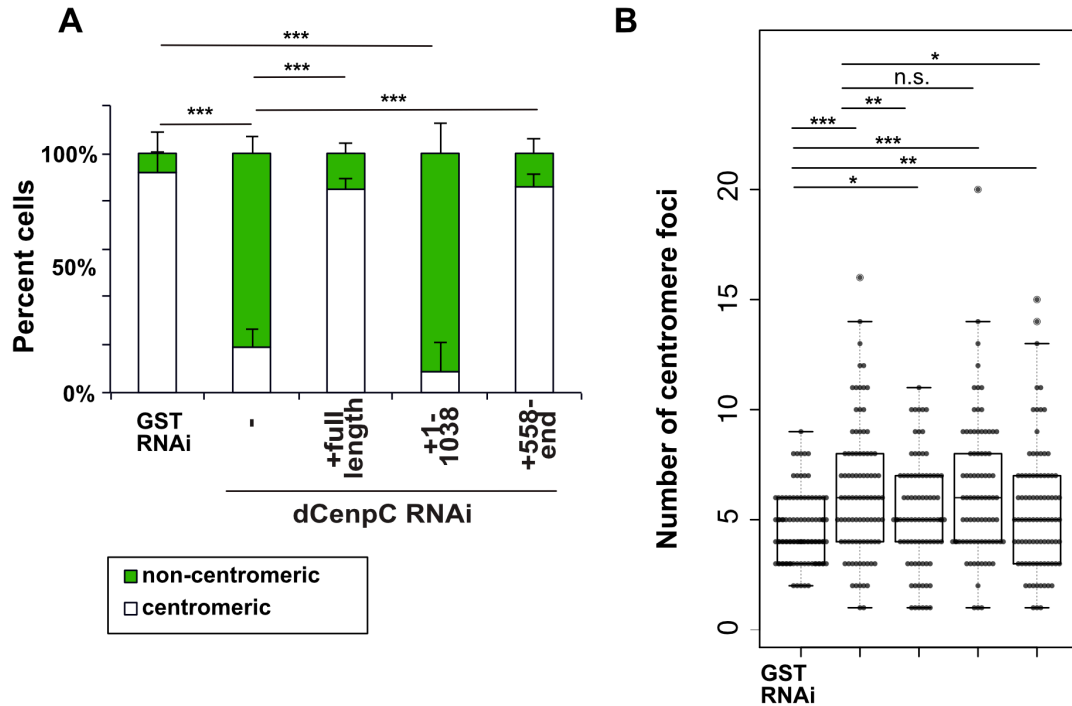
Fig. 2.38. A) Quantification of centromeric and non-centromeric localization of HMR. Localization was considered non-centromeric if less than 20% of centromeres in the cell colocalized with HMR. Linear model on centromeric fractions for each replicate with formula Value=Condition+Replicate was fit for statistical analysis. Error bars represent standard deviation. 2 independent experiments were performed, 40-50 cells quantified in each. B) Distributions of numbers of centromeric foci upon GST and dCenpC RNAi, as well as transient transfection of different RNAi-resistant dCenpC constructs. Rat anti-dCenpA 7A2 antibody was used for staining. Since dCenpA staining is reduced upon dCenpC RNAi, different exposures were used for recording images upon GST and dCenpC RNAi. Linear model on mean values for each replicate with formula Value=Condition+Replicate was fit for statistical analysis. 2 independent experiments were performed, around 40 cells quantified in each. N.s. – non-significant, * - p-value<0.05, ** - p-value<0.01, *** - p-value<0.001.

2.8 Condensins and cohesins reside in proximity to HMR and CAP-H2 binding to chromatin is reduced upon HMR+LHR overexpression.

Interestingly, among protein networks identified in HMR_{AP} vs $APEX_{NLS}$ pulldowns, we found condensin and cohesin complexes. This is particularly interesting because *Hmr* mutant larvae were shown to have mitotic defects in brain cells similar to cohesin knockouts, and hybrids between *D. melanogaster* and *D. simulans*, where HMR is known to be overexpressed, exhibit brain cells' phenotypes (under certain conditions) reminiscent of problems with chromosome condensation (Blum et al., 2017; Bolkan et al., 2007). Our proximity labeling results suggested that there is a direct molecular link between HMR and condensins/cohesins.

2.8.1 HMR peaks overlap with Rad21 and CAP-H2 peaks.

We took advantage of the ChIP-sequencing data of cohesin subunit Rad21 (vtd) and condensin subunit CAP-H2 in Kc167 cells (Van Bortle et al., 2014), as well as HMR in S2DGRC cells (Gerland et al., 2017), and compared these ChIP-sequencing profiles. Interestingly, HMR peaks overlapped well with Rad21 peaks (51%), and to a lesser degree with CAP-H2 peaks (20%) (Fig. 2.39 A). We also took advantage of the HMR ChIP-sequencing data upon HMR and LHR overexpression (HMR_{OVER}) (Cooper et al., 2019), and determined the peak overlap of HMR_{OVER} with Rad21 and CAP-H2 (Fig. 2.39 A). We found out, that approximately 57% of HMR_{OVER}-binding sites overlapped with Rad21 and 33% with CAP-H2 binding sites. Strikingly, upon overexpression HMR_{OVER} occupies nearly one-third (28.4%) of all genome-wide condensin binding sites.
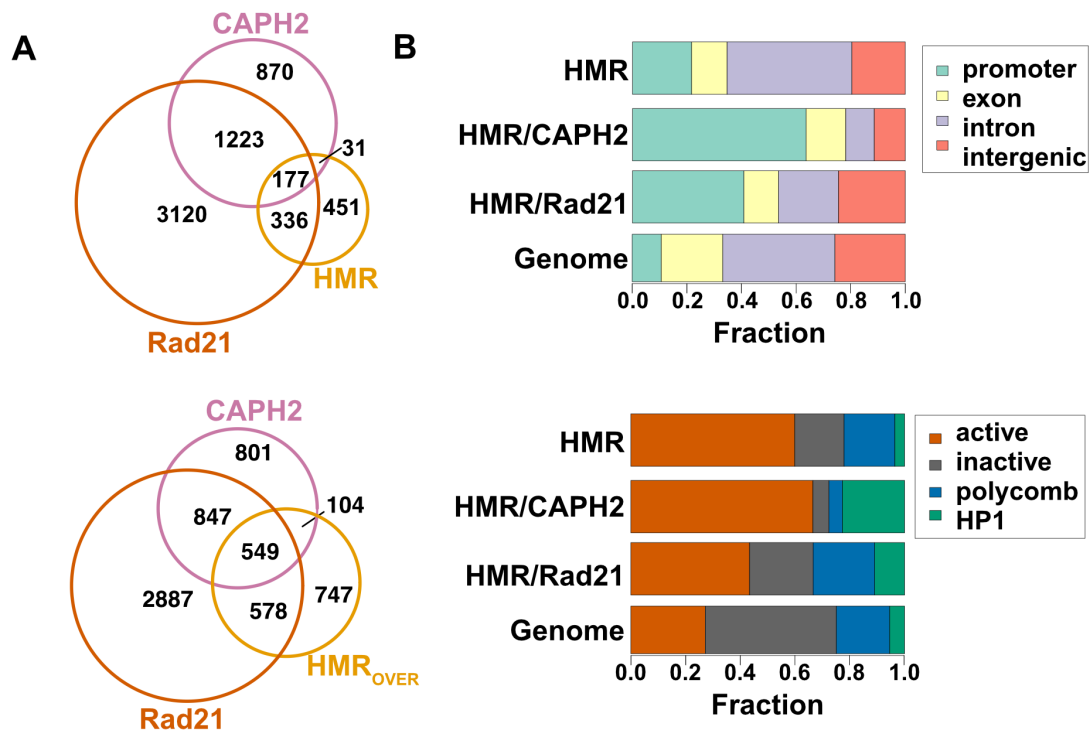
Fig. 2.39. A) overlap of HMR or HMR$_{OVER}$ peaks with Rad21 and CAPH2 peaks. B) Distribution of HMR, HMR/Rad21 and HMR/CAPH2 peaks across different genomic features and different chromatin colors.

We next plotted the distribution of HMR/Rad21/CAP-H2 binding sites across different genomic features and chromatin colors using the 5-state model (Filion et al., 2010). We found out, that HMR/Rad21 and HMR/CAP-H2 sites were mainly localized at promoters, and HMR/CAP-H2 sites were mainly localized at active and HP1a chromatin types (Fig. 2.39 B).

2.8.2 CAP-H2 resides at class I HMR binding sites but resides at other class upon HMR overexpression.

Previously, HMR binding sites were classified into two groups: class I borders HP1a at promoters of actively transcribed genes, and class II colocalizes with *gypsy* insulators, defined by the presence of insulator proteins Su(Hw), Cp190 and mod(mdg4). Notably, when we aligned the ChIP-

sequencing signals in the form of a heatmap, we found that CAP-H2 predominantly localizes to class I binding sites, and Rad21 does not display a preference for either class I or class II sites. However, if we compare Rad21 and CAP-H2 binding profiles upon HMR native conditions with the one of HMR$_{OVER}$, CAP-H2 goes to the HMR$_{OVER}$ binding sites which do not belong to class I (Fig. 2.40).
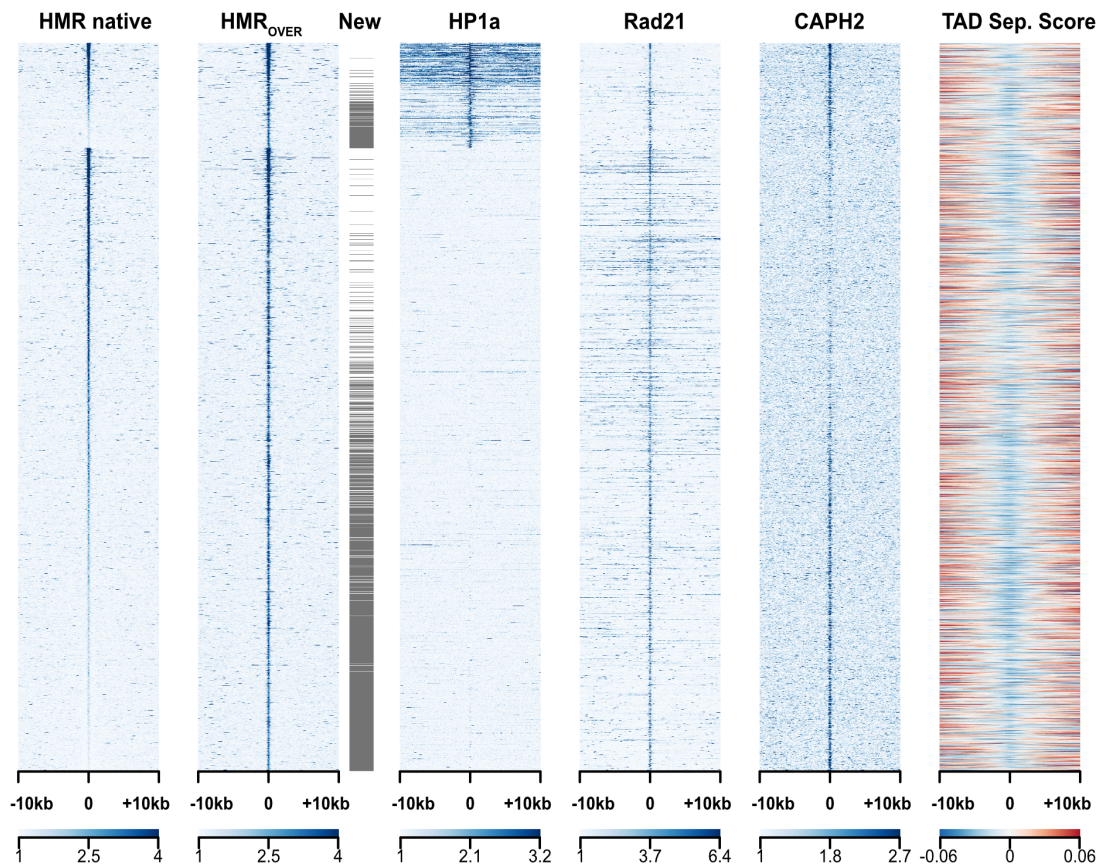
Fig. 2.40. Heatmaps of HMR, HMR$_{OVER}$, HP1a, Rad21 and CAPH2 binding profiles and TAD separation score. Peaks are sorted by HP1a signal.
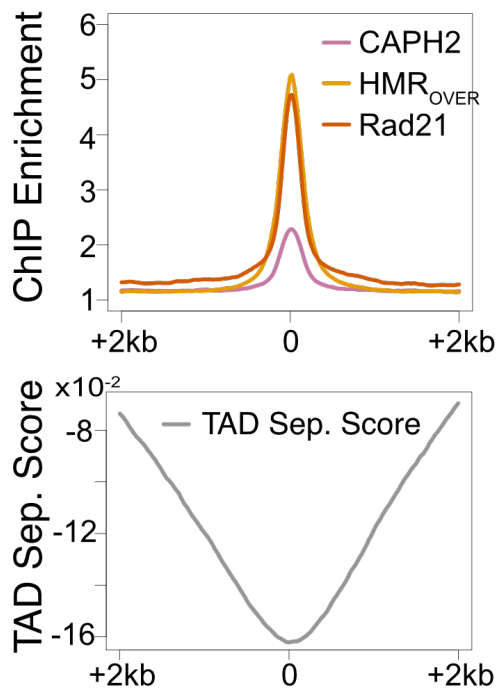
Fig. 2.41. Composite plot of HMR_{OVER}, Rad21 and CAPH2 ChIP-sequencing profiles, as well as TAD separation score.

2.8.3 HMR colocalizes with TAD boundaries.

Cohesins and condensins, known to play a role in nuclear organization, were shown to cluster at TAD boundaries (Van Bortle et al., 2014). We thus wondered, whether HMR also resides there. We aligned HMR ChIP-sequencing signal with the TAD separation score (Ramirez et al., 2018). The minima in the score (blue on the heatmap) represents a TAD boundary, while the maxima in the score (red on the heatmap) represents a "TAD peak". Interestingly, HMR binding sites often colocalized with minima in TAD separation score (Fig. 2.40, 2.41). Moreover, upon HMR overexpression more HMR peaks colocalized with TAD boundaries. Thus, HMR resides at the genome sites, which are important for 3D organization of the nucleus.

2.8.4. HMR+LHR overexpression results in reduced CAP-H2 binding to chromatin.

HMR resides in proximity to condensins. HMR peaks overlap with CAP-H2 peaks. Hybrids, where HMR and LHR are overexpressed, exhibit possible problems with chromosome condensation. We thus wondered about the effect of HMR+LHR overexpression on CAP-H2 binding sites. We induced expression

of proteins in stable cell line expressing Flag-HA-HMR+Myc-Lhr (Thomae et al., 2013), as well as induced control untransfected cells, and measured genome-wide binding of CAP-H2 by ChIP-sequencing. We found, that CAP-H2 signal was reduced upon HMR+LHR overexpression (Fig. 2.42, 2.43 A). This was prominent both at HMR-bound and HMR-unbound sites, as judged by the pooled HMR ChIP-seq profile (Fig. 2.43 B).

In agreement with previous data, our CAP-H2 ChIP-seq showed CAP-H2 localization to TAD boundaries (Fig. 2.44 A). Fig. 2.44 B shows log2 Hi-C counts, as well as log2 ratio of observed versus expected Hi-C counts at sites, bound by CAPH2 and HMR, as well as by CAPH2 alone. As seen by the high observed versus expected Hi-C counts ratio at cross-projections of both types of sites, CAPH2-HMR bound sites on average likely contact CAPH2-only bound sites. This happens at TAD boundaries (Fig. 44 B). Thus, we hypothesize a direct effect of HMR on CAP-H2.
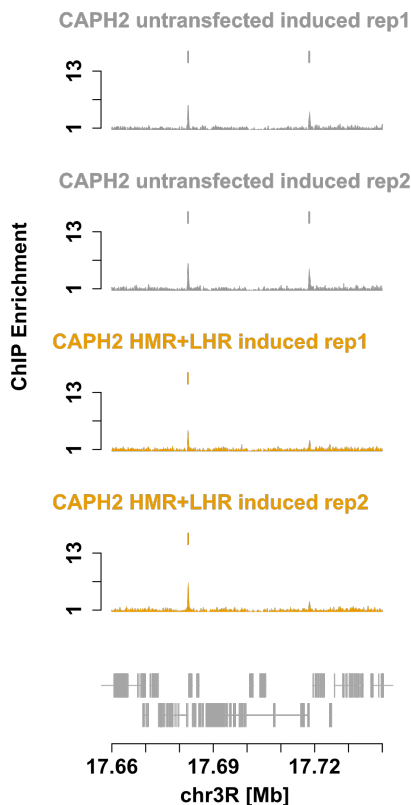
Fig. 2.42. Genomic tracks of CAP-H2 ChIP showing a region of right arm of chromosome 3.
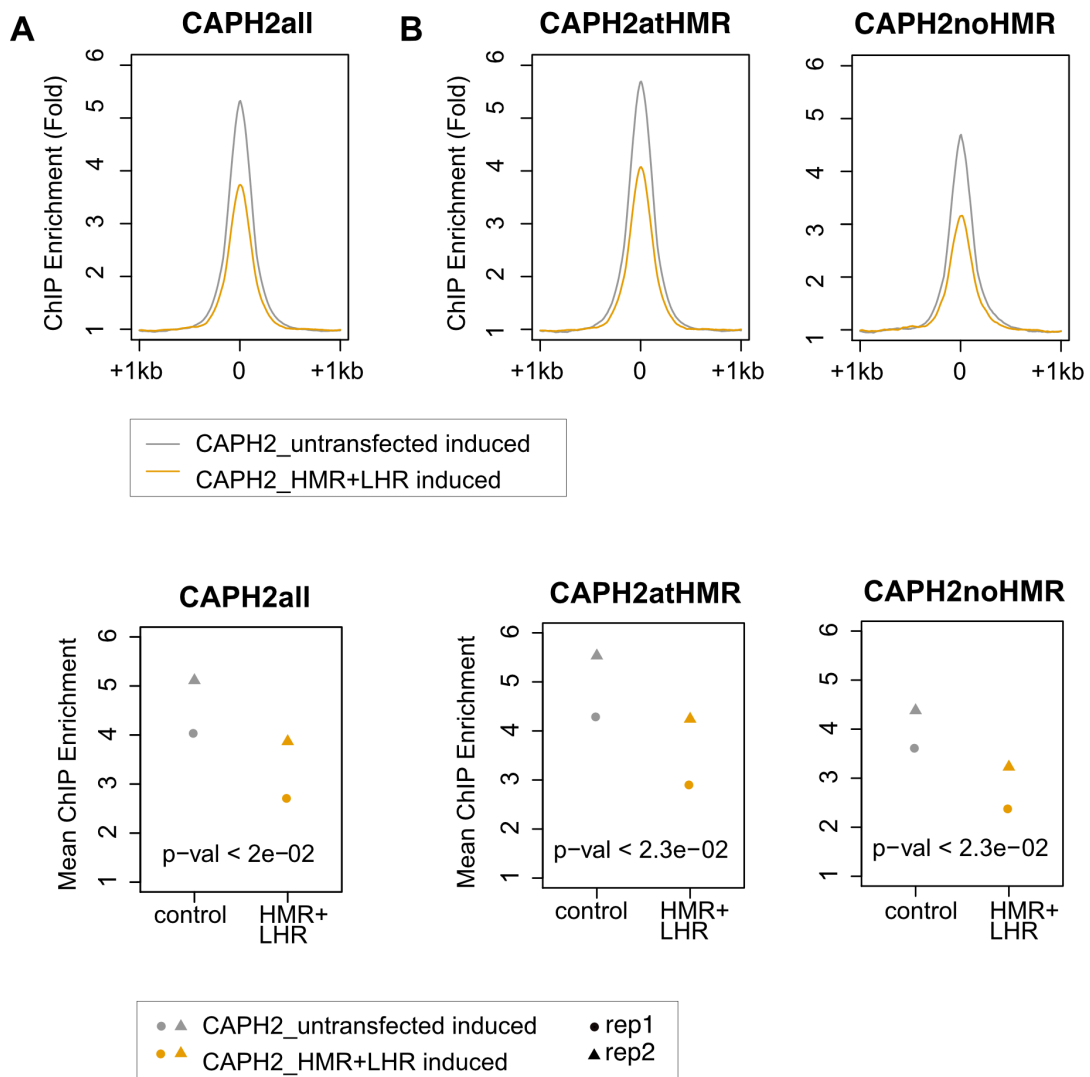
Fig. 2.43. Upper panel: ChIP enrichment of CAP-H2 in control and HMR+LHR overexpression. Lower panel: ChIP enrichment of CAP-H2 in control and HMR+LHR overexpression in different replicates. Paired t-test was used for statistical analysis. Rabbit anti-CAP-H2 antibody was used for ChIP-sequencing.
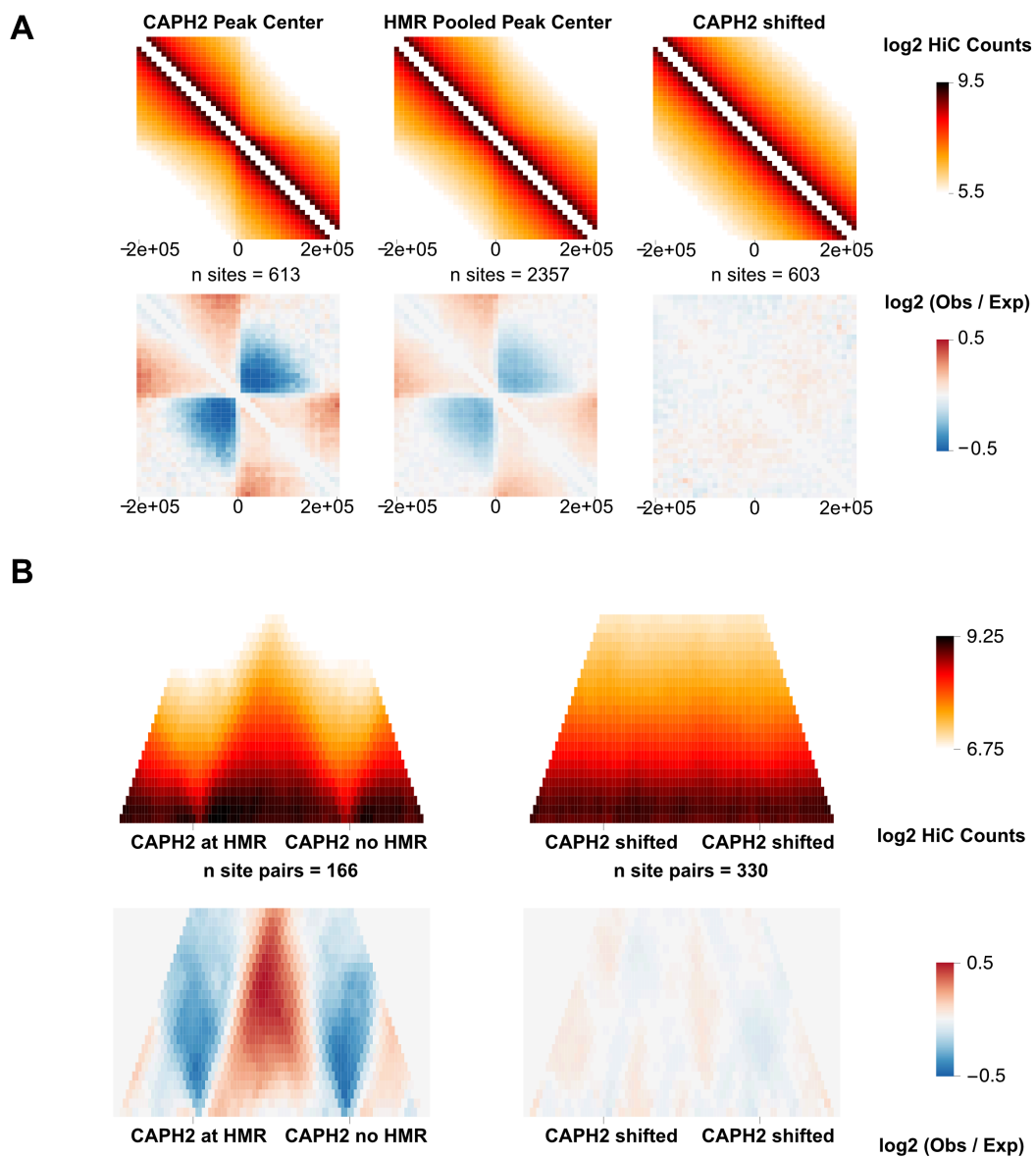
Fig. 2.44. A) averaged Hi-C maps at CAP-H2, HMR and randomly shifted CAP-H2 (as a control) peak centers. B) Averaged Hi-C maps at pairs of sites. Observed (Obs): Hi-C contact frequency, expected (Exp): average contact frequency as a function of genomic distance. Hi-C data were taken from (Ramirez et al., 2015).

2.9 Condensins and cohesins are not required for HMR´s localization to heterochromatin.

Since we found and confirmed that HMR localizes in proximity to condensins and cohesins, as well as showed HMR+LHR overexpression effect on CAP-H2, we wondered whether cohesin and condensin proteins analogously to dCenpC might contribute to centromeric residence of HMR. We performed RNAi of cohesin subunits Rad21 and SMC1 and condensin subunits CAP-H2 and SMC2 and measured HMR localization to centromere. It was not affected, suggesting that cohesins and condensins do not dramatically contribute to the centromere architecture (Fig. 2.45, 2.46 A and B).
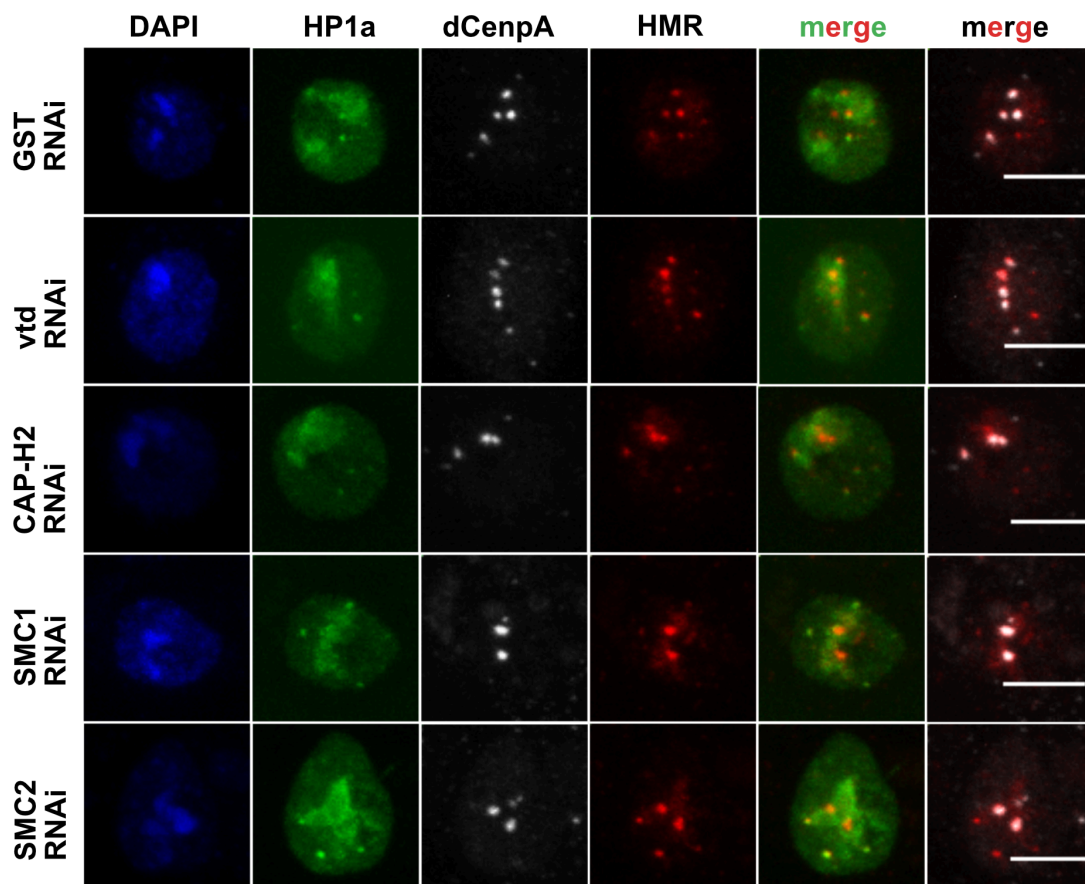
Fig. 2.45. Condensin and cohesin RNAi does not result in HMR mislocalization from the centromere. Representative immunofluorescent images upon GST, cohesin subunits Rad21 and SMC1 and condensin subunits CAPH2 and SMC2 knockdown. Cells were stained with mouse anti-HP1a C1A9, rat anti-HMR 2C10 and rabbit anti-dCenpA (Actif Motif) antibodies. Different exposures for dCenpA and HP1a might have been recorded.
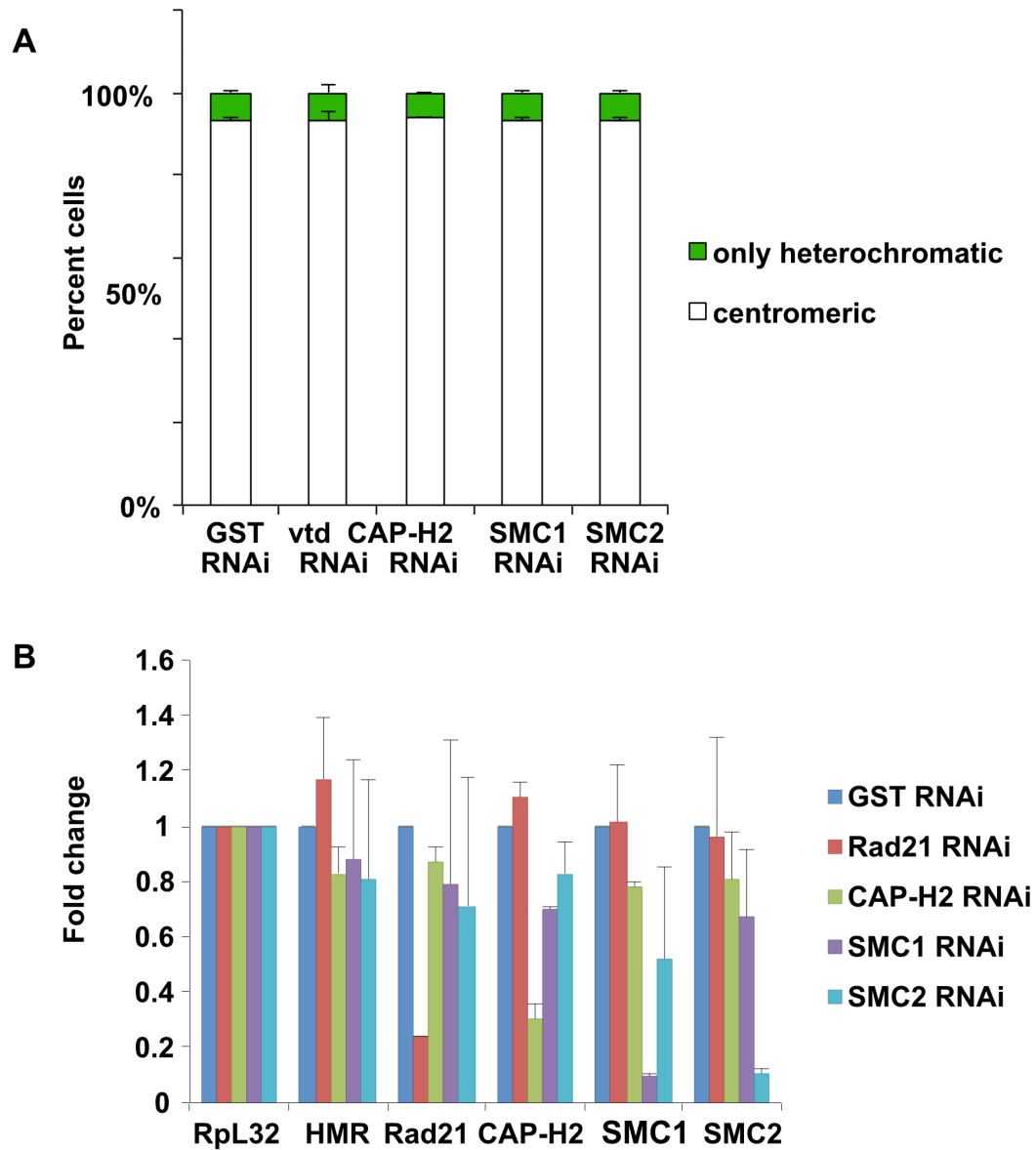
Fig. 2.46. Condensin and cohesin RNAi does not result in HMR mislocalization from the centromere. A) quantification of centromeric and only heterochromatic localization of HMR. Localization was considered only heterochromatic if less than 20% of centromeres in the cell colocalized with HMR. Error bars represent standard deviation. B) verification of cohesins and condensins knockdown by RT-qPCR. Error bars represent standard deviation. Primers are available in Table 4.

3 Discussion.

3.1 The structure of the chromocenter is more complex than thought before.

In this thesis my collaborators and I revealed a much more complex structure of the chromocenter than was described before. Previously it was considered that centromere is a round structure of dCenpA-containing chromatin. We together with (Anselm et al., 2018) showed the interdigitation between domains of different proteins.

Using the CRISPR cell line, where HMR is tagged with FLAG at C-terminus (Gerland et al., 2017), we confirmed HMR localization to the centromeres. Interestingly, HMR was associated not with all centromeres (Thomae et al., 2013) but only with a subset. It might reflect the cell-cycle regulation as well as specific localization of HMR to a subset of chromosomes (similar, for example, to the POF protein which associates with 4th chromosome only (Larsson et al., 2001)).

We investigated the structure of the chromocenter with high-resolution STED microscopy using antibodies directed against HMR, dCenpA and dCenpC. Strikingly, the experiments revealed that centromere is not a blob of dCenpA containing chromatin, but a very complex structure consisting of interdigitating domains containing individual proteins. Our experiments revealed a substantial, but not complete, colocalization between dCenpA and dCenpC, and less colocalization between HMR/dCenpA and HMR/dCenpC.

It was unexpected that dCenpA and dCenpC domains do not overlap completely. dCenpA incorporation drastically depends on dCenpC (Erhardt et al., 2008), the only known protein of CCAN in *Drosophila* except dCenpA (Barth et al., 2014). Possibly, dCenpC contributes to dCenpA incorporation indirectly, for example by maintaining a proper structure of the centromeric chromatin. Alternatively, dCenpA incorporation might start at sites of its colocalization with dCenpC and might then spread to adjacent sites already independent of dCenpC. One more possibility is that individual centromeric domains are very dynamic and only partially overlap at a particular time point.

Lack of mutual colocalization of HMR and dCenpA is in agreement with previous results that HMR does not influence dCenpA incorporation (Thomae et al., 2013) and is not found in dCenpA affinity pulldowns (Barth et al., 2014).

In agreement with numerous previous reports, we find centromeres to be embedded in HP1a-containing chromatin. Strikingly, upon confocal microscopy of HP1a and STED microscopy of HMR and dCenpA, we find that HMR domains are often bordering dCenpA domains from HP1a.

3.2 Exact molecular function of HMR remains unknown.

Previously, HMR loss in cells and flies has been associated with defects in mitosis (Blum et al., 2017; Thomae et al., 2013), upregulation of transposable elements (Satyaki et al., 2014; Thomae et al., 2013), as well as with influence on gene expression (Gerland et al., 2017; Thomae et al., 2013) and telomere length (Satyaki et al., 2014).

In more detail, the mitotic defects in HMR-deficient larval brains – broken chromosomes - resemble problems with chromosome cohesion (Blum et al., 2017). In the brains of hybrid larvae, where HMR together with its partner protein LHR is known to be overexpressed (Thomae et al., 2013), chromosomes were reported to be uncondensed and fuzzy (Orr et al., 1997). This was proposed to be an artifact, since larval brains treated not with physiological salt solution but with Schneider medium retained condensed chromosomes (Bolkan et al., 2007). However, one might argue that this is not an artifact, but just an enhancement of problems with chromosomal condensation which have started before but appear well in physiological salt solution.

Despite well described phenotypes and distribution in the cell both by immunofluorescence and ChIP-sequencing, the exact molecular mechanism by which HMR influences these phenotypes remains unknown. Simulating the hybrid situation of HMR+LHR overexpression, we have shown reduced condensin binding to chromatin upon these conditions, however the details of how HMR influences it remain to be investigated.

3.3 HMR might be bordering dCenpA chromatin from HP1a chromatin.

Previous ChIP-sequencing experiments revealed a substantial colocalization of HMR peaks with boundary factors (insulator proteins). At a subset of these sites HMR was found to border HP1a-containing chromatin at promoters of actively transcribed genes. It thus posed the question, whether HMR might serve a boundary function between two types of chromatin. The question was addressed by performing HP1a ChIP-sequencing upon HMR knockdown. The experiment did not reveal detectable HP1a spreading, however, importantly, HMR RNAi did not remove all the protein from its binding sites (Gerland et al., 2017). Thus, it will be of future interest to address the question in HMR knockout cells or flies.

Our experiments suggest that HMR domain, visible on immunofluorescent staining, might border centromeric chromatin, defined by dCenpA, from pericentric chromatin, defined by HP1a (Fig. 3.1). The suggestion comes from two sources of evidence: high-resolution microscopy and APEX2 ascorbate peroxidase proximity labeling.
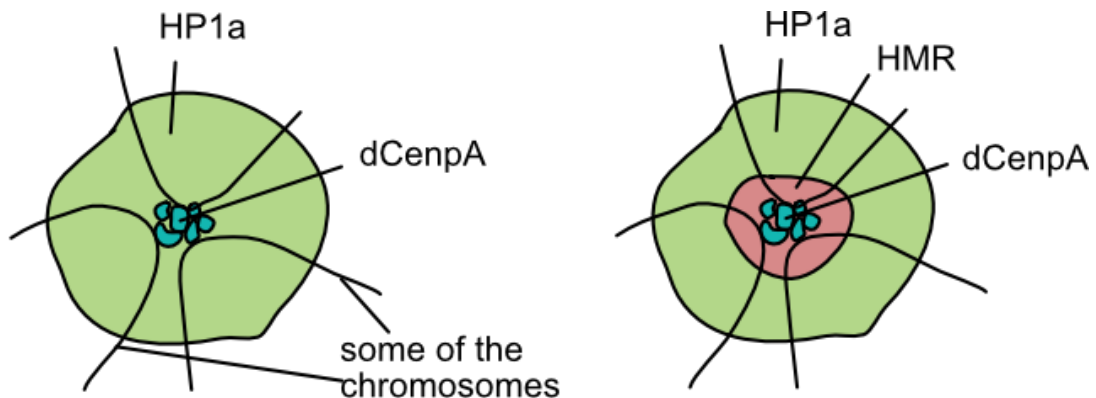
Fig. 3.1. HMR border inside the chromocenter (red, right panel), separating dCenpA from HP1a, is the novelty in our chromocenter model. In the old model dCenpA was thought to be directly embedded in HP1a chromatin (left panel). The concept of the chromocenter is taken from (Jagannathan et al., 2018).

Besides STED microscopy, HMR$_{AP}$ proximity proteome turned out to include both heterochromatic proteins and centromeric proteins, the latter forming a cluster on the HMR$_{AP}$ proximity STRING network (Fig. 2.13). In the network we also found proteins, which indirectly point to centromeric chromatin, such as factors involved in basal and active transcription (Fig. 2.13).

Accordingly, centromeres were reported to be transcribed both during interphase and mitosis (Bobkov et al., 2018; Rosic et al., 2014). Moreover, in GO-term analysis of HMR$_{AP}$ proximity proteome we find proteins associated with both "active" and "inactive" histone modifications.

This evidence in support of HMR localizing to the border between dCenpA and HP1a chromatins is interesting, since such a boundary between centromeric and pericentromeric chromatins was previously reported in fission yeast (Scott et al., 2006), but not in higher eukaryotes (to our knowledge). The existence of such a boundary in Drosophila was proposed by Olszak et al. (Olszak et al., 2011), and localization of HMR to it was suggested by Gerland et al. (Gerland et al., 2017). Given such HMR localization, it is tempting to speculate about its possible boundary function both at genome-binding sites and between centromeric and pericentric chromatins. The latter still remains to be tested in future experiments.

3.4 dCenpA$_{AP}$ and HMR$_{AP}$ proximity proteomes form STRING networks.

We tried to identify clusters, using the STRING tool, in collections of nuclear proteins enriched in dCenpA$_{AP}$, HMR$_{AP}$ and HP1a$_{AP}$, but not APEX$_{NLS}$ pulldowns. HP1a$_{AP}$ pulldowns from all three time points did not reveal many proteins enriched against APEX$_{NLS}$, and gave very few distinguishable STRING clusters.

dCenpA$_{AP}$ and HMR$_{AP}$ pulldowns, in contrast to HP1a$_{AP}$, formed STRING networks with distinguishable clusters. Interestingly, some of the dCenpA$_{AP}$ and HMR$_{AP}$ clusters belonged to the same group of proteins. For example, proteins related to transcription, nucleolus, replication, centromere and nuclear pore.

The finding of proteins involved in transcription in both STRING networks is in agreement with previous investigations, since transcription was reported at centromeres (Bobkov et al., 2018; Erhardt et al., 2008) and HMR was reported to bind at promoters of actively transcribed genes.

Although we found centromeric proteins in proximity to both dCenpA$_{AP}$ and HMR$_{AP}$, these are different centromeric proteins except dCenpC, prod and LHR. This suggests that different centromeric proteins are accessible for biotinylation by different baits, and possibly, belong to different sub-centromeric structures.

Interestingly, some of the clusters were found only in dCenpA$_{AP}$ or in HMR$_{AP}$ network. Proteins involved in RNAi were only enriched in dCenpA$_{AP}$ network. RNAi was shown to be important for heterochromatin formation in *Drosophila* and yeast (Dawe, 2003; Riddle and Elgin, 2008; Yang et al., 2018), and heterochromatin was shown to be important for the centromere formation in yeast. However, RNAi components were not previously (to our knowledge) found at dCenpA chromatin, and our results point to such possibility. Alternatively, RNAi components in heterochromatin were biotinylated because of being close to dCenpA chromatin.

HMR$_{AP}$, but not dCenpA$_{AP}$ pulldowns, included Polycomb proteins, boundary factors (which is consistent with previous report about HMR colocalization with insulator proteins (Gerland et al., 2017)), nucleosome remodelers and cohesin/condensin complex. It is expectable that HMR$_{AP}$ network turned out to be broader than dCenpA$_{AP}$ network, since HMR localizes not only to centromeres, but also to other nuclear structures.

3.5 dCenpC clusters HMR and centromeres.

dCenpC was shown to be a bona fide interactor of both dCenpA (Barth et al., 2014; Erhardt et al., 2008) and HMR (Thomae et al., 2013). Interestingly, we also found dCenpC in proximity to both proteins.

Anti-HMR antibody gives a strong centromeric and very weak heterochromatic staining. We showed that dCenpC RNAi results in only heterochromatic localization of HMR. Thus, dCenpC is required for the architecture of the chromocenter not only by being necessary for dCenpA incorporation, but also by maintaining proper localization of HMR. Since HMR is dispersed on many binding sites in the genome and at the same time gives a strong centromeric staining on immunofluorescence, we hypothesized that

dCenpC clusters HMR binding sites at the centromeres. To test it, we performed ChIP-sequencing of HMR upon dCenpC RNAi, and did not detect prominent change of HMR binding sites. Thus, dCenpC clusters HMR near the centromeric region.

We also detected declustering of centromeres, but not pericentric chromatin upon dCenpC RNAi. This suggests that centromeres' architecture has a limited effect on the architecture of pericentric heterochromatin.

Interestingly, by performing rescue experiments with full length RNAi-resistant dCenpC and protein deletion mutants, we found out that full length 1-1411 and C-terminal 558-1411 proteins rescued HMR and partially rescued centromere clustering, while N-terminal 1-1038 2xNLS protein did not show rescue. Since C-terminal mutant localizes to centromeres and N-terminal not, this suggests that centromeric localization of dCenpC is necessary for both clustering the centromeres and the near-centromeric HMR border (Fig.3.2).
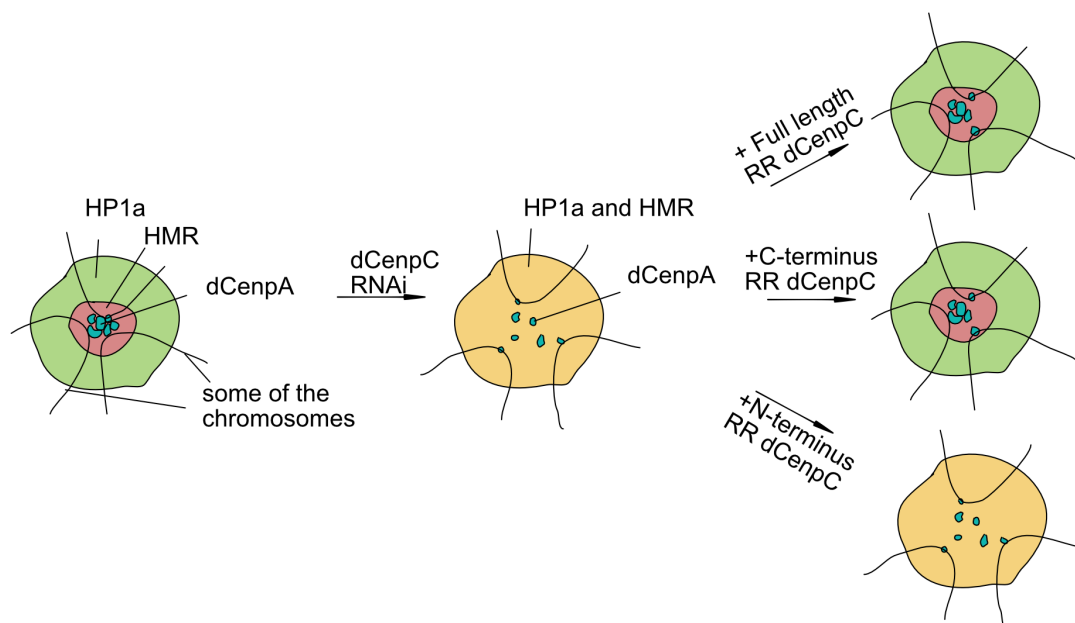
Fig. 3.2. dCenpC RNAi results in dCenpA declustering and diffusion of HMR border into HP1a chromatin. Full length and C-terminal RNAi-resistant (RR) constructs rescue HMR localization and – partially – centromere clustering. N-terminal RR construct does not rescue. The concept of the chromocenter is taken from (Jagannathan et al., 2018).

3.6 HMR association with condensins and cohesins might point to its function in pure species as well as hybrids.

It was striking to find both cohesin and condensin complexes in the STRING network of HMR$_{AP}$ proximity proteome, since phenotypes of HMR knockouts and overexpression were reminiscent of those observed upon perturbation of cohesin and condensin levels (Blum et al., 2017).

To further unravel molecular details of HMR colocalization with cohesins and condensins, we compared the available ChIP-sequencing profiles of cohesin subunit Rad21, condensin subunit CAP-H2 and HMR (Gerland et al., 2017; Van Bortle et al., 2014). Interestingly, we found HMR colocalization both with condensin and cohesin. Furthermore, this colocalization enhanced upon HMR and LHR overexpression (as determined by analysis of HMR ChIP-sequencing at these conditions (Cooper et al., 2019)).

Interestingly, bioinformatic analysis showed that CAP-H2/HMR binding sites are enriched at promoters and active or HP1a types of chromatin. This is in agreement with the fact, that when HMR is not overexpressed, condensin subunit CAP-H2 localizes predominantly to HMR class I binding sites, where HMR is bordering HP1a at promoters of actively transcribed genes. Notably, upon HMR overexpression CAP-H2 colocalizes also with other class of HMR binding sites. Rad21 binding, in contrast, does not discriminate between native HMR class I and class II binding sites, the latter of which represent *gypsy* insulators defined by the presence of Cp190, Su(Hw) and mod(mdg4) (reviewed in (Gerland et al., 2017)).

Recently, HMR upon its overexpression with LHR was shown to localize to the binding sites of another speciation protein, GFZF. More than 20% of overexpressed HMR-binding sites are also bound by GFZF (Cooper et al., 2019). It is thus tempting to speculate, that when HMR is overexpressed, HMR/GFZF and HMR/condensin / HMR/cohesin binding sites colocalize, and combination of HMR, GFZF and condensins/cohesins leads to hybrid lethality. However, this hypothesis remains to be tested both by bioinformatic analysis and cytological/mutational studies in flies.

3.7 HMR association with TAD boundaries might point towards its role in nuclear organization.

Being architectural proteins, condensins and cohesins were shown to cluster at TAD boundaries (Van Bortle et al., 2014). We thus hypothesized that HMR also resides there, and built a heat map of ChIP-sequencing signals and TAD separation score (Ramirez et al., 2018). We found that HMR indeed often colocalizes with TAD boundaries. Similar to colocalization with cohesins and condensins, this colocalization enhances when HMR is overexpressed.

It is noteworthy to mention, that there is an obvious discrepancy between hundreds of HMR-binding sites all over the genome and very focused immunofluorescent staining near the centromere. This discrepancy might be explained by two scenarios. In the first, HMR sites detected in ChIP-sequencing are distributed all over the nucleus and not detected by immunofluorescence since they are not focused together. Sites bound near the centromere, in turn, might be difficult to map in ChIP-sequencing experiments since they might be repetitive. In the second scenario, some of genome-wide binding sites of HMR at boundary elements, including TAD boundaries, cluster together in 3D to form a border between dCenpA and HP1a chromatin (Fig. 3.3).
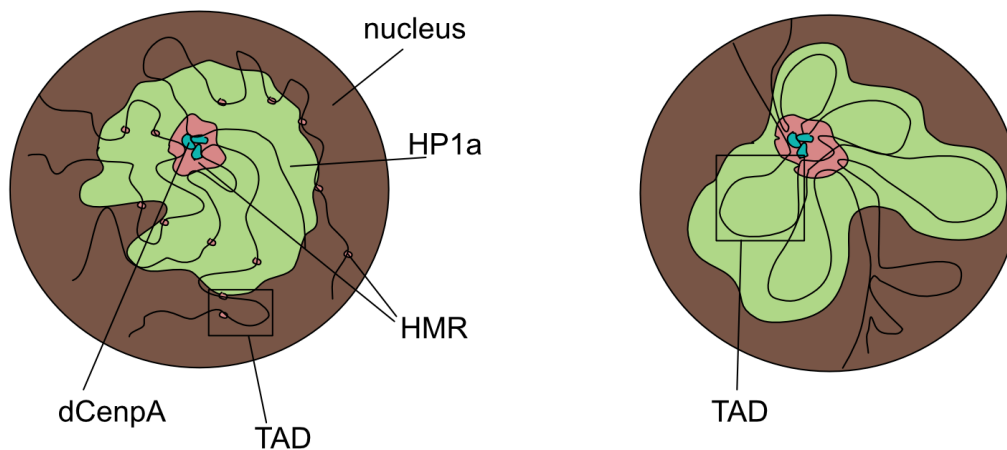
Fig. 3.3. Two models of HMR-occupied TAD boundaries positioning. Left panel: They are distributed all over the nucleus and not detected by immunofluorescence. HMR sites at the chromocenter are repetitive elements not mapped in ChIP-seq experiments. Right panel: TAD boundaries occupied by HMR are clustered at the chromocenter.

Quite some results already point to the role of cohesins in the formation of TADs, which is most probably happening by loop extrusion (reviewed in (Fudenberg et al., 2018)). It might be tempting to speculate that HMR, important for chromosome cohesion, might also contribute to the formation of TADs and 3D structure of the nucleus. In fission yeast, localization of tRNA border genes near the centromere is dependent on condensin (Iwasaki et al., 2010). It thus was logical to investigate, whether condensins/cohesins were important for HMR domain formation between dCenpA and HP1a domains, as well as chromocenter architecture and centromere clustering. However, cohesin and condensin knockdown did not affect HMR localization to centromere and did not have a dramatic effect on clustering of centromeres.

3.8 HMR+LHR overexpression reduces condensin binding to chromatin.

Native HMR ChIP signal and CAP-H2 signal seemed to be mutually exclusive at class II binding sites. Moreover, there are possible chromosome condensation defects in hybrid flies, where HMR+LHR were shown to be overexpressed. We thus wondered what happens to condensin binding sites upon HMR+LHR overexpression.

We addressed the question by ChIP-sequencing of CAP-H2 upon these conditions and found that HMR+LHR overexpression reduced condensin binding to chromatin. Interestingly, this happened at both HMR-bound and – unbound sites, but the unbound ones at TAD boundaries on average contacted the bound ones. This suggests, that HMR effect on CAP-H2 might be direct. For example, HMR could destabilize condensin at certain high concentrations of HMR molecules. It is also possible that HMR and condensin are mutually exclusive at particular context, e.g. class II binding sites. It is unlikely that these proteins are completely mutually exclusive, because in this scenario we would not have found condensin in proximity to HMR_{AP}.

Since condensin was shown to be important for the structure of mitotic chromosomes (reviewed in (Skibbens, 2019)), it is logical that HMR+LHR overexpression was shown to lead to mitotic defects (Thomae et al., 2013).

It is tempting to speculate, that the same reduction of condensin binding happens in hybrids, where HMR and LHR are overexpressed. This, in turn, might lead to observed chromosome condensation defects and hybrid lethality.

3.9 Implication of proximity labeling results to understanding of speciation.

Apart from cohesin and condensin, we found several proteins in proximity to HMR, which were reported to play a role in speciation or which could potentially do it.

Speciation proteins LHR and GFZF reside in proximity to HMR, and it would be interesting to perform proximity biotinylation with LHR and GFZF APEX2 fusions. An overlap between proximities of three speciation proteins could potentially contain other so far unknown speciation factors.

It was also interesting to find prod in proximity to HMR_{AP}. Since prod binds to *D. melanogaster* specific repeat {AATAACATAG}$_n$ and plays a role in chromocenter formation (Jagannathan et al., 2019), it might be that its binding in *D. melanogaster/D. simulans* hybrids is compromised, which might in turn result in problems with hybrid chromocenter formation.

Finally, it would be worth to mention, that Hmr2 mutant, which saves hybrids, has an impaired centromeric localization (Aruna et al., 2009; Thomae et al., 2013). It is tempting to speculate, that centromeric localization of HMR in hybrids leads to their infertility and lethality.

Interestingly, some nuclear membraneless organelles, for example heterochromatin, have been shown to form by phase separation (Larson et al., 2017; Strom et al., 2017). It would be tempting to speculate, that parts of centromere, as membraneless organelles, also form by phase separation. Proximity labeling might be an important tool in studying membraneless organelles, since it captures not only interactions, but also the proximity of the bait. It would be tempting to suggest, that problems in *D. menanogaster/D. simulans* hybrids result from impaired formation of centromeric HMR phase-separated domain. Investigating the proximity of HMR therefore might help in understanding the biology of HMR phase-separated organelle. It would be of

future interest to test the hypothesis about liquid droplets' formation impairment in fly hybrids.

3.10 Proximity labeling requires individual conditions' adaptation for every protein.

APEX2 proximity labeling technology has for the first time, to our knowledge, been applied to chromatin proteins in *Drosophila* tissue culture cells. It is noteworthy to mention, that different proteins in the nucleus have different expression levels and different biophysical characteristics. Thus, respective APEX2 fusions might require different conditions of biotinylation.

This indeed turned out to be true, since we had to apply different biotin-phenol concentrations for different protein fusions (5 mM for HMR_{AP} and $dCenpA_{AP}$ in contrast to 0.5 mM for $HP1a_{AP}$ and $APEX_{NLS}$), as well as carefully titrate the labeling time (for $HP1a_{AP}$). This points to the fact, that the technique might not always be suitable for the high-throughput studies, and requires careful "calibration" for every potential protein.

3.11 Validation and specificity of proximity labeling.

We validated the usability of proximity labeling technique. Firstly, we performed immunofluorescence of 4 factors found in proximity to $HP1a_{AP}$ and HMR_{AP}. Secondly, we measured distributions of median distances from HMR/HP1a peaks to peaks of proteins found in proximity and anti-proximity of HMR_{AP} and $HP1a_{AP}$. Proteins in proximity to HMR_{AP} and $HP1a_{AP}$ turned out to locate closer to HMR and HP1a than proteins in anti-proximity. An important part of validation was unraveling the functional link between HMR+LHR overexpression and reduced CAP-H2 binding to chromatin. Condensin is found only in proximity to HMR, but not in HMR IP. Thus, APEX2 labeling technique can point to functional links not captured by conventional IP-MS.

The APEX2 labeling technique demonstrated, that specificity, and thus, the quality of labeling might depend on individual protein fusions, in particular on their biology. HMR protein has several DNA-binding domains, and thus most

probably tightly associates with chromatin. HP1a, in contrast, has not only chromatin-bound, but also highly mobile fraction floating around the nucleus (reviewed in (Straub, 2003)). It thus seems logical, that dCenpA$_{AP}$ and HMR$_{AP}$ proteomes turned out to be rather specific in contrast to a lot of background in HP1a$_{AP}$ labeling.

One more reason why HP1a$_{AP}$ proteome after 25 minutes of labeling turned out to be less specific is mislocalization of HP1a$_{AP}$ fusion. This mislocalization is not due to the effect of H_2O_2 on heterochromatin integrity, since in unlabeled cells HP1a$_{AP}$ still forms a focused domain.

To summarize, in this thesis my collaborators and I significantly improved our understanding of the structure of the chromocenter, using STED microscopy and APEX2 proximity labeling. We validated our proteomic approach and performed follow-up studies on dCenpC and condensin, found in proximity to HMR/dCenpA and HMR respectively. We defined the role of dCenpC as an architectural protein of the chromocenter. Moreover, we pointed to the molecular mechanisms by which HMR might cause problems with chromosome condensation in hybrids.

4 Materials and Methods.

4.1 Materials.

4.1.1 Constructs.

Construct	Comment	Source
Flag-HA-APEX-2xNLS/pMT Hygro	cloned	this study
Myc-HMR-APEX/pMT Hygro	cloned	this study
Myc-dCenpA-APEX/pMT Hygro	cloned	this study
Flag-HA-HP1a-APEX/pMT Hygro	cloned	this study
GST-APEX/pGEX-6P-1	cloned	this study
Flag-HA-His-SUMO/pMT Hygro	vector used for cloning	unpublished construct from Imhof lab
Myc-Dmel HMR/pMT Hygro	vector used for cloning	Thomae et al, 2013
GST-HMR2-233/pGEX-6P-1	vector used for cloning	unpublished construct from Imhof lab
Flag-HA-HP5/pMT Hygro	cloned	this study, by Dr. Andreas W. Thomae
Flag-HA-ADD1/pMT Hygro	cloned	this study, by Dr. Andreas W. Thomae
Flag-HA-XNP/pMT Hygro	cloned	this study, by Dr. Andreas W. Thomae
Flag-HA-CG8108/pMT Hygro	cloned	this study, by Dr. Andreas W. Thomae
Flag-HA RR dCenpC41H979E/pMT Hygro	cloned	this study
Flag-HA RR dCenpC 1-1038 2xNLS/pMT Hygro	cloned	this study
Flag-HA RR dCenpC 558-end/pMT Hygro	cloned	this study

Table 1. Constructs used for cloning and cloned in this study.

4.1.2 Primers.

Primer	Sequence	Contruct
NotI_APEX_forw2	TACGCCGGCGGCCGCAAGGGAAAGTCTTAC	Flag-HA-APEX-NLS/pMT Hygro
APEX-2xNLSrev3.1	tcttcttcggtccaccgccgaccttccgcttcttcttcggGGCATCAGCAAACCC	Flag-HA-APEX-NLS/pMT Hygro
APEX-2xNLSrev3.2	GGATCCTCTAGATCAgaccttccgcttcttcttcggtcca	Flag-HA-APEX-NLS/pMT Hygro
MycNotI_Hmr_fw	AGGATCTGGGCGGCCGCGAGGAGGAGCCTGTTGC	Myc-HMR-APEX/pMT Hygro
endHmr_APEX_fw	AATCCGCCACCGCCTAAGGGAAAGTCTTAC	Myc-HMR-APEX/pMT Hygro
stop-APEXnew	CACCGGATCCTCTAGATCAGGCATCAGC	Myc-HMR-APEX/pMT Hygro; Flag-HA-HP1a-APEX/pMT Hygro; Myc-dCenpA-APEX/pMT Hygro
MycNotI_CID_fw	AGGATCTGGGCGGCCGCCCACGACACAGCAGA	Myc-dCenpA-APEX/pMT Hygro
endCID-APEX_fw	CGGGGTCGGCAATTTAAGGGAAAGTCTTAC	Myc-dCenpA-APEX/pMT Hygro
HP1-APEX-fw	TCTGATAATGAAGATGACTACAAGGATGAC	Flag-HA-HP1a-APEX/pMT Hygro
APEX-HP1rv-new	TCTGATAATGAAGATAAGGGAAAGTCTTAC	Flag-HA-HP1a-APEX/pMT Hygro
Not-HP1-fw	ATTACGCCGGCGGCCGCGGCAAGAAAATC	Flag-HA-HP1a-APEX/pMT Hygro
EcoR1.APEX.fw.new	GATCCCCGGAATTCAAGGGAAAGTCTTACCCAACTGTGAG	GST-APEX/pGEX-6P-1
Not1.APEX.rev.new	TCACGATGCGGCCGCTTAGGCATCAGCAAAC	GST-APEX/pGEX-6P-1
NotI_CenpC_fw2	TACGCCGGCGGCCGCATGTCGAAGCCCC	RNAi-resistant Flag-HA-CenpC/pMT Hygro; RNAi-

		resistant Flag-HA-CenpC 1-1038 2xNLS/pMT Hygro
CenpC fr1_rev2	GCGCATCATAAAGGC	RNAi-resistant Flag-HA-CenpC/pMT Hygro
CenpC fr1-overh_fr2_fw2	GCCTTTATGATGCGCAAACTGGCTGAGAAC	RNAi-resistant Flag-HA-CenpC/pMT Hygro
CenpC fr2_rev2	CTTTTCGGTACAGGG	RNAi-resistant Flag-HA-CenpC/pMT Hygro
CenpC fr2-overh_fr3_fw2	CCCTGTACCGAAAAGCAAAAGAGGAAGTTGC	RNAi-resistant Flag-HA-CenpC/pMT Hygro
CenpC fr3_rev	TTCGGAATGCGG	RNAi-resistant Flag-HA-CenpC/pMT Hygro
CenpC fr3-overh_fr4_fw2	TACCGCATTCCGAAAGCCTGGGATTGAG	RNAi-resistant Flag-HA-CenpC/pMT Hygro
CenpC fr4_rev	CTTGGCCTGCTTC	RNAi-resistant Flag-HA-CenpC/pMT Hygro
CenpC fr4-overh_fr5_fw2	GAAGCAGGCCAAGGTCCATCCACTTAAAC	RNAi-resistant Flag-HA-CenpC/pMT Hygro
XbaI_CenpC_rev2	GGATCCTCTAGACTAACTGCGTATACAC	RNAi-resistant Flag-HA-CenpC/pMT Hygro; RNAi-resistant Flag-HA-CenpC 558-end/pMT Hygro
CenpC_DM_fw	GCGCGGCCGCCCCAAAAAAGCCGTGGGCGG	RNAi-resistant Flag-HA-CenpC/pMT Hygro
CenpC_DM_rev	CCGCCCACGGCTTTTTTGGGGCGGCCGCGC	RNAi-resistant Flag-HA-CenpC/pMT Hygro
NLS_RRCenpC_1038rev	tcttcttcggtccaccgccgaccttccgcttcttcttcggACGCTGCAAAAACTC	RNAi-resistant Flag-HA-CenpC 1-1038 2xNLS/pMT Hygro
NotI_RRCenpC_558_fw	TACGCCGGCGGCCGCATGCTACGTAGAAATCTAATG	RNAi-resistant Flag-HA-CenpC 558-end/pMT Hygro

Table 2. Primers used for cloning in this study.

Primer name	dsRNA	Source	Primer sequence 5'->3'
Cenp-C_1_fw	Cenp-C_1	This study	TTAATACGACTCACTATAGGGAGACAAGCTTGCCGAAAATAAGCCGG
Cenp-C_1_rev	Cenp-C_1	This study	TTAATACGACTCACTATAGGGAGATTTCTCTGTGCAAGGTGTGCTGCTTATTTC
Cenp-C_2_fw	Cenp-C_2	This study	TTAATACGACTCACTATAGGGAGAATCCCTTGGCCTGAGTACCTTGACGTG
Cenp-C_2_rev	Cenp-C_2	This study	TTAATACGACTCACTATAGGGAGATTCGCTTGTTTCATGCTACGTTTTTGGTATG
GST_fw	GST	Thomae et al, 2013	TTAATACGACTCACTATAGGGAGAAGTTTGAATTGGGTTTGGAGTTTCC
GST_rev	GST	Thomae et al, 2013	TTAATACGACTCACTATAGGGAGAGGATGGTCGCCACCACCAAACGTGG
vtd.RNAi.fw	vtd	DRSC20839	TTAATACGACTCACTATAGGGAGATGGAAAGAAACTGGAGGTGTC
vtd.RNAi.rev	vtd	DRSC20839	TTAATACGACTCACTATAGGGAGATCGTCACCCATTTCATGATT
CAP-H2.RNAi.fw	CAP-H2	DRSC14908	TTAATACGACTCACTATAGGGAGAGAGCACATGACCACAAAGGA
CAP-H2.RNAi.rev	CAP-H2	DRSC14908	TTAATACGACTCACTATAGGGAGATGCATTTGAATATCGGAAAGC
SMC1.RNAi.fw	SMC1	DRSC16846	TTAATACGACTCACTATAGGGAGAAGCAAATGCTGGAAGTGGAA
SMC1.RNAi.rev	SMC1	DRSC16846	TTAATACGACTCACTATAGGGAGAGACTCCAAATCGACCATACT
SMC2.RNAi.fw	SMC2	DRSC07544	TTAATACGACTCACTATAGGGAGATCCTTAATCGCCTGTTCGAG
SMC2.RNAi.rev	SMC2	DRSC07544	TTAATACGACTCACTATAGGGAGATCGCGTTCAACAAAATGAAG

Table 3. Primers used for dsRNA generation in this study.

Target mRNA	Primer name	Primer sequence 5'->3'	Source
vtd	vtd_fw1	GGAAATATTGGCGAGATGGA	This study
vtd	vtd_rev1	CCTTTTTGAAACGCCATTGT	This study
CAP-H2	CAP-H2_fw1	GCGGCAAGATCTATGGAGAC	This study
CAP-H2	CAP-H2_rev1	CTAGGGGTCTCCTTCTGCAA	This study
SMC1	SMC1_fw1	GTCCTACCGCGGTCACATAG	This study
SMC1	SMC1_rev1	TCACGAAACTGATGGCATCC	This study
SMC2	SMC2_fw3	TCAAAACAAAGTGGGCGCC	This study
SMC2	SMC2_rev3	ACTTCATGACAGGCTCGTAA	This study
RpL32	ONTG233 RPL32 RT F	GTTCGATCCGTAACCGATGT	This study, thesis T. Gerland, 2017
RpL32	ONTG234 RPL32 RT B	CCAGTCGGATCGATATGCTAA	This study, thesis T. Gerland, 2017
Hmr	ON223 Hmr RT fw 160	AATCGCTTGCGAAGAACACT	This study, thesis T. Gerland, 2017
Hmr	ON224 Hmr RT rev 160	ACTGGCCGTGGACAAGTTAC	This study, thesis T. Gerland, 2017

Table 4. Primers used for RT-qPCR in this study. All primers were designed with Primer3.

4.1.3 Cell lines.

Cell line	Selection
Myc-dCenpA-APEX/pMT Hygro	100 ug/ml Hygromycin
Myc-dCenpA-APEX/pMT Hygro clone 8	100 ug/ml Hygromycin
Myc-Hmr-APEX/pMT Hygro	100 ug/ml Hygromycin
Flag-HA-HP1a-APEX/pMT Hygro	100 ug/ml Hygromycin
Flag-HA-HP1a-APEX/pMT Hygro clone 29	100 ug/ml Hygromycin

Flag-HA-APEX-2xNLS/pMT Hygro	100 ug/ml Hygromycin
Flag-HA-HP5/pMT Hygro	100 ug/ml Hygromycin
Flag-HA-ADD1-PA/pMT Hygro	100 ug/ml Hygromycin
Flag-HA-XNP/pMT Hygro	100 ug/ml Hygromycin
Flag-HA-CG8108/pMT Hygro	100 ug/ml Hygromycin

Table 5. Cell lines, generated and used in this study.

4.1.4 Antibodies.

Primary antibody	Source and (optionally) comments	Dilution
rat anti-HMR 2C10	Helmgoltz Zentrum, Thomae et al, 2013	WB, IF 1:25, STED 1:5, ChIP 1 ml/IP
rat anti-APEX 20H10	Helmgoltz Zentrum, raised in this study	WB, IF 1:50
rat anti-dCenpA 7A2	Helmgoltz Zentrum	IF 1:100, STED 1:50
rabbit anti-dCenpA	Actif Motif	IF 1:500, STED 1:250
mouse anti-HP1a 1:100	kind gift from Sarah Elgin	WB, IF 1:100
rabbit anti-dCenpC	kind gift from Christian Lehner	IF 1:5000, STED 1:1000
mouse anti-FLAG	Sigma M2, 1mg/ml	IF 1:100
rabbit anti-BubR1	kind gift from Claudio Sunkel	IF 1:1000
Streptavidin Alexa 555	Thermo Scientific	IF 1:400
rabbit anti-D1	kind gift from Harmit Malik	IF 1:500
rabbit anti-CAP-H2	kind gift from Giovanni Bosco	ChIP 6 ul/IP
mouse anti-HA 12CA5	Helmgoltz Zentrum	IF 1:1000
rat anti-HA 3F10	Helmgoltz Zentrum	IF 1:100
rabbit anti-GFZF	kind gift from uNitin Phadnis	WB 1:500
Anti-rat bridging antibody	Dianova	ChIP 6 ul/IP
Secondary antibody	**Source and (optionally) concentration**	**Dilution**
sheep anti-mouse IgG HRP Linked Whole Ab	GE Healthcare	WB 1:10000

donkey anti-rabbit IgG HRP Linked Whole Ab	GE Helthcare	WB 1:10000
goat anti-rat IgG HRP Linked Whole Ab	GE Helthcare	WB 1:10000
donkey anti-rabbit Alexa 488	Jackson ImmunoResearch	IF 1:500
donkey anti-rabbit Alexa 647	Jackson ImmunoResearch, 0,75 mg/ml in 50% glycerol, preabsorbed	IF 1:300
donkey anti-rat Alexa 488	Jackson ImmunoResearch	IF 1:300
donkey anti-mouse Alexa 488	Jackson ImmunoResearch, very highly preabsorbed	IF 1:300
donkey anti-rat Cy3	Jackson ImmunoResearch, preabsorbed	IF 1:800
donkey anti-rat Alexa 594	Thermo Scientific	STED 1:300
goat anti-rabbit Abberior STAR 635P	Sigma	STED 1:300

Table 6. Antibodies, used in this study.

4.1.5 Reagents.

Reagent	Company
1 kb DNA ladder	NEB
100 bp DNA ladder	NEB
2-iodoacetamide	Merck
Acetonitrile, HPLC grade	J.T. Baker®
Agarose Universal	Bio&SELL
Ampicillin	Roth
Aprotinin	Genaxxon bioscience
Benzonase	Merck
beta-Mercaptoethanol	Sigma
Biotin-phenol	Iris Biotech
Bromophenol Blue	Sigma

CaCl2	Calbiochem
Clarity™ Western ECL Substrate	Biorad
Colchicine ≥95% (HPLC)	Sigma
CpG2006	TIB MOLBIOL
DAPI	Life Technologies
Deoxynucleotide (dNTP) Solution Mix	NEB
Dimethylsulfoxide	Sigma
DNase I recombinant, RNase-free	Roche
DTT	Roth
EcoRI-HF	NEB
EcoRV-HF	NEB
EDTA	AppliChem
EGTA	Roth
Ethanol	Sigma
Ethidium bromide	Merck
Fetal Bovine Serum Low in Endotoxin A. H	Sigma
Formaldehyde, 37% (w/v) solution	Sigma
Formic acid, 98-100%	Merck
Gel Loading Dye, Purple (6X)	NEB
Glycerol	AppliChem
Glycin	Merck
H2O2 30%	Sigma
HCl	VWR
HEPES	Serva
HindIII-HF	NEB
Hygromycin B in PBS 50mg/ml	Invitrogen
Image-iT FX signal enhancer	Invitrogen
KCl	AppliChem
KH2PO4	Merck
LB-Agar-Pulver	Diagonal
LB-Medium Pulver	Diagonal
Leupeptin	Genaxxon bioscience

Lysyl Endopeptidase®, Mass Spectrometry Grade	Wako
Methanol	VWR
Methanol, HPLC grade	Roth
MG132	Enzo Life Sciences
MgCl2	VWR
Micrococcal Nuclease	Sigma
Midori Green Direct	NIPPON Genetics
Na2HPO4	Merck
NaCl	neoFroxx
NaOH	neoFroxx
NH4Ac	Sigma
Non-fat dry milk	Heirler
Normal Goat Serum	Dianova
NotI-HF	NEB
NP-40	Fluka
Penicillin-Streptomycin	Sigma
Pepstatin	Genaxxon bioscience
Phusion High-Fidelity DNA Polymerase	NEB
PMSF	Sigma
Ponceau S solution	Sigma
PowerUp™ SYBR™ Green Master Mix	Applied Biosystems™
ProLong™ Diamond Antifade	Thermo Scientific
Protein Marker V	Serva
Proteinase K	Roche
RNAse A	Sigma
Schneider Drosophila medium	Life Technologies
SDS	Serva
Sequencing Grade Modified Trypsin, Lyophilized	Promega
Sodium ascorbate	Sigma
Sodium azide	Merck
Sodium Deoxycholate	Sigma
Taq DNA Polymerase with ThermoPol® Buffer	NEB

TBE 5x	VWR
Trifluoroacetic acid, LC-MS grade	Thermo Scientific
Tris	neoFroxx
Triton-X-100	Sigma
Trolox	Sigma
Tween 20	Sigma
Urea	Roth
Vectaschield	Vector Labs
Water, HPLC grade	VWR
X-tremeGENE™ HP DNA Transfection Reagent	Roche
XbaI	NEB

Table 7. Reagents, used in this study.

4.1.6 Kits.

Kit	Company
In-Fusion® HD Cloning Kit	Clontech
QIAprep Spin Miniprep Kit	QUIAGEN
NucleoSpin Plasmid EasyPure	Macherey Nagel
MEGAscript RNA kit	Invitrogen
SuperScript™ III First-Strand Synthesis System	Invitrogen
RNAeasy mini kit	QUIAGEN
NucleoSpin Gel and PCR Clean-up	Macherey Nagel
QIAquick PCR Purification Kit	QUIAGEN
MicroPlex Library Preparation Kit 12 indexes	Diagenode

Table 8. Kits, used in this study.

4.1.7 Consumables and devices

Consumable/device	Company

1.5 ml eppendorfs	Greiner, Sarstedt
18 mm 3.5 kD MWCO dialysis membrane	Spectra/Por
2.0 ml tubes	Sarstedt
384 Well Lightcycler Plate	Sarstedt
384 Well Lightcycler Plate Sealing Tape, optically clear	Sarstedt
6 well cell culture plate	Sarstedt
96 well cell culture plate	Sarstedt
Adaptors for C18 and HILIC columns	Glygen
AFA Tubes (Tubes for Covaris S220 instrument)	Covaris S-Series Tube & Cap 12 x 24 mm
AMPure XP beads	Beckman Coulter
Braun S Pestle 5 ml	B. Braun
C18 solid phase extraction disk	Empore
CELLSTAR® Cell Culture Dishes 15 cm	Greiner
Column 120 x 0.075 mm, in house packed with Reprosil-C18, 2.4 µm	Dr. Maisch GmbH
Corning® 250mL polypropylene (PP) centrifuge tubes	Corning
Coverslips 12mm in diameter	Sigma
Dialysis clamps	Spectra/Por
Disposable cell scrapers	Sarstedt
Falcon tubes 15 ml	Sarstedt
Falcon Tubes 50 ml	Sarstedt
Filter papers	Whatman
Forceps Dumont 110mm K342.1	Roth
Inoculation loops	Sarstedt
Microscope Slides	Roth
Nail polish	Essence
Nitrocellulose membrane	Amersham
Parafilm	Brand PARAFILM
PCR tubes	Greiner
Pierce Streptavidin Magnetic Beads	Thermo Scientific
Pipette boy RF3000	Heathrow Scientific

Pipettes	Gilson
Precast gels	Serva
Protein LowBind Tubes 1.5 ml	Eppendorf
Roller bottles	Greiner
Sepharose Protein A beads	GE Healthcare
Sepharose Proteins G beads	GE Healthcare
Small petridishes	Sarstedt
T175 flasks	Greiner
T25 flasks	Greiner
T75 flasks	Greiner
TopTip PolyHydroxyethyl A (HILIC) 1-10 ul	Glygen
Tubes 1.5 mL, DNase-, Rnase free	Biozym
Western blotting chambers	Li-Cor

Table 9. Consumables and devices, used in this study.

4.1.8 Technical devices

Technical device	Company and model
-20 °C freezer	Miele, Liebherr
-80 °C freezer	GFL
26 °C incubator	LMS
26 °C roller bottles incubator	Bellco-Tecnomara
4 °C refrigerator	Liebherr
Big centrifuge	Thermo Scientific, Heraeus Multifuge X3R
BlueLight Table	Serva
Cell counter	OMNI Life Science, CASY
ChemiDoc™ Imaging System	BioRad
Concocal microscope	Leica, TCS SP5
Confocal and STED microscope	Leica, SP8X WLL
DNA electrophoresis chambers	University workshop
Electrophoresis power supply	PQ Lab, EV243

Epifluorescence microscope	Zeiss, Axiovert 200
Ice machine	Ziegra
Incubator shaker	Innova, 42
Laminar flow hood	BERNER FlowSafe
Liquid nitrogen tank for cells	Thermo Scientific, 7403
Magnetic stirrers	Bachofer Ika-Combimag Reo
Mass spectrometer	Thermo Scientific, Q-Exactive
Microwave	SEVERIN
PCR machine	Applied Biosystems, 2720 thermal cycler
pH-meter	inoLab, pH 720
Protein electrophoresis chambers	Serva, bv 104
Quantitative Real-Time PCR instrument	Roche LightCycler 480 II
Rotators	NeoLab
Scales	Sartorius, TE 153S
Scales	KERN, ABJ-NM/ABS-N
Scales	KERN, PCB
Shaker	NeoLab, DOT10L
Shaker and thermomixer	Eppendorf, comfort
Sonicator	Covaris S220 Focused-ultrasonicator
Table top centrifuge	Eppendorf, 5424
Table top centrifuge	Eppendorf, MiniSpin
Table top centrifuge	Eppendorf, 5430R
Table top centrifuge	Eppendorf, 5804R
Thermomixer	Eppendorf, MTP
Tissue culture centrifuge	Hettich, Rotanta 460
Vacuum centrifuge	ScanVac, Scan Speed 40
Water bath	B. Braun Thermomix 1420
Western blotting chambers with accessories	Biorad

Table 10. Technical devices, used in this study.

4.1.9 Software

Application	Software
Creating STRING networks	STRING database (https://string-db.org)
Creating STRING networks	Cytoscape
Genome-wide data analysis	R studio
Genome-wide data analysis	bowtie2 (version 2.2.9)
Genome-wide data analysis	samtools (version 1.3.1)
Genome-wide data analysis	Homer (version 4.9)
GO-term analysis	GO consortium (geneontology.org/)
Hi-C data analysis	HiC-Pro (version 2.9.0)
Image analysis	Adobe Illustrator
Image analysis	Adobe affinity designer
Image analysis	Adobe Photoshop
Image analysis	Bio-Rad Image Lab
Microscopic image analysis	ImageJ
Microscopic image analysis	Huygens 17.10 p2
Microscopic image analysis	Leica Application Suite X
Office tools	Microsoft Word
Office tools	Microsoft PowerPoint
Office tools	Microsoft Excel
Primer design	Primer3 (Rozen and Skaletsky, 2000)
Raw proteomics data processing	MaxQuant version 1.5.3.12
RT-qPCR	Roche LightCycler 480 SW 1.5
Working with sequences	Softonic Serial Cloner 2-6-1

Table 11. Software, used in this study.

4.1.10 GSE of data used for ChIP distance to peak analysis.

GEO	Name	Proximity	Comparison
GSE101554	Ez	proximal	HMR
GSE101554	M1BP	proximal	HMR
GSE102339	Psc_WT	proximal	HMR
GSE105009	GFZF	proximal	HMR
GSE109384	Wapl_BG3	proximal	HMR
GSE116806	Upf1	proximal	HMR
GSE118699	CLAMP	proximal	HMR
GSE23537	Trl_Kc	proximal	HMR
GSE23537	Ttk_Kc	proximal	HMR
GSE27078	LID	proximal	HMR
GSE29206	l3mbt	proximal	HMR
GSE30820	Ash1C	proximal	HMR
GSE33546	jarid2	proximal	HMR
GSE37864	MOF	proximal	HMR
GSE37864	MSL1	proximal	HMR
GSE41440	Lpt	proximal	HMR
GSE47250	Suvar37	proximal	HMR
GSE47263	Chro	proximal	HMR
GSE47298	MBDR2	proximal	HMR
GSE47330	KDM4A	proximal	HMR
GSE49102	row	proximal	HMR
GSE51989	Smc3	proximal	HMR
GSE54529	Rad21	proximal	HMR
GSE56101	HIPP1	proximal	HMR
GSE66183	Scm	proximal	HMR
GSE76997	Pita	proximal	HMR
GSE80700	Ibf2	proximal	HMR

GSE85741	Taf2	proximal	HMR
GSE93828	EPc	proximal	HMR
GSE100613	TFIIB	control	HMR
GSE102043	Rbf	control	HMR
GSE114092	pnt	control	HMR
GSE118484	MED30	control	HMR
GSE19025	HSF	control	HMR
GSE28065	MCM	control	HMR
GSE33546	Suz12	control	HMR
GSE39393	Dp1	control	HMR
GSE40797	Shep	control	HMR
GSE41950	Rrp40	control	HMR
GSE47294	HP1b	control	HMR
GSE60428	DSP1	control	HMR
GSE83959	Lark	control	HMR
GSE87022	MLF	control	HMR
GSE92383	Mago	control	HMR
ENCSR637GDK	Suvar_2_10	Proximal	HP1a
GSE56101	ADD1	Proximal	HP1a
GSE56101	HIPP1	Proximal	HP1a
GSE50364	CG8478	Control	HP1a
GSE19025	HSF	Control	HP1a
GSE37864	MLE	Control	HP1a
GSE39393	Dp1	Control	HP1a
GSE66639	Rrp6	Control	HP1a
GSE83959	Lark	Control	HP1a
GSE89459	Trr	Control	HP1a

Table 12. List of GSE numbers, used for ChIP distance to peak analysis.

4.2 Methods.

Many methods used were published in (Kochanova et al., 2018). Methods cited from the preprint are indicated in italic.

4.2.1 Cloning of APEX containing constructs.

dCenpA, HP1a, Hmr and APEX2 genes were amplified with PCR in a way, that all bait genes had overhangs overlapping with pMT vector on 5' and with APEX2 on 3', and APEX2 gene had an overhang overlapping with pMT vector on 3'. Vectors were linearized with XbaI and NotI-HF restriction, and the new constructs were assembled with In-Fusion kit from required vector and required fragments of bait and APEX2. The product of In-Fusion reaction was transformed into competent bacteria, and colonies from the plate were expanded the next day into separate 5 ml of LB medium. Plasmids from grown bacteria were purified with Miniprep kit and checked by HindIII-HF restriction. Plasmids that gave a needed restriction pattern were sequenced with Eurofins.

GST-APEX2 construct in pGEX-6P-1 vector was cloned the same way, with the difference that the vector was cut with EcoRI and NotI, and the control restriction was performed with EcoRI-HF and EcoRV-HF.

RNAi-resistant dCenpC construct was cloned the same way, being assembled from 5 fragments, 2 of which were designed to be RNAi-resistant with Dr. Tamas Schauer SeqMixer App (https://tschauer.shinyapps.io/SeqMixer/) and were synthesized by Eurofins. Deletion Mutants were amplified by PCR from full length construct PCR product and cloned as described above.

The details of cloning of Flag-HA- HP5, ADD1-PA, XNP and CG8108 are available upon request.

4.2.1.1 PCR and agarose gel electrophoresis.

For amplification of fragments used for cloning, PCR reactions containing 1 ul 10 pmol each primer, 1 ul 100 ng template, 1 ul 10 mM dNTPs, 10 ul 5x Phusion Polymerase buffer, 0.5 ul Phusion Polymerase and 35.5 ul of

water, were incubated in a PCR thermocycler according to Phusion polymerase protocol. PCR products were separated on an 1.5% agarose gel, made by dissolving agarose in TBE (90 mM Tris, 90 mM Boric acid, 2 mM EDTA) in a microwave. Electrophoresis was performed at 90V until all samples entered the gel, and then at 110V. Bands of expected size were excised with scalpel, and DNA was purified from gel with QIAquick PCR Purification Kit (QIAGEN) or NucleoSpin Gel and PCR Clean-up (Macherey Nagel).

4.2.2 Cell culture and generation of stable cell lines.

Drosophila L2-4 cells and L2-4 stable cell lines were grown in Schneider medium, supplemented with ampicillin/streptomycin and 10% fetal calf serum, at 26°C.

For generation of stable cell lines, cell were transfected with plasmids mixed with X-tremeGENE™ HP DNA Transfection Reagent according to manufacture's instructions. All transfected plasmids contained hygromycin B resistance gene, and cell lines were selected with 100 ug/ml hygromycin B. Cells were optionally induced with 250 µM $CuSO_4$ 12-24 hours before the experiment.

$dCenp_{AP}$ and $HP1a_{AP}$ cell lines were diluted in 20% conditioned medium (Böttcher et al., 2014), and clones originating from several cells were grown separately. Clones were checked by immunofluorescence and those, which contained less hugely overexpressing cells, were selected for further work (clone 8 for $dCenp_{AP}$ and clone 29 for $HP1a_{AP}$).

For long-term storage, 40 million cells/vial were resuspended in a solution containing 50% FCS, 40% medium with antibiotics and FCS, and 10% DMSO. Vials were put into isopropanol box overnight for slow freezing, at -80°C. Next day cells were transferred into liquid nitrogen tank and kept there for storage. For defreezing, one vial of cells was resuspended in 13 ml medium and put in a medium flask. After cells adhered, the medium was changed. Hygromycin B was added on the 2nd-3rd day of defreezing, if necessary.

4.2.3 Proximity labeling coupled to proteomics.

4.2.3.1 Proximity labeling.

Cells were grown in roller bottles up to 5 million cells/ml density. For biotinylation in solution cells were counted, and 10^9 cells were spun down 250g 20 minutes and resuspended in 200 ml DMSO or biotin-phenol in PBS. After 0,5 hours incubation, H_2O_2 was added to the concentration of 1mM, and cells were spun down 250g 20 minutes. Solution was aspirated, and cells were resuspended in quenching solution (5 mM Trolox, 10 mM sodium azide, 10 mM sodium ascorbate). Cells were washed two more times in quenching solution, and the last washing step was performed in 15 ml falcons. After washes the cells were subjected to nuclear extraction.

For biotinylation on plates 2 bottles (400 ml) of 5 million cells/ml were adhered on 40 15-cm plates for mammalian cells for 40 minutes. Then the medium was removed and 15 ml of biotin-phenol/PBS or DMSO/PBS per plate was added. In half an hour H_2O_2 was added to biotin-phenol/PBS to the concentration of 1mM for 1.5 or 5 minutes. The solution was aspirated and 15 ml of quenching solution per plate was added. Cells were scraped off in quenching solution, spun down 20 min 250g in 250 ml conic tubes, and washed once more in 10 ml of quenching solution.

4.2.3.2 Nuclear extraction.

Nuclear extraction was performed as in (Barth et al., 2014) with modifications. Protease inhibitors were added to all buffers. Cells were resuspended in 3 packed cell volumes (PCV) (for example, 2.1 ml) of hypotonic buffer (10 mM NaCl, 10 mM Tris pH 7.6, 1.5 mM $MgCl_2$, 0.1 mM EDTA) and left on ice for 30 minutes. Swelled cells were centrifuged 250g 10 minutes at 4°C, and supernatant was aspirated. Cells were resuspended in 3 PCV hypotonic buffer with 0.2% NP-40. Cells were rotated at 4°C for 10 minutes for the lysis of plasma membrane, and nuclei were centrifuged at 4°C 10 minutes 1000g. Nuclei were further resuspended in 2 ml of quenching solution, and spun down

at 4°C 10 min 1500 g. Supernatant was aspirated, and nuclei were snap-frozen in liquid nitrogen.

Next day nuclei were resuspended in Tris-Ex100 buffer (100 mM NaCl, 10 mM Tris pH 7.6, 1.5 mM $MgCl_2$, 0.5 mM EGTA and 10% v/v glycerol), in 3 ml per 0.7 ml PCV. 1500 units Mnase, 1500 units Benzonase and 2mM $CaCl_2$ were added and nucleic acids were digested 20 min at 26°C. Reaction was stopped by addition of EDTA and EGTA on ice to 10 mM each. Nuclei were disrupted by douncing 10 times in tight-fitting Braun pestle. To solubilize and extract chromatin, NaCl was added to 600 mM final concentration (f.c.), Triton-X-100 to 1%, sodiumdeoxycholate (SOD) to 0.5% and SDS to 0.1%. Chromatin was rotated for 1 hour at 4°C, and then nuclear extracts were centrifuged at 4°C 20 minutes 10000g to get rid of insoluble material. Gained nuclear extracts were dialyzed 4 hours at 4°C through 3.5 MWCO Millipore membranes against Tris-Ex100 buffer without glycerol and with detergents, supplemented with 0.2 mM PMSF and 1 mM DTT. Dialyzed extracts were snap-frozen.

4.2.3.3 Immunoprecipitation.

For immunoprecipitation at room temperature (RT) protease inhibitors were added freshly to defrozen nuclear extracts. 500 ul per sample of streptavidin beads (Pierce) were washed twice in 500 ul Tris-Ex100 + detergents, and washed beads were mixed with input material. 200 ul aliquot of nuclear extract was kept for further Western blotting.

Immunoprecipitation lasted 1.5 hours at RT on a rotating wheel. Afterwards beads were washed 2 times with 500 ul Tris-Ex100 + detergents with protease inhibitors, one time in 1 ml 10 mM Tris 2M urea, and two more times with 500 ul Tris-Ex100 + detergents with protease inhibitors. After the last washing step 50 ul of beads were stored for further Western blotting, and remaining 450 ul beads were further processed for mass spectrometry.

4.2.3.4 Mass spectrometry.

Before on-beads digestion, beads were washed 3 times in 500 ul of 50 mM Tris pH 7.5, 4M urea. Proteins on beads were reduced in 500 µl of 20 mM DTT in 50 mM Tris, 2M urea pH 7.5 with Lys C 450 ng/sample at 27°C for 1h. Next, 50 mM iodoacetamide fc was added and proteins were alkylated for 1 h 25°C shaking 900 rpm in the dark. The reaction was stopped by addition of DTT to 10 mM fc. The samples were shaken 900 rpm at 25°C for two more hours for better LysC digestion. After that urea concentration was reduced to 1.5M by addition of 300 ul water. 1.5 ug of trypsin and 2 mM fc $CaCl_2$ were added for digestion overnight at 25°C shaking 900 rpm. In the morning another 1.5 ug of trypsin were added, and the digestion was performed for 4 more hours. After digestion the supernatant was collected, and beads were washed with 100 µl of 20 mM Tris 50 mM NaCl 25% ACN 2 times for elution of loosely-bound peptides from the beads. Washes were combined with the supernatant, and evaporated at less than 28°C. Next day samples were resuspended in 100 ul 0.1% formic acid (FA) and desalted. The second elution from the beads was performed in 300 µl 0.05% SDS, 0.1% FA at 80°C for 10 min. The elution was dried and subjected to HILIC chromatography.

Desalting was performed the following way: C18 Stage tips (3 white discs, use one stage tip per sample) were placed in Eppendorf vials using the adaptors and washed with 2x 50 µl MeOH and 3x 70 µl 0.1% trifluoroacetic acid (TFA) using an Eppendorf 5804R centrifuge. Centrifugation was performed at 1500-2500 rpm at 20-25°C for each step until the liquid passed the C18 filters. Sample, redissolved in 0.1% FA, was applied on the top of the stage tip and centrifuged at 800-1500 rpm until it completely passed the C18 tips. The filter was washed 2x with 0.1% FA and dried by centrifugation (1500-2500 rpm). The stage tips were placed in the fresh 1.5 ml Eppendorf vials and 2x 100 ul elution solvent (70% ACN, 0.1% FA) was added on top of the stage tips and centrifuged at 800-1500 rpm.

For HILIC purification second elution was redissolved in 100 ul 85% ACN 15 mM NH4Ac, loaded on the HILIC column and centrifuged using the adaptors 800 rpm using a Eppendorf 5804R centrifuge. The columns were washed 2x with 60 ul loading solvent (85% ACN, 15 mM NH4Ac), and eluted

with 2x60 ul Elution II (55% ACN 15 mM NH4Ac pH 3.5) and 3x60ul Elution V (10% ACN 15 mM NH4Ac pH 3.5).

HILIC elutions were combined with C18 elutions, vacuum-dried and redissolved in 45 ul 0.1% FA. Samples were centrifuged at 4°C 30 min 20000 g, and 40 ul of supernatant was taken for mass spectrometry analysis. Desalted peptide mixtures from tryptic digestion were subjected to nano-reversed phase liquid chromatography (nRP-LC) separation coupled to online tandem mass spectrometry (MS/MS) analysis on an Ultimate 3000 nano chromatography system coupled to a QExactive HF mass spectrometer (both Thermo Fisher Scientific). Of each sample 2-4 technical replicates were acquired. For direct injection onto the separation column (Picotip emitter tips, 120 x 0.075 mm, in house packed with ReprosilAQ-C18, Dr. Maisch GmbH, 2.4 µm), samples were loaded at a flow rate of 0.3 µl/min. The peptides were separated by a linear gradient generated over 50 min from 3% ACN to 40% ACN. For online detection of peptides, the outlet of the column served as electrospray ionization emitter to transfer the peptide ions directly into the mass spectrometer. The QExactive mass spectrometer was operated in a data-dependent duty cycle to detect intact peptide ion in positive ion mode in the initial survey scan and perform peptide fragmentation experiments for up to 10 precursors per cycle. Mass spectra were recalibrated using the signals of ambient siloxanes. The survey scan was acquired at a resolution of 60,000 and an AGC target of 3 e9 ions. In order to select suitable precursor ions, the charge state was defined from the previous full scan. Ions with charge states between 2+ and 5+ and minimal abundance of 67,000 ions were isolated in a 2 Da window and subjected to higher-energy collisional fragmentation in the HCD-Trap. MS/MS spectra were generated from 1.5 e5 ions and were acquired at a resolution of 15,000. For peptide fragmentation, normalized collision energy of 27 was applied. Precursors were excluded from repeated fragmentation for 20 seconds to avoid acquisition of redundant MS/MS data from highly abundant peptide species. For each technical repeat, one raw data file was generated which included all survey and fragment ion data.

4.2.4 Microscopy.

1 million cells were adhered on the 12 mm glass coverslips for 30 minutes at RT. Cells were washed in PBS for 5 minutes and fixed in 3.7% formaldehyde 0.3% Triton-X-100/PBS for 12 minutes (or just in 3.7% formaldehyde/PBS for 10 minutes for BubR1 intensity quantification staining, as well as stainings after RNAi). Cells were immediately rinsed with PBS after fixation and washed in PBS for 5 minutes. Next cells were permeabilized on ice with 0.25% Triton-X-100/PBS for 6 minutes. After this cells were immediately rinsed 2 times with PBS and washed with PBS 2 more times 5 minutes each. Blocking was performed for 45 minutes with Image-iT FX signal enhancer in a humidified chamber. After rinse in PBS, primary antibody, diluted in 5% NGS, was incubated on coverslips for 1 hour at RT (or overnight at 4°C for STED microscopy and after RNAi experiments). After 6 minutes' wash in 0.1% Triton-X-100/PBS and following 5 minutes' wash in PBS, secondary antibody, diluted in 5% NGS, was incubated on coverslips. Cells were next washed for 6 minutes with 0.1% Triton-X-100/PBS, rinsed 2 times and washed for 5 minutes with PBS, stained with DAPI/PBS (200 ng/mL or 50 ng/mL for STED microscopy), washed 2 more times 5 minutes with PBS, and mounted in VECTASHIELD (usually) or ProLong™ Diamond Antifade (only for STED microscopy).

For in situ biotinylation on coverslips followed by immunofluorescence, cells were adhered as described previously. Next, cells were incubated with biotin-phenol/PBS or DMSO/PBS, followed by addition of H_2O_2 to 1 mM f.c. for a defined time. The biotinylation reaction was stopped by aspirating the solution and addition of quenching solution, after which cells were processed for immunofluorescence as described previously. Biotinylation in solution was performed with 10^6 cells in 200 ul volume. Cells were spun down during the biotinylation procedure for 10 minutes 250 g, resuspended in 200 ul quenching solution, and adhered on coverslips for 15 minutes, followed by the standard immunofluorescence procedure.

Images were acquired on a Zeiss Axiovert 200 epifluorescence microscope with a CCD Camera (AxioCam MR, Zeiss). For confocal microscopy Leica TCS SP5 microscope or Leica SP8X WLL microscope, equipped with 405 nm laser, WLL2 laser (470 - 670 nm) and acousto-optical

beam splitter, were used. The microscopes reside at Core Facility Bioimaging of Biomedical Center. Acquisition of STED images was performed with a 100x1.4 objective, 24-25 nm pixel size and following settings: DAPI (excitation 405 nm; emission 415-470 nm), Alexa Fluor 594 (590 nm; 600-625) and Abberior STAR 635P (635; 645-720). Sequential recording was performed not to allow channel misalignment or channel crosstalk. For recordings with hybrid photo detectors were used in a counting mode. For STED and some confocal images deconvolution was performed with Huygens 17.10 p2. For image processing ImageJ was used.

4.2.5 Antibody generation.

Wistar rats were immunized subcutaneously (s.c.) and intraperitonially (i.p.) with 50 µg of GST-APEX fusion protein dissolved in 500µl PBS, 5 nmol CpG2006 (TIB MOLBIOL) an equal volume of incomplete Freund's adjuvant. 6 weeks after immunization a 50 µg boost injection was applied i.p. and s.c. three days before fusion. Fusion of the splenic B cells and the myeloma cell line P3X63Ag8.653 was performed using polyethylene glycol 1500 according to standard protocols (Kohler and Milstein, 1975). Hybridoma supernatants were tested by solid-phase enzyme-linked immunoassay (ELISA) using the recombinant GST-fusion protein and verified by Western blotting of whole cell extracts from APEX2 fusions-expressing cell lines (Fig. 2B). Hybridoma cell line from specifically reacting supernatants were cloned twice by limiting dilution. Experiments in this study were performed with clone 20H10 (rat IgG2a/к).

4.2.6 Data analysis.

4.2.6.1 Protein MaxQuant search.

MaxQuant search was done with MaxQuant version 1.5.3.12 against dmel-all-translation-r6.08.fasta database from Flybase. Technical replicates were assigned to one experiment (to one biological replicate). The search parameters were left default except choosing iBAQ and LFQ quantitations and

setting "Match between runs". The files corresponding to different baits and different time points of biotinylation were run separately.

4.2.6.2 Data sources.

Genomic coordinates were converted from the dm3 to dm6 release using the liftover tool from UCSC. ChIP-seq datasets were taken from GEO with the following numbers: GSE86106 (HMR native and HP1a), GSE118291 (HMR overexpressed), GSE54529 (Rad21 and CAPH2). List of GSE numbers for distance to peak analysis is available in Table 12. The chromatin colors 5-state model was taken from (Filion et al., 2010), and "red" and "yellow" chromatin states were fused in "active". Types of genomic fragments (introns, promoters, etc.) were taken from Flybase (version r6.17). TAD separation score was used from http://chorogenome.ie-freiburg.mpg.de/data_sources.html. Hi-C data from S2 cells was taken from GSE58821.

4.2.6.3 Proteomics data analysis.

LFQ values were log2-transformed and missing values were imputed using impute.MinProb function (imputeLCMD R package v2.0) using q = 0.05. The values were median normalized after imputation. Statistical tests were performed using the lmFit and eBayes functions from the limma R package (version 3.34.9). Volcano plots were built using Dr. Tamas Schauer LabeledPlots App (https://tschauer.shinyapps.io/LabledPlots/). Code is available upon request. GO-term analysis was performed using the Gene Ontology consortium tool (http://geneontology.org). Only lowest-hierarchy terms were considered. The protein-protein interaction network was built using STRING database (Szklarczyk et al., 2017), and interactions were taken from experiments, gene fusion, databases, co-expression and co-occurrence. The minimum required interaction score was 0.7. The STRING network was further imputed into Cytoscape and additional interactions were added from Flybase as dotted lines.

4.2.6.4 ChIP-sequencing and Hi-C data analysis.

Sequencing reads were mapped to the Drosophila genome (version dm6) using bowtie2 (version 2.2.9) and filtered by quality using samtools (version 1.3.1). ChIP-seq tracks were generated by Homer (version 4.9) and normalized to sequencing depth and to input. Peaks were identified by Homer with parameters -style factor -F 2 -size 200 (except for Rad21 -F 4). The number of overlapping peaks was visualized as Venn diagrams with the Vennerable R package (version 3.1.0.9). HMR peaks were classified by the overlap with Rad21 or CAPH2 peaks. Such peak groups were further characterized by genomic regions (i.e. promoters, exons, introns or intergenic) and epigenetic domains (i.e. active, inactive, polycomb or HP1-type). ChIP-seq tracks were visualized as heatmaps or average plots centered at the pool of native and overexpressed HMR peaks. Heatmaps were sorted by HMR native ChIP enrichment and clustered by HP1a ChIP enrichment.

ChIP-seq profiles for distance comparison (see Table 12 for GSE numbers) were processed using Homer as described above with optimized -F peak finding parameters. Peak coordinates were imported to R and distances between peak centers were calculated using the distanceToNearest function (GenomicRanges package version 1.36.1).

CAP-H2 and HMR ChIP-seq data (for knockdown / overexpression experiments) were processed using Homer as described above with -F 4 peak finding parameter. Peak finding was carried out on the pool of the reads from replicates. ChIP-seq tracks were visualized as average plots centered at the pool of HMR peaks or CAP-H2 peaks in a 2 kb window. Statistical analysis was performed on the mean ChIP signal at the center of the peak for each replicate. Paired t-test was performed on the mean values. The ChIP-seq data was deposited at GSE137194.

Hi-C raw data were processed using HiC-Pro software (version 2.9.0) with ICE normalization. The Hi-C contact matrix was imported to R using HiTC package (version 1.28.0). For average Hi-C signal surrounding HMR or CAPH2 sites, the Hi-C matrix was centered in 400 kb windows and averaged across sites. For average Hi-C signal at pairs of sites, the Hi-C matrix was subset for each site pair and interpolated to the same scale using interp.surface.grid

function (fields package version 9.8-6) and averaged across sites. For both analyses, expected matrix was calculated as the average contact frequency as a function of genomic distance.

Figures were plotted by R graphics.

4.2.7 Western blotting of extracts from whole cells.

15 (or 1.5) million cells were collected, washed twice in PBS and frozen. Next day the cells were resuspended on ice in 80 (or 20) ul of RIPA buffer (10 mM Tris pH 7.6, 140 mM NaCl, 1 mM EDTA, 1% Triton-X-100, 0.1% SOD, 0.1% SDS) with protease inhibitors. 30 units of benzonase were added and nucleic acids were digested on ice for 30 minutes. After this 20 (or 5) ul of 5x Laemmli buffer (250 mM Tris pH 6.8, 10% SDS, 500 mM DTT, 0.5% Bromophenol blue, 25% Glycerol) were added and lysates were boiled 10 minutes 96°C. 10 ul of boiled lysates were loaded together with marker on the Serva gel in a chamber with running buffer (25 mM Tris, 192 mM glycine, 0,1% SDS) and separated by electrophoretic mobility. Electrophoresis was performed at 90V until all samples entered the gel, and then at 135V. After the run gels were put in a sandwich with a membrane between two Whatman papers between two sponges (all materials soaked in transfer buffer), and proteins were transferred to the membrane at 4°C in Western blotting chambers containing a pack of ice and filled with Western blotting buffer (20 mM Tris, 192 mM glycine, 20% methanol, 0.02% SDS). Transfer was performed for 2.5 hours at 200V and 400mA. Membranes were blocked in 5% milk for 1 hour at RT, and incubated with primary antibody in 1% milk overnight at 4°C or for 2 hours at RT. After two washes in 0.1% Tween 20/PBS (PBST) for 5 minutes, the membranes were incubated with secondary antibody, washed with PBST two more times and developed using Bio-Rad Clarity™ Western ECL Substrate in ChemiDoc™ Imaging System.

4.2.8 RNAi.

Double stranded RNAs (dsRNAs) were synthesized according to the manufacture's instructions with MEGAscript RNA kit. Primers are available in Table 3. RNAi was performed similarly to (Thomae et al., 2013). 1 million cells per well were seeded in a 6-well plate and grown overnight. The next day the medium was removed, and 10 ug dsRNA (or 5 ug of each dsRNA for Cenp-C RNAi) were added in 1 ml of serum free medium. Cells were gently shaken on a platform at RT for 10 minutes, which was followed by 50 minutes incubation at 26°C. After that 2 ml of medium with serum was added. Cells were optionally split on day 4 (for experiments with BubR1 quantification upon HMR RNAi) and were collected on day 6-7. In the case of BubR1 quantification upon HMR RNAi cells were treated with 0.025 mM colchicine 16 hours before harvesting (Godinho and Tavares, 2008). In case of rescue experiments after dCenpC RNAi, transient transfection was performed on day 4 after RNAi and cells were collected on day 7.

4.2.9 cDNA synthesis and RT-qPCR.

4 million cells were harvested and frozen. RNA was extracted with RNAeasy mini kit according to the manufacture's instructions. SuperScript™ III First-Strand Synthesis System kit was used for cDNA synthesis from 1 ug of RNA, treated with Dnase for 1 hour. cDNA was treated with RNAse H at 37°C for 20 minutes, was diluted 1:20 and qPCR was performed in 96 well plate sealed with a foil, in Roche LightCycler 480 II. In every well reaction mix, consisting of 1 ul 3 mM each primer, 2 ul template, 1 ul water and 5 ul PowerUp™ SYBR™ Green Master Mix, was used. Triplicates were performed for all reactions.

Primer design was performed with primer3. Primers against cohesin and condensin subunits' cDNA were designed with default parameters, except setting product size to 80-120 base pairs (bp) and annealing temperature (Tm) 49-61°C. Primers against RpL32 were designed with default parameters except product size: min: 90, opt: 120, max: 140; primer Tm: min: 58, opt: 60, max: 61; max Tm difference: 1.0; primer GC%: min: 50, max: 60; max poly-X: 3; CG

Clamp: 1. All primers (Table 4) were titrated, melting and fluorescence curves were analyzed and primer efficiencies were calculated. Calculation was performed using Roche LightCycler 480 II software.

During RT-qPCR reaction Ct values were calculated by the machine. Ct is the number of PCR cycle when fluorescence intensity of the product becomes above the background. Using the Ct values it is possible to determine the change in the transcript level relative to the normalizer (e.g. GST RNAi) and reference housekeeping gene (e.g. RpL32). The difference was calculated the following way: ΔCt = Ct(gene)-Ct(reference gene). ΔΔCt= ΔCt(RNAi condition)-ΔCt(GST RNAi). Expression fold change = (primer efficiency)$^{\Delta\Delta Ct}$.

4.2.10 ChIP-sequencing.

ChIP-sequencing was performed as described in (Gerland et al., 2017). Cells were washed with PBS and crosslinked in 1% formaldehyde for 5 minutes at room temperature. The crosslinking was quenched by addition of 12% f.c. glycine. The crosslinked material was washed 2 times with PBS and resuspended in ChIP buffer A (10 mM HEPES pH 7.6, 10 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.25% Triton-X-100), supplemented with protease inhibitors and MG132. Cells were lysed 10 minutes on rotation wheel in the cold room, after which the chromatin was pelleted and resuspended in ChIP buffer B (10 mM HEPES pH 7.6, 100 mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.01% Triton-X-100), supplemented with protease inhibitors and MG132. Chromatin was rotated 10 minutes on rotation wheel in the cold room, after which was aliquoted, pelleted and frozen with material from 50 million cells/tube.

1 day before the IP protein A/G beads were mixed in 1:1 ratio and 30 ul beads/IP were precoupled to the antibodies overnight. In the case of HMR ChIP, HMR antibodies were precoupled via an anti-rat bridging antibody (of which 6 ul was incubated in RIPA buffer with the beads 1 hour at room temperature).

At the day of IP 1 vial of chromatin per sample was thawed by resuspending in 1 ml of TE buffer (10 mM Tris pH 7.5, 1 mM EDTA pH 8.0),

supplemented with protease inhibitors and MG132. Chromatin was pelleted and resuspended in 1 ml TE buffer with protease inhibitors and MG132+0.1% SDS. The solution was transferred to Covaris tube and sheared in the Covaris machine with the following settings: 10 min, 140W, 5% duty, 200 cycles per burst. The sheared chromatin was transferred to 1.5 ml tubes, the buffer was adjusted to RIPA, and protease inhibitors PMSF and MG132 were refreshed. In case of HMR ChIP replicate 1, SDS concentration was adjusted to 0.11% instead of 0.1%. Chromatin was precleared 2 times by centrifugation and further by incubating with 30 ul mix of protein A/protein G beads for 1 hour in the coldroom. Afterwards precleared chromatin was centrifuged 2 times again, 1/10 of sample was taken as an input (100 ul) and stored overnight in the fridge, and the rest of the sample was mixed with beads precoupled to the antibodies. IP was performed overnight.

Next day the supernatant was trashed and beads were washed 5 times with RIPA buffer supplemented with protease inhibitors and MG132. 200 ul TE buffer was added to the beads, and 100 ul TE buffer was added to 100 ul input. 4 ul 10 mg/ml RNAse A was added, and samples were incubated at 37°C for 30 minutes shaking. 0.5% f.c. SDS and 20 ul 10 mg/ml Proteinase K were added, and samples were incubated shaking at 56°C for 2 hours and at 65°C overnight for protein digestion and reverse crosslinking respectively. Next day DNA was purified using AMPure beads.

Libraries were prepared from 1 ng DNA with MicroPlex Diagenode Kit according to manufacture's instructions, without size selection, by Angelika Zabel. Sequencing was performed at LAFUGA facility, by Dr. Stefan Krebs.

5 Abbreviations.

ACN	acetonitrile
ADD1	ADD domain-containing protein 1
APC	anaphase-promoting complex
APEX	ascorbate peroxidase
APEX$_{NLS}$	ascorbate peroxidase fused to nuclear localization signal
BirA	bifunctional ligase/repressor birA
Bub3	budding uninhibited by benzimidazoles 3
BubR1	budding uninhibited by benzimidazole-related 1
C. elegans	*Caenorhabditis elegans*
Cal1	chromosome alignment defect 1
CAP-H2	chromosome-associated protein H2
CCAN	constitutive centromere-associated network
Cdc20	cell division cycle protein 20
Cenp-A	centromere protein A
Cenp-C	centromere protein C
CenH3	centromere specific histone H3
ChIP-sequencing	chromatin immunoprecipitation-sequencing
CID	centromere identifier
Cp190	centrosomal protein 190 kDa
CPC	chromosomal passenger complex
CRISPR	clustered regularly interspaced short palindromic repeats
Ct	cycle threshold
CTCF	CCCTC-binding factor
D. melanogaster	*Drosophila melanogaster*
D. simulans	*Drosophila simulans*
DAPI	4′,6-diamidino-2-phenylindole
dCas9	nuclease-deficient CRISPR-associated protein 9
dCenpA	*Drosophila* Cenp-A
dCenpA$_{AP}$	*Drosophila* Cenp-A fused to APEX

dCenpC	*Drosophila* Cenp-C
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotides
DMSO	dimethyl sulfoxide
DFG	Deutsche Forschungsgemeinschaft
dsRNA	double-stranded RNA
FISH	fluorescence in situ hybridization
FCS	fetal calf serum
f.c.	final concentration
GEO	Gene Expression Omnibus
GFZF	GST-containing FLYWCH zinc-finger protein
GFZF$_{AP}$	GST-containing FLYWCH zinc-finger protein fused to APEX
GFP	green fluorescent protein
GST	glutathione S-transferase
GO	gene ontology
H2A	histone 2A
H2AZ	histone 2AZ
H2B	histone 2B
H3	histone 3
H3K4me1/2	histone 3 lysine 4 mono/dimethylation
H3K9me1/2/3	histone 3 lysine 9 mono/di/trimethylation
H3K27me1/2/3	histone 3 lysine 27 mono/di/trimethylation
H3K36me2/3	histone 3 lysine 36 di/trimethylation
H3K64me3	histone 3 lysine 64 trimethylation
H4	histone 4
H4K5Ac	histone 4 lysine 5 acetylation
H4K12Ac	histone 4 lysine 12 acetylation
H4K20me1/2/3	histone 4 lysine 20 mono/di/trimethylation
HBO1	histone acetyltransferase bound to ORC 1
HCP-3	histone H3-like centromeric protein
HILIC	hydrophilic interaction liquid chromatography
HJURP	holiday junction recognition protein
HMR	hybrid male rescue

HMR$_{AP}$	HMR fused to APEX
HP1a	heterochromatic protein 1a
HP1a$_{AP}$	HP1a fused to APEX
HP1β	heterochromatic protein 1β
HP3	heterochromatic protein 3
HP4	heterochromatic protein 4
HP5	heterochromatic protein 5
HP6	heterochromatic protein 6
HPLC	high performance liquid chromatography
Hygro	Hygromycin
i.p.	intraperitonially
iBAQ	intensity-based absolute quantification
Incenp	inner centromere protein
IP	immunoprecipitation
IP-MS	immunoprecipitation – mass spectrometry
kb	kilobase
KAT7	lysine acetyltransferase 7
LFQ	label free quantitation
LHR	lethal hybrid rescue
LHR$_{AP}$	LHR fused to APEX
Mad1	mitotic arrest deficiency 1
Mad2	mitotic arrest deficiency 2
MNase	micrococcal nuclease
mod(mdg4)	modifier of mdg4
mRNA	messenger RNA
MS/MS	tandem mass spectrometry
MWCO	molecular weight cut-off
MYST2	MYST domain containing lysine acetyltransferase
NLP	nucleoplasmin-like protein
NLS	nuclear localization. signal
nRP-LC	nano-reversed phase liquid chromatography
n.s.	non-significant
NP-40	nonyl phenoxypolyethoxylethanol
OdsH	Ods-site homeobox

PCR	polymerase chain reaction
PCV	packed cell volume
pMT	metallothionein promoter
PolII	RNA polymerase II
prod	proliferation disruptor
Rad21	radiation 21
RIPA buffer	radioimmunoprecipitation assay buffer
RNA	ribonucleic acid
RNAi	RNA interference
RpL32	ribosomal protein L32
RT	room temperature
RT-qPCR	real time quantitative reverse transcription PCR
SAC	spindle assembly checkpoint
s.c.	subcutaneously
SDS	sodium dodecyl sulfate
SMC1	structural maintenance of chromosomes 1
SMC2	structural maintenance of chromosomes 2
snRNP	small nuclear ribonucleoproteins
SOD	sodium deoxycholate
STED	stimulated emission depletion microscopy
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
Su(Hw)	suppressor of hairy wings
Su(var)3-9	suppressor of variegation 3-9
TAD	topologically associated domain
QBM	Graduate School of Quantitative Biosciences Munich
USCS	University of California Santa Cruz
Vtd	verthandi
WLL	white light laser
Zhr	zygotic hybrid rescue

6 References.

Ainsztein, A.M., Kandels-Lewis, S.E., Mackay, A.M., and Earnshaw, W.C. (1998). INCENP Centromere and Spindle Targeting: Identification of Essential Conserved Motifs and Involvement of Heterochromatin Protein HP1. The Journal of Cell Biology *143*, 1763-1774.

Akhtar, A., Zink, D., and Becker, P.B. (2000). Chromodomains are protein-RNA interaction modules. Nature *407*, 405-409.

Alekseyenko, A.A., Gorchakov, A.A., Zee, B.M., Fuchs, S.M., Kharchenko, P.V., and Kuroda, M.I. (2014). Heterochromatin-associated interactions of Drosophila HP1a with dADD1, HIPP1, and repetitive RNAs. Genes Dev *28*, 1445-1460.

Allshire, R.C., and Madhani, H.D. (2018). Ten principles of heterochromatin formation and function. Nat Rev Mol Cell Biol *19*, 229-244.

Anselm, E., Thomae, A.W., Jeyaprakash, A.A., and Heun, P. (2018). Oligomerization of Drosophila Nucleoplasmin-Like Protein is required for its centromere localization. Nucleic Acids Res *46*, 11274-11286.

Aruna, S., Flores, H.A., and Barbash, D.A. (2009). Reduced fertility of Drosophila melanogaster hybrid male rescue (Hmr) mutant females is partially complemented by Hmr orthologs from sibling species. Genetics *181*, 1437-1450.

Bailey, A.O., Panchenko, T., Shabanowitz, J., Lehman, S.M., Bai, D.L., Hunt, D.F., Black, B.E., and Foltz, D.R. (2016). Identification of the Post-translational Modifications Present in Centromeric Chromatin. Mol Cell Proteomics *15*, 918-931.

Barth, T.K., Schade, G.O.M., Schmidt, A., Vetter, I., Wirth, M., Heun, P., Thomae, A.W., and Imhof, A. (2014). Identification of novel Drosophila centromere-associated proteins. Proteomics *14*, 2167-2178.

Bayes, J.J., and Malik, H.S. (2009). Altered heterochromatin binding by a hybrid sterility protein in Drosophila sibling species. Science *326*, 1538-1541.

Bendayan, M. (2001). Worth Its Weight in Gold. Science *291*, 1363-1365.

Bergmann, J.H., Rodriguez, M.G., Martins, N.M., Kimura, H., Kelly, D.A., Masumoto, H., Larionov, V., Jansen, L.E., and 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.

Bersuker, K., Peterson, C.W.H., To, M., Sahl, S.J., Savikhin, V., Grossman, E.A., Nomura, D.K., and Olzmann, J.A. (2018). A Proximity Labeling Strategy Provides Insights into the Composition and Dynamics of Lipid Droplet Proteomes. Dev Cell *44*, 97-112 e117.

Bickmore, W.A., and van Steensel, B. (2013). Genome Architecture: Domain Organization of Interphase Chromosomes. Cell *152*, 1270-1284.

Blattes, R., Monod, C., Susbielle, G., Cuvier, O., Wu, J.H., Hsieh, T.S., Laemmli, U.K., and Käs, E. (2006). Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. EMBO J *25*, 2397-2408.

Blower, M.D. (2004). Centromeric Transcription Regulates Aurora-B Localization and Activation. Cell Rep *15*, 1624–1633.

Blum, J.A., Bonaccorsi, S., Marzullo, M., Palumbo, V., Yamashita, Y.M., Barbash, D.A., and Gatti, M. (2017). The Hybrid Incompatibility Genes Lhr and Hmr Are Required

for Sister Chromatid Detachment During Anaphase but Not for Centromere Function. Genetics *207*, 1457-1472.

Bobkov, G.O.M., Gilbert, N., and Heun, P. (2018). Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation. J Cell Biol *217*, 1957-1972.

Boettiger, A.N., Bintu, B., Moffitt, J.R., Wang, S., Beliveau, B.J., Fudenberg, G., Imakaev, M., Mirny, L.A., Wu, C.T., and Zhuang, X. (2016). Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. Nature *529*, 418-422.

Bolkan, B.J., Booker, R., Goldberg, M.L., and Barbash, D.A. (2007). Developmental and cell cycle progression defects in Drosophila hybrid males. Genetics *177*, 2233-2241.

Böttcher, R., Hollmann, M., Merk, K., Nitschko, V., Obermaier, C., Philippou-Massier, J., Wieland, I., Gaul, U., and Förstemann, K. (2014). Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells. Nucleic Acids Res *42*, e89.

Branon, T.C., Bosch, J.A., Sanchez, A.D., Udeshi, N.D., Svinkina, T., Carr, S.A., Feldman, J.L., Perrimon, N., and Ting, A.Y. (2017). Directed evolution of TurboID for efficient proximity labeling in living cells and organisms. BioArxiv.

Brideau, N.J., Flores, H.A., Wang, J., Maheshwari, S., Wang, X., and Barbash, D.A. (2006). Two Dobzhansky-Muller genes interact to cause hybrid lethality in Drosophila. Science *314*, 1292-1295.

Brown, S.W. (1996). Heterochromatin. Science *151*, 417-425.

Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M., Madhani, H.D., Narlikar, G.J., and Al Sady, B. (2011). Chromodomain-Mediated Oligomerization of HP1 Suggests a Nucleosome-Bridging Mechanism for Heterochromatin Assembly. Mol Cell *41*, 67-81.

Chan, F.L., Marshall, O.J., Saffery, R., Kim, B.W., Earle, E., Choo, K.H.A., and 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.

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.*, et al.* (2015). Establishment of Centromeric Chromatin by the CENP-A Assembly Factor CAL1 Requires FACT-Mediated Transcription. Dev Cell *34*, 73-84.

Chiolo, I., Minoda, A., Colmenares, S.U., Polyzos, A., Costes, S.V., and Karpen, G.H. (2011). Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. Cell *144*, 732-744.

Conde, C., Osswald, M., Barbosa, J., Moutinho-Santos, T., Pinheiro, D., Guimaraes, S., Matos, I., Maiato, H., and Sunkel, C.E. (2013). Drosophila Polo regulates the spindle assembly checkpoint through Mps1-dependent BubR1 phosphorylation. EMBO J *32*, 1761-1777.

Cooper, J.C., Lukacs, A., Reich, S., Schauer, T., Imhof, A., and Phadnis, N. (2019). Altered Localization of Hybrid Incompatibility Proteins in Drosophila. Mol Biol Evol *36*, 1783-1792.

Cosgrove, M.S. (2012). Writers and readers: deconvoluting the harmonic complexity of the histone code. Nat Struct Mol Biol *19*, 739-740.

Dawe, R.K. (2003). RNA interference, transposons, and the centromere. Plant Cell *15*, 297-301.

Disteche, C.M. (2012). Dosage compensation of the sex chromosomes. Annu Rev Genet *46*, 537-560.

Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. Mol Cell *62*, 668-680.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Earnshaw, W., Bordwell, B., Marino, C., and Rothfield, N. (1986). Three human chromosomal autoantigens are recognized by sera from patients with anti-centromere antibodies. J Clin Invest *77*, 426-430.

Eissenberg, J.C., and Elgin, S.C. (2014). HP1a: a structural chromosomal protein regulating transcription. Trends Genet *30*, 103-110.

Eissenberg, J.C., and Reuter, G. (2009). Cellular Mechanism for Targeting Heterochromatin Formation in Drosophila. Int Rev Cell Mol Biol *273*, 1-47.

Erhardt, S., Mellone, B.G., Betts, C.M., Zhang, W., Karpen, G.H., and Straight, A.F. (2008). Genome-wide analysis reveals a cell cycle-dependent mechanism controlling centromere propagation. J Cell Biol *183*, 805-818.

Eskiw, C.H., Cope, N.F., Clay, I., Schoenfelder, S., Nagano, T., and Fraser, P. (2010). Transcription Factories and Nuclear Organization of the Genome. Cold Spring Harbor Symposia on Quantitative Biology *75*, 501-506.

Ferree, P.M., and Barbash, D.A. (2009). Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in Drosophila. PLoS Biol *7*, e1000234.

Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R., Bussemaker, H.J., *et al.* (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell *143*, 212-224.

Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature *438*, 1116-1122.

Fodor, B.D., Shukeir, N., Reuter, G., and Jenuwein, T. (2010). Mammalian Su(var) genes in chromatin control. Annu Rev Cell Dev Biol *26*, 471-501.

Folco, H.D., Pidoux, A.L., Urano, T., and Allshire, R.C. (2008). Heterochromatin and RNAi Are Required to Establish CENP-A Chromatin at Centromeres. Science *319*, 94-97.

Foltz, D.R., Jansen, L.E., Black, B.E., Bailey, A.O., Yates, J.R., 3rd, and Cleveland, D.W. (2006). The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol *8*, 458-469.

Franke, A., and Baker, B.S. (1999). The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. Mol Cell *4*, 117-122.

Franklin, S.G., and Zweidler, A. (1977). Non-allelic variants of histones 2a, 2b and 3 in mammals. Nature *266*, 273-275.

Freire, E., Schon, A., Hutchins, B.M., and Brown, R.K. (2013). Chemical denaturation as a tool in the formulation optimization of biologics. Drug Discov Today *18*, 1007-1013.

Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A., and Mirny, L.A. (2018). Emerging Evidence of Chromosome Folding by Loop Extrusion. Cold Spring Harb Symp Quant Biol.

Gao, X.D., Tu, L.C., Mir, A., Rodriguez, T., Ding, Y., Leszyk, J., Dekker, J., Shaffer, S.A., Zhu, L.J., Wolfe, S.A., *et al.* (2018). C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2. Nat Methods.

Garrigues, J.M., Sidoli, S., Garcia, B.A., and Strome, S. (2015). Defining heterochromatin in C. elegans through genome-wide analysis of the heterochromatin protein 1 homolog HPL-2. Genome Res *25*, 76-88.

Gerland, T.A., Sun, B., Smialowski, P., Lukacs, A., Thomae, A.W., and Imhof, A. (2017). The Drosophila speciation factor HMR localizes to genomic insulator sites. PLoS One *12*, e0171798.

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C., *et al.* (2018). A pathway for mitotic chromosome formation. Science *359*.

Godinho, S., and Tavares, A.A. (2008). A role for Drosophila Polo protein in chromosome resolution and segregation during mitosis. Cell Cycle *7*, 2529-2534.

Goutte-Gattat, D., Shuaib, M., Ouararhni, K., Gautier, T., Skoufias, D., Hamiche, A., and 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.

Greaves, I.K., Rangasamy, D., Ridgway, P., and Tremethick, D.J. (2006). H2A.Z contributes to the unique 3D structure of the centromere. Proc Natl Acad Sci U S A *104*, 525–530.

Green, N.M. (1975). Avidin. *29*, 85-133.

Greil, F., de Wit, E., Bussemaker, H.J., and van Steensel, B. (2007). HP1 controls genomic targeting of four novel heterochromatin proteins in Drosophila. EMBO J *26*, 741-751.

Gupta, R., Somyajit, K., Narita, T., Maskey, E., Stanlie, A., Kremer, M., Typas, D., Lammers, M., Mailand, N., Nussenzweig, A., *et al.* (2018). DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. Cell.

Hayakawa, T., Haraguchi, T., Masumoto, H., and Hiraoka, Y. (2003). Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase. J Cell Sci *116*, 3327-3338.

Heeger, S., Leismann, O., Schittenhelm, R., Schraidt, O., Heidmann, S., and Lehner, C.F. (2005). Genetic interactions of separase regulatory subunits reveal the diverged Drosophila Cenp-C homolog. Genes Dev *19*, 2041–2053.

Henikoff, S., Henikoff, J.G., Sakai, A., Loeb, G.B., and Ahmad, K. (2009). Genome-wide profiling of salt fractions maps physical properties of chromatin. Genome Res *19*, 460-469.

Heun, P., Erhardt, S., Blower, M.D., Weiss, S., Skora, A.D., and Karpen, G.H. (2006). Mislocalization of the Drosophila centromere-specific histone CID promotes formation of functional ectopic kinetochores. Dev Cell *10*, 303-315.

Hori, T., Shang, W.H., Toyoda, A., Misu, S., Monma, N., Ikeo, K., Molina, O., Vargiu, G., Fujiyama, A., Kimura, H., *et al.* (2014). Histone H4 Lys 20 monomethylation of the CENP-A nucleosome is essential for kinetochore assembly. Dev Cell *29*, 740-749.

Hung, V., Zou, P., Rhee, H.W., Udeshi, N.D., Cracan, V., Svinkina, T., Carr, S.A., Mootha, V.K., and Ting, A.Y. (2014). Proteomic mapping of the human mitochondrial intermembrane space in live cells via ratiometric APEX tagging. Mol Cell *55*, 332-341.

Hutter, P., and Ashburner, M. (1987). Genetic rescue of inviable hybrids between Drosophila melanogaster and its sibling species. Nature *327*, 331-333.

Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods *9*, 999-1003.

Iwasaki, O., Tanaka, A., Tanizawa, H., Grewal, S.I.S., and Noma, K.I. (2010). Centromeric Localization of Dispersed Pol III Genes in Fission Yeast. Mol Biol Cell *21*, 254–265.

Izuta, H., Ikeno, M., Suzuki, N., Tomonaga, T., Nozaki, N., Obuse, C., Kisu, Y., Goshima, N., Nomura, F., Nomura, N.*, et al.* (2006). Comprehensive analysis of the ICEN (Interphase Centromere Complex) components enriched in the CENP-A chromatin of human cells. Genes Cells *11*, 673-684.

Jagannathan, M., Cummings, R., and Yamashita, Y.M. (2018). A conserved function for pericentromeric satellite DNA. Elife *7*.

Jagannathan, M., Cummings, R., and Yamashita, Y.M. (2019). The modular mechanism of chromocenter formation in Drosophila. Elife *8*.

Jankovics, F., Bence, M., Sinka, R., Farago, A., Bodai, L., Pettko-Szandtner, A., Ibrahim, K., Takacs, Z., Szarka-Kovacs, A.B., and Erdelyi, M. (2018). Drosophila small ovary gene is required for transposon silencing and heterochromatin organization, and ensures germline stem cell maintenance and differentiation. Development *145*.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T.*, et al.* (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480-486.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-cell dynamics of genome-nuclear lamina interactions. Cell *153*, 178-192.

Kochanova, N.Y., Schauer, T., Mathias, G.P., Lukacs, A., Schmidt, A., Flatley, A., Schepers, A., Thomae, A.W., and Imhof, A. (2018). Intricate structure of the interphase chromocenter revealed by the analysis of a factor involved in species formation. bioRxiv.

Kohler, G., and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. Nature *256*, 495-497.

Kurzban, G.P., Bayer, E.A., Wilchek, M., and Horowitz, P.M. (1991). The quaternary structure of streptavidin in urea. J Biol Chem *266*, 14470-14477.

Lacoste, N., Woolfe, A., Tachiwana, H., Garea, A.V., Barth, T., Cantaloube, S., Kurumizaka, H., Imhof, A., and 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.

Lallemand-Breitenbach, V., and de The, H. (2010). PML nuclear bodies. Cold Spring Harb Perspect Biol *2*, a000661.

Lam, A.L., Boivin, C.D., Bonney, C.F., Rudd, M.K., and 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.

Lam, S.S., Martell, J.D., Kamer, K.J., Deerinck, T.J., Ellisman, M.H., Mootha, V.K., and Ting, A.Y. (2015). Directed evolution of APEX2 for electron microscopy and proximity labeling. Nat Methods *12*, 51-54.

Lambert, J.P., Tucholska, M., Go, C., Knight, J.D., and Gingras, A.C. (2015). Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. J Proteomics *118*, 81-94.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. Nature *547*, 236-240.

Larsson, J., Chen, J.D., Rasheva, V., Rasmuson-Lestander, A., and Pirrotta, V. (2001). Painting of fourth, a chromosome-specific protein in Drosophila. Proc Natl Acad Sci U S A *98*, 6273-6278.

Li, Y., Danzer, J.R., Alvarez, P., Belmont, A.S., and Wallrath, L.L. (2003). Effects of tethering HP1 to euchromatic regions of the Drosophila genome. Development *130*, 1817-1824.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science *326*, 289-293.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Maddox, P.S., Oegema, K., Desai, A., and Cheeseman, I.M. (2004). "Holo"er than thou: Chromosome segregation and kinetochore function in C. elegans. Chromosome Res *12*, 641-653.

Markmiller, S., Soltanieh, S., Server, K.L., Mak, R., Jin, W., Fang, M.Y., Luo, E.C., Krach, F., Yang, D., Sen, A.*, et al.* (2018). Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. Cell *172*, 590-604 e513.

Mendiburo, M.J., Padeken, J., Fulop, S., Schepers, A., and Heun, P. (2011). Drosophila CENH3 is sufficient for centromere formation. Science *334*, 686-690.

Mick, D.U., Rodrigues, R.B., Leib, R.D., Adams, C.M., Chien, A.S., Gygi, S.P., and Nachury, M.V. (2015). Proteomics of Primary Cilia by Proximity Labeling. Dev Cell *35*, 497-512.

Minc, E., Allory, Y., Worman, H.J., Courvalin, J.C., and Buendia, B. (1999). Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells. Chromosoma *108*, 220–234.

Morris, G.E. (2008). The Cajal body. Biochim Biophys Acta *1783*, 2108-2115.

Muller, S., and Almouzni, G. (2017). Chromatin dynamics during the cell cycle at centromeres. Nat Rev Genet *18*, 192-208.

Musacchio, A. (2015). The Molecular Biology of Spindle Assembly Checkpoint Signaling Dynamics. Curr Biol *25*, R1002-1018.

Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B.R., Mirny, L.A., and Dekker, J. (2013). Organization of the mitotic chromosome. Science *342*, 948-953.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Obuse, C., Yang, H., Nozaki, N., Goto, S., Okazaki, T., and Yoda, K. (2004). Proteomics analysis of the centromere complex from HeLa interphase cells: UV-damaged DNA binding protein 1 (DDB-1) is a component of the CEN-complex, while BMI-1 is

transiently co-localized with the centromeric region in interphase. Genes Cells 9, 105-120.

Ohzeki, J., Shono, N., Otake, K., Martins, N.M., Kugou, K., Kimura, H., Nagase, T., Larionov, V., Earnshaw, W.C., and Masumoto, H. (2016). KAT7/HBO1/MYST2 Regulates CENP-A Chromatin Assembly by Antagonizing Suv39h1-Mediated Centromere Inactivation. Dev Cell 37, 413-427.

Olszak, A.M., van Essen, D., Pereira, A.J., Diehl, S., Manke, T., Maiato, H., Saccani, S., and Heun, P. (2011). Heterochromatin boundaries are hotspots for de novo kinetochore formation. Nat Cell Biol 13, 799-808.

Orr, H.A., Madden, L.D., Coyne, J.A., Goodwin, R., and Hawley, R.S. (1997). The Developmental Genetics of Hybrid Inviability - A Mitotic Defect in Drosophila Hybrids. Genetics 145, 1031-1040.

Padeken, J., Mendiburo, M.J., Chlamydas, S., Schwarz, H.J., Kremmer, E., and Heun, P. (2013). The nucleoplasmin homolog NLP mediates centromere clustering and anchoring to the nucleolus. Mol Cell 50, 236-249.

Palmer, D.K., O'Day, K., Wener, M.H., Andrews, B.S., and Margolis, R.L. (1987). A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. The Journal of Cell Biology 104, 805-815.

Pederson, T. (2011). The nucleolus. Cold Spring Harb Perspect Biol 3, a000638.

Phadnis, N., Baker, E.P., Cooper, J.C., Frizzell, K.A., Hsieh, E., de la Cruz, A.F., Shendure, J., Kitzman, J.O., and Malik, H.S. (2015). An essential cell cycle regulation gene causes hybrid inviability in Drosophila. Science 350, 1552-1555.

Pirrotta, V., and Li, H.B. (2012). A view of nuclear Polycomb bodies. Curr Opin Genet Dev 22, 101-109.

Quenet, D., and Dalal, Y. (2014). A long non-coding RNA is required for targeting centromeric protein A to the human centromere. Elife 3, e03254.

Ramirez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., Gruning, B.A., Villaveces, J., Habermann, B., Akhtar, A., and Manke, T. (2018). High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nat Commun 9, 189.

Ramirez, F., Lingg, T., Toscano, S., Lam, K.C., Georgiev, P., Chung, H.R., Lajoie, B.R., de Wit, E., Zhan, Y., de Laat, W., et al. (2015). High-Affinity Sites Form an Interaction Network to Facilitate Spreading of the MSL Complex across the X Chromosome in Drosophila. Mol Cell 60, 146-162.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159, 1665-1680.

Regnier, V., Vagnarelli, P., Fukagawa, T., Zerjal, T., Burns, E., Trouche, D., Earnshaw, W., and Brown, W. (2005). CENP-A is required for accurate chromosome segregation and sustained kinetochore association of BubR1. Mol Cell Biol 25, 3967-3981.

Rhee, H.W., Zou, P., Udeshi, N.D., Martell, J.D., Mootha, V.K., Carr, S.A., and Ting, A.Y. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science 339, 1328-1331.

Riddle, N.C., and Elgin, S.C. (2008). A role for RNAi in heterochromatin formation in Drosophila. Curr Top Microbiol Immunol 320, 185-209.

Riddle, N.C., Jung, Y.L., Gu, T., Alekseyenko, A.A., Asker, D., Gui, H., Kharchenko, P.V., Minoda, A., Plachetka, A., Schwartz, Y.B., et al. (2012). Enrichment of HP1a on

Drosophila chromosome 4 genes creates an alternate chromatin structure critical for regulation in this heterochromatic domain. PLoS Genet *8*, e1002954.

Rosic, S., Kohler, F., and Erhardt, S. (2014). Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. J Cell Biol *207*, 335-349.

Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol *196*, 801-810.

Saksouk, N., Simboek, E., and Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin *8*.

Satyaki, P.R., Cuykendall, T.N., Wei, K.H., Brideau, N.J., Kwak, H., Aruna, S., Ferree, P.M., Ji, S., and Barbash, D.A. (2014). The Hmr and Lhr hybrid incompatibility genes suppress a broad range of heterochromatic repeats. PLoS Genet *10*, e1004240.

Sawamura, K. (2012). Chromatin evolution and molecular drive in speciation. Int J Evol Biol *2012*, 301894.

Schmidtmann, E., Anton, T., Rombaut, P., Herzog, F., and Leonhardt, H. (2016). Determination of local chromatin composition by CasID. Nucleus *7*, 476-484.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of Drosophila SU(VAR)3–9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J *21*, 1121-1131.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev *18*, 1251-1262.

Schubert, T., Pusch, M.C., Diermeier, S., Benes, V., Kremmer, E., Imhof, A., and Langst, G. (2012). Df31 protein and snoRNAs maintain accessible higher-order structures of chromatin. Mol Cell *48*, 434-444.

Scott, K.C., Merrett, S.L., and Willard, H.F. (2006). A heterochromatin barrier partitions the fission yeast centromere into discrete chromatin domains. Curr Biol *16*, 119-129.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.

Shang, W.H., Hori, T., Westhorpe, F.G., Godek, K.M., Toyoda, A., Misu, S., Monma, N., Ikeo, K., Carroll, C.W., Takami, Y.*, et al.* (2016). Acetylation of histone H4 lysine 5 and 12 is required for CENP-A deposition into centromeres. Nat Commun *7*, 13465.

Skibbens, R.V. (2019). Condensins and cohesins - one of these things is not like the other! J Cell Sci *132*.

Spector, D.L., and Lamond, A.I. (2011). Nuclear speckles. Cold Spring Harb Perspect Biol *3*.

Steiner, F.A., and Henikoff, S. (2015). Diversity in the organization of centromeric chromatin. Curr Opin Genet Dev *31*, 28-35.

Straub, T. (2003). Heterochromatin dynamics. PLoS Biol *1*, E14.

Straub, T., and Becker, P.B. (2007). Dosage compensation: the beginning and end of generalization. Nat Rev Genet *8*, 47-57.

Straub, T., Neumann, M.F., Prestel, M., Kremmer, E., Kaether, C., Haass, C., and Becker, P.B. (2005). Stable chromosomal association of MSL2 defines a dosage-compensated nuclear compartment. Chromosoma *114*, 352-364.

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. Nature *547*, 241-245.

Sullivan, B.A., and 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.

Swenson, J.M., Colmenares, S.U., Strom, A.R., Costes, S.V., and Karpen, G.H. (2016). The composition and organization of Drosophila heterochromatin are heterogeneous and dynamic. Elife *5*.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., *et al.* (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res *45*, D362-D368.

Talbert, P.B., Kasinathan, S., and Henikoff, S. (2018). Simple and Complex Centromeric Satellites in Drosophila Sibling Species. Genetics *208*, 977-990.

Thomae, A.W., Schade, G.O., Padeken, J., Borath, M., Vetter, I., Kremmer, E., Heun, P., and Imhof, A. (2013). A pair of centromeric proteins mediates reproductive isolation in Drosophila species. Dev Cell *27*, 412-424.

Török, T., Harvie, P.D., Buratovich, M., and Bryant, P.J. (1997). The product of proliferation disrupter is concentrated at centromeres and required for mitotic chromosome condensation and cell proliferation in Drosophila. Genes Dev *11*, 213-225.

Van Bortle, K., Nichols, M.H., Li, L., Ong, C.T., Takenaka, N., Qin, Z.S., and Corces, V.G. (2014). Insulator function and topological domain border strength scale with architectural protein occupancy. Genome Biology *15*, R82.

Wiblin, A.E., Cui, W., Clark, A.J., and Bickmore, W.A. (2005). Distinctive nuclear organisation of centromeres and regions involved in pluripotency in human embryonic stem cells. J Cell Sci *118*, 3861-3868.

Yang, J., Sun, S., Zhang, S., Gonzalez, M., Dong, Q., Chi, Z., Chen, Y.H., and Li, F. (2018). Heterochromatin and RNAi regulate centromeres by protecting CENP-A from ubiquitin-mediated degradation. PLoS Genet *14*, e1007572.

Zaidan, N.Z., Walker, K.J., Brown, J.E., Schaffer, L.V., Scalf, M., Shortreed, M.R., Iyer, G., Smith, L.M., and Sridharan, R. (2018). Compartmentalization of HP1 Proteins in Pluripotency Acquisition and Maintenance. Stem Cell Reports *10*, 627-641.

Zuzow, N., Chosh, A., Leonard, M., Liao, J., Yang, B., and Bennett, E.J. (2018). Mapping the mammalian ribosome quality control complex interactome using proximity labeling approaches. Mol Biol Cell.

Curriculum Vitae.

Personal information:
Name: Natalia Kochanova
Citizenship: Russian
Date of birth: 23.02.1992
Place of birth: Russian Federation, Moscow

Education:

2009-2014 (specialist degree, equivalent to the old diploma system): Virology Department, Biological Faculty, M.V. Lomonosov Moscow State University, Leninskie Gory, Moscow 119 991, Russian Federation
Major field of studies: biochemistry.
Minor field of studies: virology.
GPA: 4.89/5

2014-2019 doing a PhD Prof. Axel Imhof´s group Molecular Biology Unit, Adolf Butenandt Institute and Center for Integrated Protein Science, Ludwig Maximilians University, Munich, Germany / Graduate School of Quantitative Biosciences (QBM).

Research experience:
September 2014-present *Prof. Axel Imhof's group, Molecular Biology Unit, Adolf Butenandt Institute and Center for Integrated Protein Science, Ludwig Maximilians University, Munich, Germany / Graduate School of Quantitative Biosciences (QBM).*
October 2011- June 2014 *Prof. Dmitri Kramerov's group, Laboratory of eukaryotic genome evolution, Engelhardt Institute of Molecular Biology of Russian Academy of Sciences, Moscow, Russia.*
July-September 2013 *Prof. Axel Imhof's group, Molecular Biology Unit, Adolf Butenandt Institute and Center for Integrated Protein Science, Ludwig Maximilians University, Munich, Germany.*

Conferences:
Presentation of poster "Profiling of S-adenosyl-L-homocysteine binding proteins using a novel compound based strategy":
Amgen Scholars European Symposium –September 8-11, 2013, Cambridge University, UK; Amgen Scholars LMU Symposium –September 13, 2013 LMU Biocenter, Munich, Germany; Modern Biology and Future Biotech Winter School – January 2014, Moscow Region.

EMBO Conference Series "From functional genomics to systems biology" 8–11 November 2014 in Heidelberg, Germany

Presentation of poster "Intricate structure of the interphase chromocenter revealed by the analysis of factor involved in species formation".

Abcam conference on epigenetics "Chromatin and Epigenetics: Inheritance and Design" 01-03 April 2019 Helmholz Zentrum, Munich
FEBS Chromatin Proteomics Workshop 22-27 September 2019 Crete

Talk "Intricate structure of the interphase chromocenter revealed by the analysis of a factor involved in species formation". - *FEBS Chromatin Proteomics Workshop 22-27 September 2019 Crete*

Grants and awards:

2012 - A prize in the popular scientific articles contest "Bio/mol/text" for the popular scientific review "*Alu*: the story of one sequence";

2013 – First prize in the popular scientific articles contest "Bio/mol/text" for the popular scientific article "The mysterious journey of *Xist* RNA across the X-chromosome";

2012, 2013 – An academic state scholarship for students of Russian Ministry of Education and Science;

2013 – fully funded summer internship: *LMU Amgen Scholars Programme.*

2013-2014 – participation in the RFFI grant (Russian Fund of Fundamental Research) "Expression, posttranscriptional modification and degradation of RNA of short SINE retrotransposons and relative genes"*.*

2014-2017 – DFG/QBM fellowship for studying and doing research in Graduate School of Quantitative Biosciences (QBM).