UNIVERSITÀ
DI PAVIA

Dipartimento di Biologia e Biotecnologie "Lazzaro Spallanzani"

Laurea Magistralis in Molecular Biology and Genetics

Molecular characterization of centromeres in lymphoma cell lines

Caratterizzazione molecolare dei centromeri in linee cellulari di linfoma

Supervisor:
Prof.ssa Elena Giulotto

Co-supervisors:
Dott.ssa Maria Francesca Piras
Dott.ssa Eleonora Cappelletti

Experimental thesis by
Doğukan Işık

Academic Year 2025/2026

ABSTRACT

Centromeres are essential chromosomal regions that ensure accurate chromosome segregation and are epigenetically specified by the presence of CENP-A, the centromere-specific histone H3 variant. In humans, canonical centromeres are typically associated with α-satellite DNA sequences and are bound by CENP-B, the only centromeric protein with sequence-specific DNA binding activity. However, satellite sequences are neither necessary nor sufficient to specify centromere identity, and functional centromeres can therefore arise sporadically outside the conventional sequence context, giving rise to neocentromeres. This epigenetic plasticity may become particularly relevant in pathological contexts, as centromere dysfunction contributes to chromosome instability and aneuploidy, key features of cancer genomes. Although deregulated expression of centromeric proteins is common in cancer and has been associated with poor prognosis, the genomic and epigenomic organization of centromeres in cancer remains largely unexplored. Centromeric chromatin may therefore be particularly susceptible to remodeling during tumor evolution, especially in cancers characterized by extensive genome rearrangements, such as chromothripsis.

This study investigated centromere organization in lymphoma cell lines from the LL-100 panel, focusing on the relationship between centromere remodeling, genome instability, extensive chromosomal rearrangements and chromothripsis. Transcriptomic analysis of centromere-associated genes across the LL-100 panel revealed recurrent deregulation of centromeric protein expression, without a clear subtype-specific expression pattern. Several centromeric proteins, including CENP-A, CENP-H, CENP-W, CENP-M, CENP-N and CENP-X, were overexpressed in multiple cell lines, whereas CENP-C generally displayed lower expression compared with healthy lymph-node controls. Interestingly, CENP-B showed marked cell-line-specific variation and was undetectable in the Burkitt lymphoma BJAB cell line. Based on these findings, three cell lines were selected for more detailed analyses of CENP-A and CENP-B distribution and centromere organization in relation to chromosomal rearrangements: the Hodgkin lymphoma L-1236 and HDLM-2 cell lines, characterized by CENP-A overexpression, detectable CENP-B expression, and chromothripsis, and BJAB, which lacks detectable CENP-B expression and shows normal CENP-A levels. Western blot analysis confirmed the absence of detectable CENP-B protein in BJAB, whereas CENP-B was predominantly detected in the nuclear fractions of HDLM-2 and L-1236.

Cytogenetic analyses revealed alterations in centromere organization across the three cell lines. In BJAB, CENP-A remained localized at centromeres despite the absence of detectable

CENP-B, indicating that CENP-A localization can be maintained independently of CENP-B. In HDLM-2 and L-1236, canonical CENP-A/CENP-B colocalization was accompanied by ectopic CENP-B-positive/CENP-A-negative chromosomal sites. In L-1236, combined CENP-A immunofluorescence and pan-centromeric α-satellite DNA FISH identified CENP-A-positive sites lacking detectable satellite DNA, suggesting the presence of potential neocentromeres. Conversely, ectopic alpha-satellite-DNA signals lacking detectable CENP-A were also observed, suggesting the presence of relocated or inactive centromeric satellite sequences.

CENP-A ChIP-seq analysis provided further evidence of cell-line-specific remodeling of centromeric chromatin, including changes in the extent, distribution and genomic positioning of CENP-A domains. In BJAB, CENP-A was enriched over α-satellite regions lacking CENP-B boxes, which are not centromeric in normal cells, further supporting the uncoupling of CENP-A and CENP-B observed by cytogenetic analysis. In HDLM-2, a non-canonical CENP-A-enriched domain was identified on chromosome 13, extending into HSat1A, a DNA sequence normally associated with pericentromeric regions. This peculiar distribution may represent neocentromere formation in the context of chromosomal rearrangement. Despite sharing a Hodgkin lymphoma background and increased CENP-A expression, HDLM-2 and L-1236 displayed distinct CENP-A distribution patterns, indicating that centromere remodeling does not follow a uniform pattern even within the same lymphoma subtype.

Overall, these findings reveal multiple forms of centromere remodeling in lymphoma cells and support centromere plasticity as an additional feature of highly rearranged cancer genomes. In particular, the coexistence of altered centromere organization with previously reported chromothripsis in HDLM-2 and L-1236 raises the possibility of an association between catastrophic genome restructuring and centromere remodeling, including the establishment of non-canonical CENP-A domains. These findings provide a basis for further investigating the relationship between chromothripsis, centromere reorganization and neocentromere formation in lymphoma.

ABSTRACT

I centromeri sono regioni cromosomiche essenziali che garantiscono un'accurata segregazione dei cromosomi e sono specificati epigeneticamente dalla presenza di CENP-A, la variante centromerica dell'istone H3. Nell'uomo, i centromeri canonici sono tipicamente associati a sequenze di DNA α-satellite e sono legati da CENP-B, l'unica proteina centromerica dotata di attività di legame al DNA specifica per sequenza. Tuttavia, le sequenze satelliti non sono né necessarie né sufficienti a specificare l'identità centromerica e, pertanto, centromeri funzionali possono insorgere sporadicamente al di fuori del contesto di sequenza convenzionale, dando origine a neocentromeri. Questa plasticità epigenetica può assumere particolare rilevanza in contesti patologici come il cancro, considerata l'associazione tra alterazioni della funzione centromerica, instabilità cromosomica e aneuploidia, caratteristiche chiave dei genomi tumorali. Sebbene la deregolazione dell'espressione delle proteine centromeriche sia comune nei tumori e sia stata associata a una prognosi sfavorevole, l'organizzazione genomica ed epigenomica dei centromeri nel cancro rimane ancora in gran parte inesplorata. La cromatina centromerica potrebbe pertanto essere particolarmente suscettibile a fenomeni di rimodellamento durante l'evoluzione tumorale, soprattutto nei tumori caratterizzati da estesi riarrangiamenti genomici, come la cromotripsi.

Questo studio ha analizzato l'organizzazione dei centromeri nelle linee cellulari di linfoma appartenenti al pannello LL-100, concentrandosi sulla relazione tra rimodellamento centromerico, instabilità genomica, estesi riarrangiamenti cromosomici e cromotripsi. L'analisi trascrittomica dei geni codificanti proteine centromeriche nell'intero pannello LL-100 ha evidenziato una ricorrente deregolazione dell'espressione di queste proteine, senza tuttavia mostrare un chiaro pattern di espressione specifico per sottotipo. Diverse proteine centromeriche, tra cui CENP-A, CENP-H, CENP-W, CENP-M, CENP-N e CENP-X, sono sovraespresse in molteplici linee cellulari, mentre CENP-C mostra generalmente livelli di espressione inferiori rispetto ai controlli costituiti da linfonodi sani. È interessante notare che CENP-B è caratterizzata da una marcata variabilità tra le diverse linee cellulari e la sua espressione risulta assente nella linea BJAB di linfoma di Burkitt. Sulla base di questi risultati, sono state selezionate tre linee cellulari per analisi più approfondite della distribuzione di CENP-A e CENP-B e dell'organizzazione centromerica in relazione ai riarrangiamenti cromosomici: le linee di linfoma di Hodgkin L-1236 e HDLM-2, caratterizzate da sovraespressione di CENP-A, presenza di CENP-B e cromotripsi, e BJAB, priva di espressione di CENP-B e caratterizzata da livelli normali di CENP-A. L'analisi

mediante Western blot ha confermato l'assenza della proteina CENP-B in BJAB, mentre CENP-B è stata rilevata prevalentemente nelle frazioni nucleari di HDLM-2 e L-1236.

Analisi citogenetiche successive hanno evidenziato alterazioni dell'organizzazione centromerica in tutte e tre le linee cellulari. In BJAB, CENP-A rimane localizzata ai centromeri nonostante l'assenza di CENP-B rilevabile, indicando che la localizzazione di CENP-A può essere mantenuta indipendentemente da CENP-B. In HDLM-2 e L-1236, CENP-B è presente sia in associazione canonica con CENP-A sia in siti cromosomici ectopici non centromerici. In L-1236, esperimenti di immuno-FISH con un anticorpo anti-CENP-A e una sonda pancentromerica per il DNA α-satellite hanno mostrato centromeri con segnale di CENP-A ma privi di DNA satellite rilevabile, suggerendo la presenza di potenziali neocentromeri. Al contrario, sono stati osservati anche segnali ectopici di DNA α-satellite privi di CENP-A rilevabile, suggerendo la presenza di centromeri inattivati.

L'analisi ChIP-seq di CENP-A ha fornito ulteriori evidenze di un rimodellamento della cromatina centromerica in ciascuna linea cellulare, comprendente variazioni nell'estensione, nella distribuzione e nel posizionamento genomico dei domini di CENP-A. In BJAB sono stati individuati domini di CENP-A in regioni α-satellite prive di CENP-B box e non centromeriche nelle cellule normali. In HDLM-2 è stato identificato sul cromosoma 13 un dominio centromerico non canonico localizzato in corrispondenza del satellite HSat1A, tipicamente pericentromerico nelle cellule normali. Questa peculiare distribuzione potrebbe riflettere la formazione di un neocentromero nel contesto di un riarrangiamento cromosomico. Le linee HDLM-2 e L-1236, nonostante derivino entrambe da linfoma di Hodgkin, mostrano pattern distinti di distribuzione di CENP-A, indicando che il rimodellamento centromerico non segue un pattern uniforme neppure all'interno dello stesso sottotipo di linfoma.

Nel complesso, questi risultati rivelano molteplici forme di rimodellamento centromerico nelle cellule di linfoma e supportano la plasticità centromerica come ulteriore caratteristica dei genomi tumorali altamente riarrangiati. In particolare, la coesistenza di un'organizzazione centromerica alterata con la cromotripsi, precedentemente descritta in HDLM-2 e L-1236, suggerisce una possibile associazione tra una ristrutturazione genomica catastrofica e il rimodellamento centromerico, compresa l'instaurazione di domini non canonici di CENP-A. Questi risultati forniscono una base per ulteriori studi volti a chiarire la relazione tra cromotripsi, riorganizzazione centromerica e formazione di neocentromeri nel linfoma.

TABLE OF CONTENTS

ABBREVIATIONS

APS: ammonium persulfate

BL: Burkitt lymphoma

bp: base pair

BSA: bovine serum albumin

CATD: CENP-A targeting domain

CCAN: constitutive centromere-associated network

CCD: charge-coupled device

CENP: centromere protein

CENP-A: centromere protein A

CENP-B: centromere protein B

ChIP: chromatin immunoprecipitation

ChIP-seq: chromatin immunoprecipitation followed by sequencing

cHL: classical Hodgkin lymphoma

CIN: chromosomal instability

CREST: calcinosis, Raynaud phenomenon, esophageal dysmotility, sclerodactyly and telangiectasia

DAPI: 4′,6-diamidino-2-phenylindole

DNA: deoxyribonucleic acid

DSMZ: German Collection of Microorganisms and Cell Cultures

EBV: Epstein–Barr virus

ECL: enhanced chemiluminescence

EDTA: ethylenediaminetetraacetic acid

ENC: evolutionarily new centromere

FBS: fetal bovine serum

FIMO: Find Individual Motif Occurrences

FISH: fluorescence *in situ* hybridization

H3K4me2: histone H3 lysine 4 dimethylation

H3K9me3: histone H3 lysine 9 trimethylation

HJURP: Holliday junction recognition protein

HOR: higher-order repeat

HP1: heterochromatin protein 1

HRS: Hodgkin and Reed–Sternberg

HSat: human satellite DNA

ICC: International Consensus Classification

INCENP: inner centromere protein

JAK/STAT: Janus kinase/signal transducer and activator of transcription

kb: kilobase

KMN: KNL1–MIS12–NDC80 network

MACS2: Model-based Analysis for ChIP-Seq 2

Mb: megabase

NF-κB: nuclear factor kappa B

NK: natural killer

NP-40: Nonidet P-40

PBS: phosphate-buffered saline

PBST: phosphate-buffered saline containing Tween 20

PCR: polymerase chain reaction

PNA: peptide nucleic acid

RNA: ribonucleic acid

RNA-seq: RNA sequencing

RPKM: reads per kilobase per million mapped reads

RPMI: Roswell Park Memorial Institute medium

SDS: sodium dodecyl sulfate

SDS-PAGE: sodium dodecyl sulfate–polyacrylamide gel electrophoresis

SSC: saline–sodium citrate

T2T: telomere-to-telomere

TBS: Tris-buffered saline

TEMED: N,N,N′,N′-tetramethylethylenediamine

TPM: transcripts per million

WHO: World Health Organization

INTRODUCTION

1. THE CENTROMERE

1.1 The centromere: a paradoxical locus

The centromere is a specialized nucleoprotein structure of eukaryotic chromosomes that plays an essential role in the correct segregation of genetic material during mitosis and meiosis (Choo, 1997). It is recognized as the primary constriction of metaphase chromosomes and represents the chromosomal site where the kinetochore is assembled. Through the kinetochore, centromeres establish interactions with spindle microtubules that are required for the accurate separation and transmission of chromosomes to daughter cells (Choo, 1997; Choo, 2000), as schematically illustrated in Figure 1.

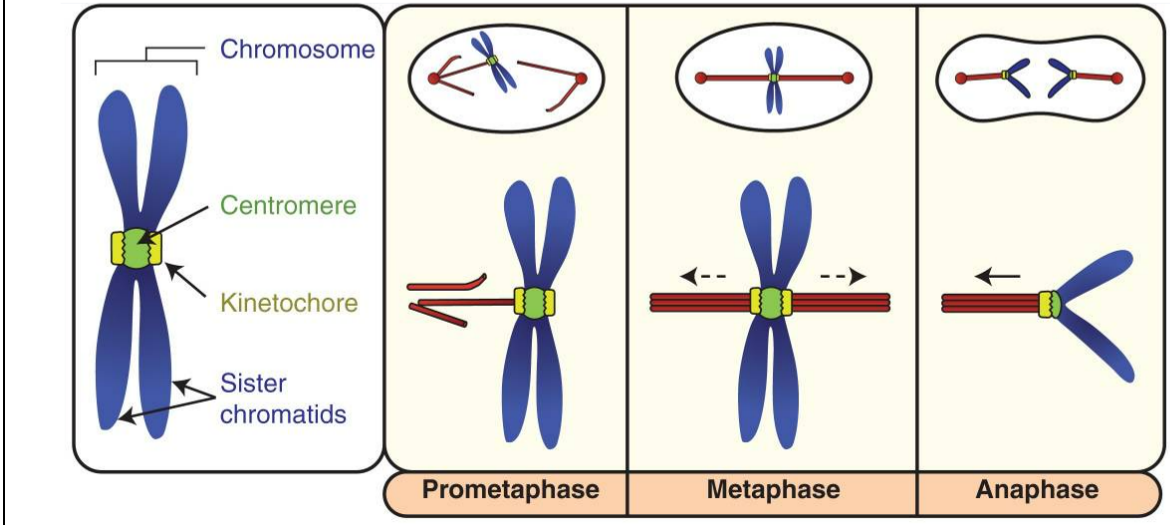

Figure 1. Schematic representation of the centromere and its role during chromosome segregation. The centromere is located at the primary constriction of the chromosome and provides the site for kinetochore assembly. During cell division, spindle microtubules attach to the kinetochore, allowing chromosome alignment at metaphase and the subsequent separation of sister chromatids during anaphase. Adapted from Westhorpe and Straight (2015).

The correct organization and function of the centromere are essential for the maintenance of genomic integrity. Defects affecting centromeric chromatin, kinetochore assembly or centromeric DNA can interfere with chromosome segregation and promote aneuploidy and chromosomal instability. Moreover, centromeric regions themselves are particularly susceptible to replication-associated stress and DNA damage, making the preservation of their structural and functional integrity important for genome stability (Nassar *et al.*, 2023). These alterations are especially relevant in cancer, where chromosome missegregation and

chromosomal instability contribute to the continuous reshaping of the tumor genome (Bakhoum and Cantley, 2018).

Despite this essential and highly conserved function, centromeres show an unusual degree of evolutionary variability. The DNA sequences associated with centromeric regions can differ substantially among species and can evolve rapidly, whereas the chromosome-segregation function of the centromere remains conserved. This apparent contradiction between functional conservation and rapid evolution of centromeric DNA sequences has been defined as the "centromere paradox" (Henikoff *et al.*, 2001). Rather than being universally specified by a particular DNA sequence, centromere identity in most eukaryotes is epigenetically specified, depending primarily on the organization and inheritance of specialized centromeric chromatin.

1.2 The epigenetic determinant of centromere identity: CENP-A

CENP-A was identified as a centromere-associated protein recognized by autoantibodies present in sera from patients with CREST (calcinosis, Raynaud phenomenon, esophageal dysmotility, sclerodactyly and telangiectasia) syndrome and was subsequently shown to be associated with nucleosomal particles (Palmer *et al.*, 1987). It is a centromere-specific variant of histone H3 and represents the main epigenetic determinant of centromere identity in most eukaryotes. By replacing canonical H3 in a subset of nucleosomes, CENP-A establishes a specialized chromatin domain that marks the position of the active centromere and provides the foundation for kinetochore assembly (Black *et al.*, 2007; McKinley and Cheeseman, 2016).

Although CENP-A retains the general histone-fold organization of histone H3, it contains specific structural features that distinguish centromeric nucleosomes from conventional chromatin. In particular, the CENP-A targeting domain (CATD), encompassing loop L1 and the α2 helix of the histone-fold domain, is important for the specific recognition and centromeric targeting of CENP-A. Other regions of the protein contribute to interactions with centromere-associated proteins, allowing the CENP-A nucleosome to function not only as a structural component of centromeric chromatin, but also as a molecular platform for the assembly of the centromere–kinetochore machinery (Black *et al.*, 2004; French and Straight, 2013).

Because centromere identity depends on the presence of CENP-A, its deposition and inheritance must be tightly regulated during the cell cycle.

1.3 CENP-A deposition and maintenance

The incorporation of newly synthesized CENP-A into centromeric chromatin is mediated by a dedicated loading machinery. A central component of this process is HJURP (Holliday Junction Recognition Protein), a specific histone chaperone that recognizes the CATD of CENP-A and binds CENP-A/H4 before nucleosome assembly. HJURP recruitment to centromeres is coordinated by the Mis18 complex, which helps define where and when new CENP-A can be incorporated (Dunleavy *et al.*, 2009; Foltz *et al.*, 2009; Pukało *et al.*, 2025).

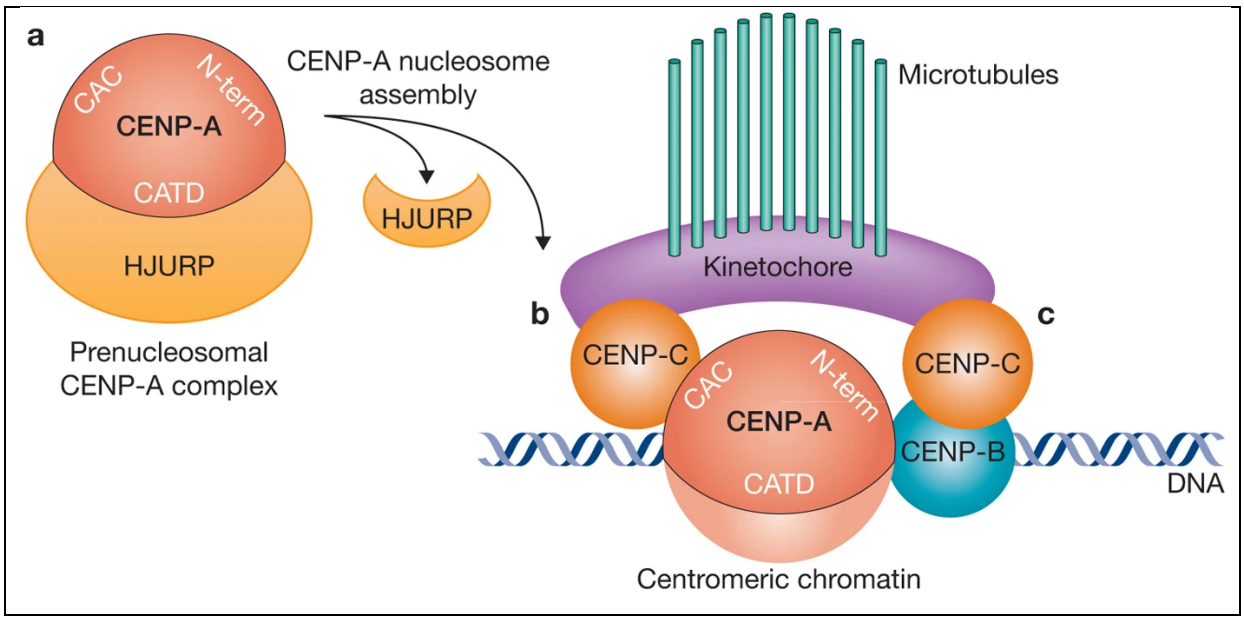

Figure 2. Selected molecular interactions involved in CENP-A deposition and centromere function. (a) The CENP-A targeting domain (CATD) interacts with the histone chaperone HJURP and contributes to the centromeric deposition of CENP-A. (b) Once incorporated into centromeric chromatin, the C-terminal region of CENP-A interacts with CENP-C, promoting kinetochore assembly. (c) The N-terminal region of CENP-A may provide an additional pathway for CENP-C recruitment involving CENP-B. Reproduced from French and Straight (2013).

CENP-A deposition is also tightly linked to cell-cycle progression. During DNA replication, pre-existing CENP-A nucleosomes are distributed between the two sister chromatids, reducing the amount of CENP-A at each newly replicated centromere. In human cells, this pool is restored after mitotic exit, mainly during early G1, when newly synthesized CENP-A is deposited through the HJURP/Mis18 pathway (Jansen *et al.*, 2007; Dunleavy *et al.*, 2009; Foltz *et al.*, 2009). This regulated replenishment enables the CENP-A domain to be propagated from one cell generation to the next, providing a mechanism for the epigenetic maintenance of centromere identity.

1.4 Centromere-associated proteins and kinetochore assembly

CENP-A-containing chromatin provides the platform for the assembly of the kinetochore, the multiprotein structure that connects chromosomes to spindle microtubules during cell division. A central component of this organization is the constitutive centromere-associated network (CCAN), which remains associated with centromeric chromatin throughout the cell cycle and forms the inner kinetochore. The CCAN is composed of several interconnected CENP proteins that link the CENP-A nucleosome to the outer kinetochore machinery (Foltz *et al.*, 2006; McKinley and Cheeseman, 2016) (Figure 3).

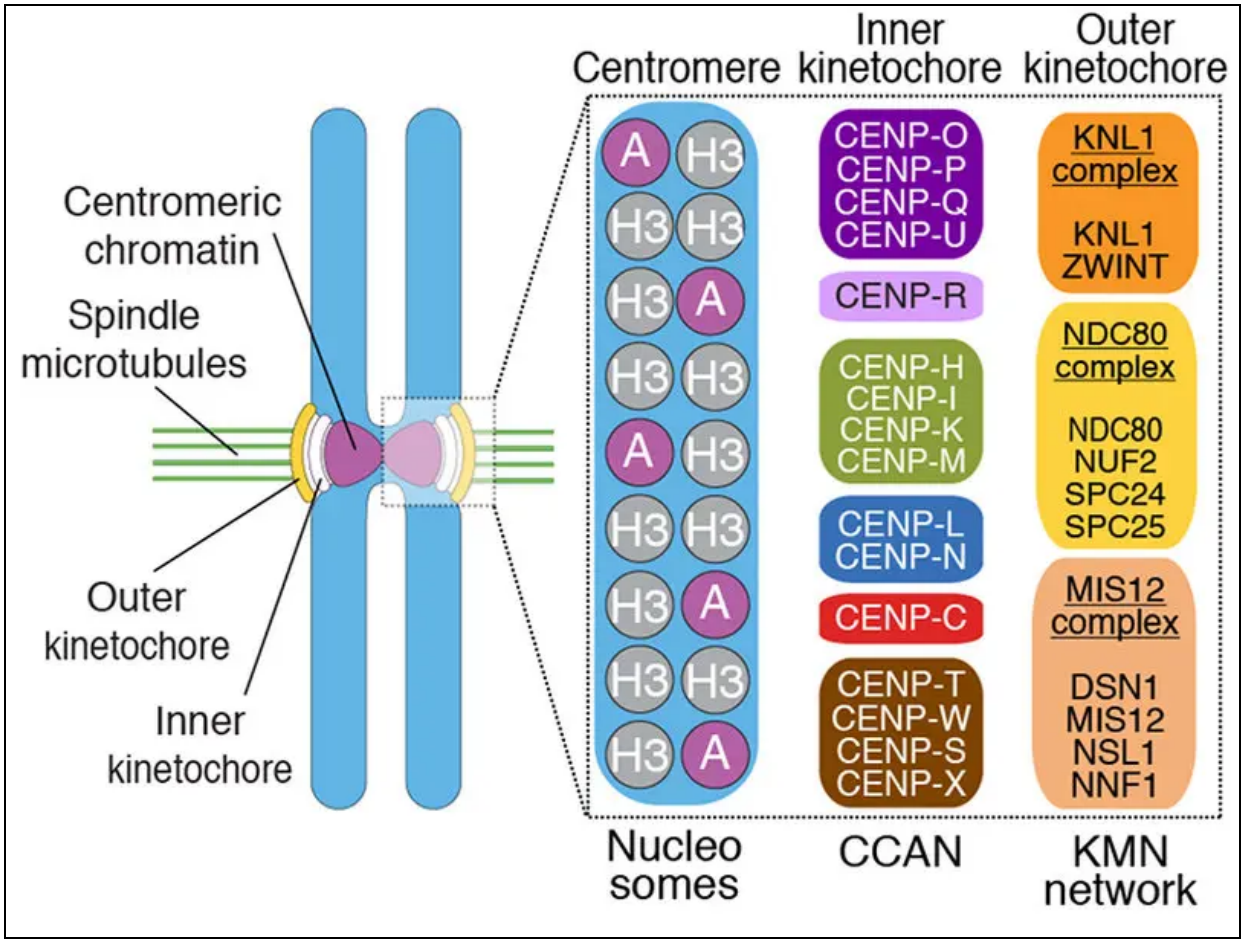

Figure 3. Schematic representation of centromere and kinetochore organization. CENP-A-containing chromatin provides the foundation for assembly of the constitutive centromere-associated network (CCAN), which connects the centromeric chromatin to the outer kinetochore. The KMN network forms a major component of the outer kinetochore and mediates the interaction with spindle microtubules. Adapted from Basilico *et al.* (2014).

Among these proteins, CENP-C and CENP-N directly recognize CENP-A-containing nucleosomes and contribute to the recruitment and organization of additional centromeric components (Carroll *et al.*, 2009; Carroll *et al.*, 2010). The CENP-T/W/S/X complex provides

an additional connection between centromeric chromatin and the outer kinetochore (Pesenti *et al.*, 2022). Together, these pathways promote recruitment of the KMN network, composed of KNL1, the MIS12 complex and the NDC80 complex. The KMN network forms a major microtubule-binding interface and is essential for stable chromosome–spindle attachment and accurate chromosome segregation (Cheeseman *et al.*, 2006). More recent structural studies have further revealed the highly integrated organization of the human CCAN and its connections with the outer kinetochore (Pesenti *et al.*, 2022).

In addition to the CCAN components, CENP-B represents a distinct sequence-dependent centromeric protein that contributes to centromeric chromatin organization and is discussed in detail in the following section.

1.5 CENP-B and centromeric DNA recognition

CENP-B is unusual among centromeric proteins because it recognizes centromeric DNA in a sequence-specific manner. It binds a 17-bp motif known as the CENP-B box, which is present within several mammalian satellite DNA families. Early biochemical studies demonstrated direct binding of CENP-B to this sequence and suggested that the protein may contribute to the higher-order organization of the human α-satellite DNA (Muro *et al.*, 1992).

The CENP-B/CENP-B-box interaction also contributes to the establishment and stability of centromeric chromatin. Experiments using human artificial chromosomes showed that intact CENP-B boxes are important for the de novo assembly of centromeric chromatin on α-satellite DNA (Ohzeki *et al.*, 2002). Later work demonstrated that sequence-dependent binding of CENP-B helps maintain appropriate levels of CENP-C at the centromere and improves chromosome segregation fidelity (Fachinetti *et al.*, 2015).

More recent studies have refined this view by showing that CENP-B does not simply bind centromeric DNA, but can also influence its chromatin environment. CENP-A-containing chromatin is relatively dynamic and accessible, which facilitates CENP-B binding to its target sequence. In turn, bound CENP-B can further increase chromatin accessibility and promote DNA unwrapping around neighboring nucleosomes. These observations suggest a cooperative relationship between CENP-A and CENP-B in shaping the physical organization of centromeric chromatin (Nagpal *et al.*, 2023).

However, CENP-B is not an absolute requirement for centromere activity. The clearest evidence comes from naturally occurring centromeres in which CENP-A and kinetochore function are maintained despite the absence of CENP-B. Studies from our laboratory have

shown that several equid centromeres lack CENP-B at the functional centromeric domain while still recruiting CENP-A and CENP-C. In these species, CENP-B-box-containing satellite sequences can instead be found outside the active centromeric core, demonstrating that the positions of CENP-A, CENP-B and satellite DNA do not necessarily coincide (Cappelletti *et al.*, 2025b). This finding extends earlier work from the same laboratory showing a natural uncoupling between satellite DNA and centromeric function in the genus *Equus* (Piras *et al.*, 2010).

The opposite situation is also possible: CENP-B binding does not by itself establish an active centromere. A de novo α-satellite insertion identified in the human genome recruited CENP-B but lacked detectable CENP-A and did not acquire centromeric activity (Giannuzzi *et al.*, 2021). Taken together, these observations indicate that CENP-B has an important role in the organization and stability of many satellite-based centromeres, but its localization is not sufficient to define centromere identity. Instead, the functional centromere remains primarily associated with the presence and maintenance of CENP-A chromatin.

1.6 Centrochromatin: epigenetic identity and transcriptional activity

The functional centromere is associated with a specialized chromatin domain that differs from both conventional euchromatin and the surrounding heterochromatin. This domain, often referred to as centrochromatin, contains CENP-A nucleosomes interspersed with canonical histone H3 nucleosomes and is characterized by a relatively permissive epigenetic environment. In particular, active centromeric chromatin is enriched in histone modifications such as H3K4me2, which distinguish the centromeric core from the more repressive chromatin surrounding it (Sullivan and Karpen, 2004) (Figure 4).

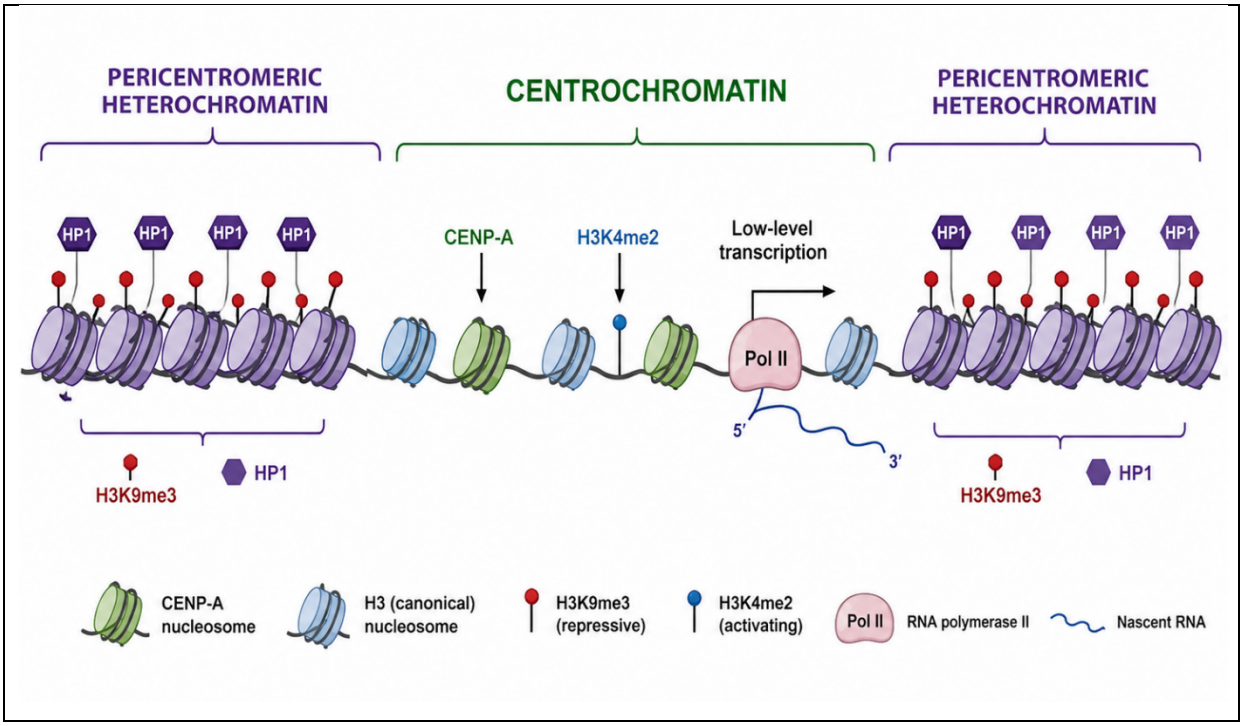

Figure 4. Schematic representation of centrochromatin and surrounding pericentromeric heterochromatin. The centromeric core is composed of specialized chromatin containing CENP-A nucleosomes together with canonical H3 nucleosomes and is associated with permissive features such as H3K4me2 and low levels of RNA polymerase II-dependent transcription. In contrast, the surrounding pericentromeric regions are characterized by a more compact heterochromatic organization enriched in H3K9me3 and HP1. Schematic created based on concepts described by Sullivan and Karpen (2004), McNulty *et al.* (2017) and Naughton *et al.* (2022).

In contrast, the chromatin flanking the centromeric core corresponds to pericentromeric heterochromatin, which is more compact and is enriched in repressive features such as H3K9me3 and HP1. The coexistence of these two domains—a specialized centromeric core and repressive surrounding chromatin—is important for centromere stability. Experimental studies have shown that altering this balance, either by promoting excessive heterochromatin formation within the centromeric domain or by interfering with the chromatin environment required for centromere function, can impair CENP-A maintenance and kinetochore activity (Sullivan and Karpen, 2004; Molina *et al.*, 2016).

Although centromeres were long considered transcriptionally silent because of their repetitive DNA content, it is now clear that low levels of transcription are a normal feature of active centromeres. In human cells, α-satellite DNA can give rise to centromeric transcripts that associate with centromere proteins such as CENP-A and CENP-C and contribute to centromere maintenance (McNulty *et al.*, 2017). This transcription is tightly regulated and appears to support the organization of centromeric chromatin rather than oppose it. Consistent

with this idea, centromere repositioning in a human neocentromere was shown to be accompanied by RNA polymerase II recruitment and a more open chromatin structure at the active centromeric domain (Naughton *et al.*, 2022). Together, these observations indicate that centromere identity depends not only on CENP-A chromatin itself, but also on a specific epigenetic and transcriptional environment.

1.7 Centromeric and pericentromeric satellite DNA

Human centromeric and pericentromeric regions contain several classes of repetitive DNA, of which α-satellite DNA is the major centromeric component (Figure 5). α-Satellite DNA consists of monomers of approximately 171 bp. Individual monomers are often organized into larger units known as higher-order repeats (HORs), which are themselves repeated many times to form arrays that can extend over several megabases. The organization and sequence of these arrays differ among chromosomes and can also vary considerably between individuals (Willard and Waye, 1987; Altemose *et al.*, 2022; Logsdon *et al.*, 2024).

Not all α-satellite arrays on a chromosome are functionally equivalent. Although multiple HOR arrays may occur on the same chromosome, CENP-A is generally concentrated within one array, referred to as the active HOR array, where the functional kinetochore is assembled. Other related HOR arrays that do not show comparable CENP-A association are considered inactive. Even within the active HOR, CENP-A does not necessarily occupy the entire repeat array but can be restricted to a smaller domain. Complete T2T-CHM13 maps have made it possible to distinguish these active and inactive arrays and to define their relationship with CENP-A chromatin at sequence-level resolution (Altemose *et al.*, 2022).

Outside the large and relatively homogeneous HOR arrays, α-satellite sequences can become more divergent and may occur as monomeric α-satellite DNA. Human pericentromeric regions also contain several other satellite families, including HSat1A, HSat1B, HSat2 and HSat3, as well as smaller blocks of beta- and gamma-satellite DNA (Figure 5). These repeat families have different sequence compositions and chromosome distributions and are commonly associated with the heterochromatic regions surrounding the functional centromere. Their highly repetitive organization meant that large portions of these regions were absent or poorly represented in earlier human reference genomes. The completion of the T2T-CHM13 assembly therefore provided the first continuous view of the organization of α-satellite and other satellite families across human centromeric and pericentromeric regions (Altemose *et al.*, 2022).

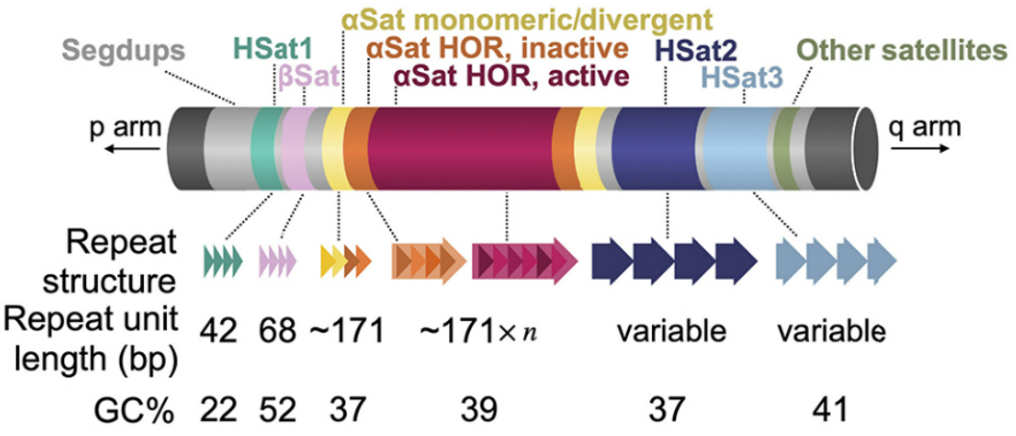

Figure 5. Organization of human centromeric and pericentromeric satellite DNA. Schematic representation of the principal repeat classes present in human centromeric regions. The centromeric core is composed mainly of α-satellite DNA arranged as higher-order repeats (HORs), which can be distinguished into active and inactive arrays, and is flanked by monomeric/divergent α-satellite DNA and other satellite families, including HSat1, HSat2, HSat3, beta satellite, and additional repetitive sequences. Segmental duplications are also present in pericentromeric regions. The lower part of the scheme summarizes the repeat structure and approximate repeat-unit length of the indicated sequence classes. Adapted from Altemose *et al.* (2022).

Despite their close association with centromeres, satellite repeats do not themselves determine centromere identity. Instead, they provide a genomic environment that can contribute to centromeric chromatin organization, while the position of the functional centromere is primarily defined by CENP-A. This distinction is supported by the existence of functional centromeres lacking canonical satellite DNA in humans (Voullaire *et al.*, 1993; Murillo-Pineda *et al.*, 2021), and in other mammals, where satellite-free centromeres have been identified in several species, particularly within the genus *Equus* (Piras *et al.*, 2010; Nergadze *et al.*, 2018; Cappelletti *et al.*, 2022). Recent complete human centromere assemblies have also revealed substantial variation in α-satellite HOR array size, sequence organization and kinetochore position among individuals (Logsdon *et al.*, 2024). These non-canonical forms of centromere organization are discussed in greater detail in the following section on neocentromeres.

2. NEOCENTROMERES

2.1 Neocentromeres and the process of centromerization

Neocentromeres are functional centromeres that arise at genomic locations that were not previously associated with centromeric activity. They can assemble a functional kinetochore and support chromosome segregation despite frequently lacking the large arrays of α-satellite DNA that characterize canonical human centromeres. Their existence demonstrates that centromere identity is not determined solely by the underlying DNA sequence, but can instead be established and maintained through epigenetic mechanisms, principally through the formation of a stable CENP-A-containing chromatin domain (Amor and Choo, 2002; Warburton, 2004).

The establishment of centromeric activity at a previously non-centromeric locus is referred to as centromerization (Choo, 2000). A central event in this process is the establishment and stable propagation of a CENP-A-containing chromatin domain, which provides the foundation for recruitment of the constitutive centromere-associated network and assembly of a functional kinetochore. Formation of a new centromere can also involve remodeling of the local chromatin environment rather than simply the deposition of CENP-A at an ectopic site. Experimental induction of human neocentromere formation showed that CENP-A establishment was accompanied by loss of local H3K9me3, recruitment of cohesin and RNA polymerase II, and subsequent maturation of centromeric functions during long-term propagation (Murillo-Pineda *et al.*, 2021). Centromerization can therefore be considered a progressive epigenetic process through which a previously non-centromeric region acquires and stably maintains the chromatin and protein organization required for chromosome segregation. This epigenetic distinction is illustrated in Figure 6: whereas the canonical centromere contains CENP-A chromatin associated with α-satellite DNA and CENP-B, the experimentally established neocentromere contains CENP-A in the absence of α-satellite DNA and CENP-B binding (Carty and Dunleavy, 2021).

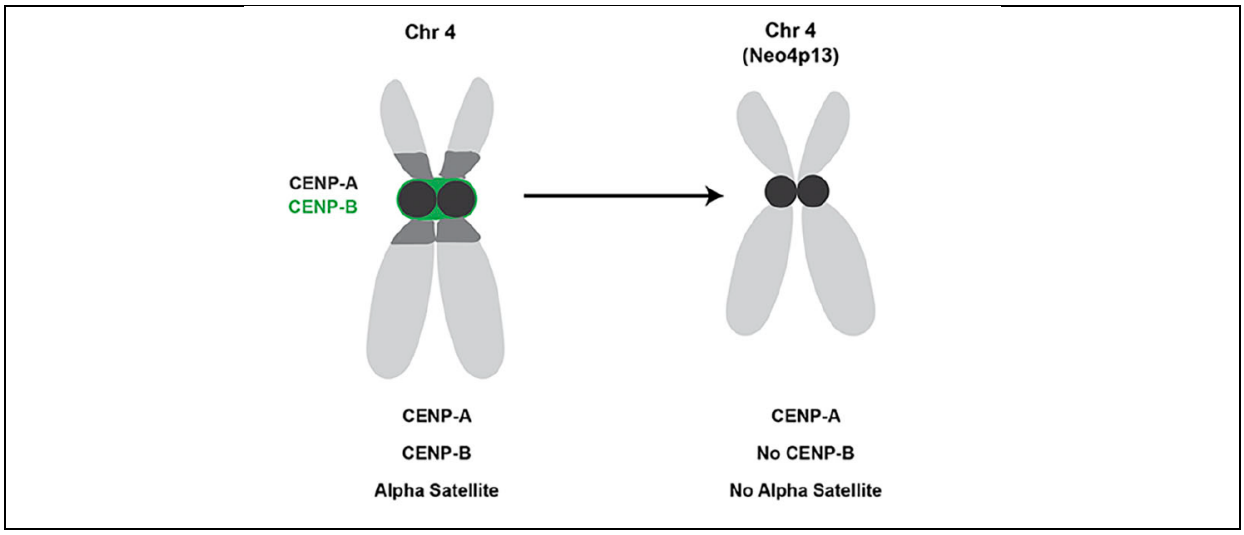

Figure 6. Comparison between canonical centromere and neocentromere organization. The canonical centromere is characterized by CENP-A-containing chromatin associated with α-satellite higher-order repeats, CENP-B binding and surrounding pericentromeric heterochromatin. In contrast, the experimentally established neocentromere contains CENP-A but lacks α-satellite DNA and CENP-B binding, illustrating the epigenetic basis of centromere establishment. Adapted from Carty and Dunleavy (2021), based on Murillo-Pineda *et al.* (2021).

Neocentromeres can arise in different biological contexts and are commonly discussed as human clinical neocentromeres or evolutionarily new centromeres (ENCs). Although both involve the establishment of centromeric activity at a new genomic position, they differ in their biological context and long-term fate. Evolutionarily new centromeres become stably inherited and fixed during evolution, whereas human clinical neocentromeres arise in association with individual chromosome abnormalities and can also occur in highly rearranged cancer genomes. These two contexts are considered separately in the following sections.

2.2 Evolutionarily new centromeres and centromere repositioning

Evolutionarily new centromeres (ENCs) arise when centromeric function shifts to a new chromosomal position and the new centromere subsequently becomes stably inherited within a lineage. This process, known as centromere repositioning, demonstrates that centromere position can change during evolution without necessarily requiring major alteration of the surrounding gene order (Montefalcone *et al.*, 1999; Rocchi *et al.*, 2012). Studies in *Equus* have provided important examples of this phenomenon. Work from our laboratory identified and characterized numerous satellite-free centromeres in equids, showing that newly established centromeres can function in the absence of detectable satellite DNA and supporting a model in which repetitive sequences may be acquired only after centromeric

activity has been established (Wade *et al.*, 2009; Piras *et al.*, 2010; Nergadze *et al.*, 2018; Cappelletti *et al.*, 2022; Piras *et al.*, 2022; Piras *et al.*, 2023; Cappelletti *et al.*, 2025a). Beyond equids, a polymorphic α-satellite-free neocentromere has been characterized in orangutans and is present in both the Bornean orangutan (*Pongo pygmaeus*) and the Sumatran orangutan (*Pongo abelii*) (Tolomeo *et al.*, 2017; de Gennaro *et al.*, 2026). In addition, non-canonical centromeres with a markedly different repetitive architecture from the canonical α-satellite-rich centromeres of great apes were characterized in the eastern hoolock gibbon (Hartley *et al.*, 2025). These observations provide evolutionary evidence for the positional and sequence plasticity of centromere identity and its independence from canonical satellite DNA.

2.3 Human clinical neocentromeres

Human clinical neocentromeres have been identified mainly during cytogenetic analysis of individuals carrying structurally abnormal chromosomes. In many constitutional cases, chromosome rearrangements generate a fragment that lacks a conventional centromere; acquisition of neocentromeric activity can stabilize this otherwise acentric material and allow its transmission during cell division. Neocentromeres are therefore frequently associated with supernumerary marker chromosomes, particularly those produced by inverted duplications or other complex rearrangements. The clinical consequences are generally related to the accompanying chromosomal imbalance, such as duplication or deletion of genomic material, rather than to the presence of the neocentromere itself (Amor and Choo, 2002; Marshall *et al.*, 2008).

Most characterized human clinical neocentromeres lack detectable α-satellite DNA but nevertheless form a primary constriction and recruit the centromere proteins required for a functional kinetochore. They have been identified at different positions throughout the human genome, and even neocentromeres arising within the same chromosomal region can occupy different CENP-A-associated domains. This variability argues against a unique DNA-sequence requirement for neocentromere formation, although recurrent cytogenetic hotspots have been identified in particular genomic regions (Amor and Choo, 2002; Marshall *et al.*, 2008).

Acquired neocentromeres have also been identified in cancer, where they can contribute to the stabilization of highly rearranged chromosomes. They have been described in solid tumors, including liposarcoma and lung sarcomatoid carcinoma, as well as in hematological malignancies such as angioimmunoblastic T-cell lymphoma (Blom *et al.*, 2010; Macchia *et al.*, 2018). These observations demonstrate that neocentromere formation is not restricted to

constitutional chromosome abnormalities but can also occur during cancer-associated genome restructuring. The mechanisms and biological consequences of this process in malignant cells, however, remain incompletely understood.

3. CENTROMERE DYSFUNCTION AND CANCER

3.1 Centromere dysfunction and chromosomal instability

Faithful chromosome segregation depends on the correct organization of centromeric chromatin, assembly of the kinetochore and establishment of stable interactions between chromosomes and the mitotic spindle. Alterations affecting centromere structure, kinetochore–microtubule attachment or sister-chromatid cohesion can compromise this process and result in segregation defects, including lagging chromosomes, chromosome bridges and unequal chromosome distribution between daughter cells (Manning *et al.*, 2010; Funk *et al.*, 2016). When these errors persist over successive cell divisions, they can contribute to chromosomal instability (CIN), a common feature of many cancers that promotes continuous changes in chromosome number and structure (Bakhoum and Cantley, 2018).

Repeated chromosome segregation errors can also generate aneuploidy, defined as an abnormal number of chromosomes. Although aneuploidy and CIN are closely related, they describe different concepts: aneuploidy represents a cellular state, whereas CIN refers to an ongoing elevated rate of chromosome segregation errors during cell proliferation. This continuous generation of new karyotypic configurations can increase genetic heterogeneity within a tumor and contribute to cancer evolution (Figure 7) (Thompson and Compton, 2008; Ben-David and Amon, 2020).

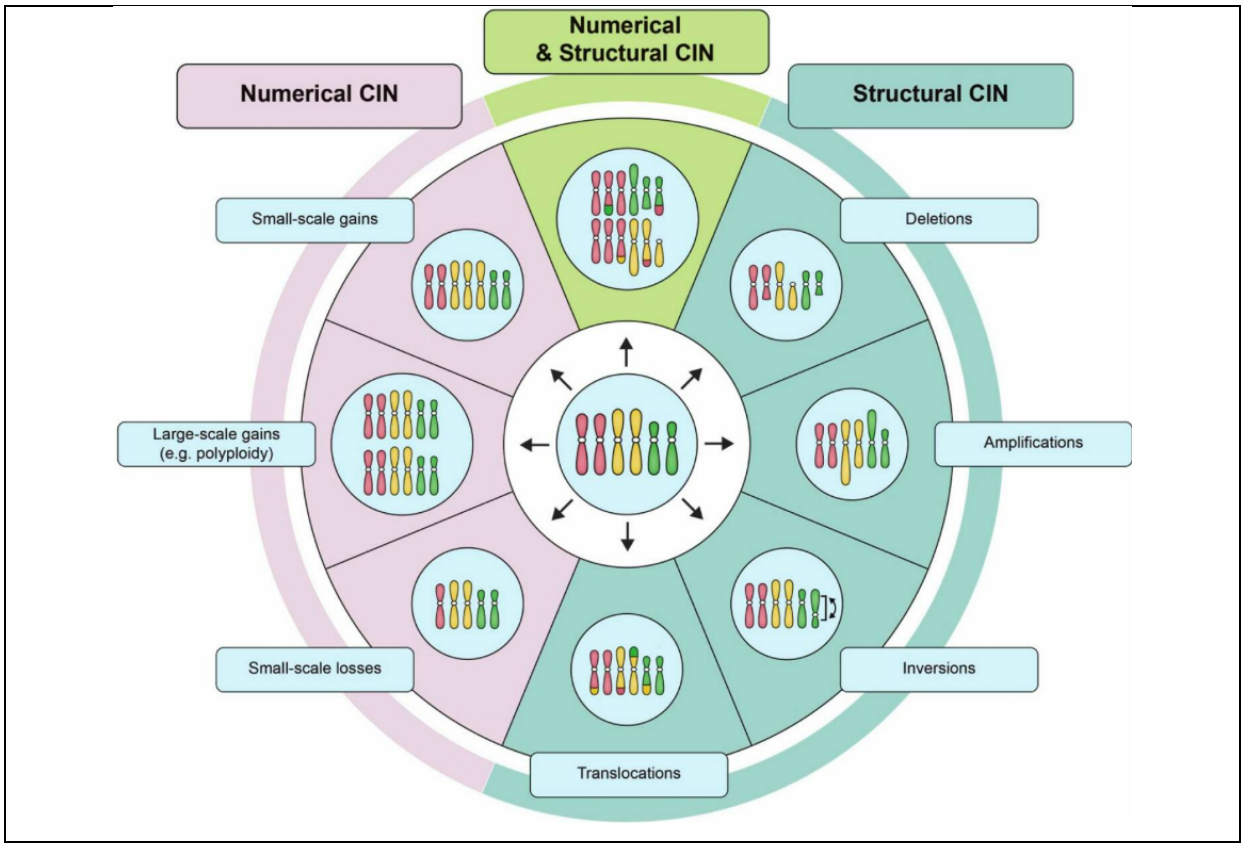

Figure 7. Numerical and structural chromosomal instability. Schematic representation of chromosome alterations generated by chromosomal instability (CIN). Numerical CIN results from gains or losses of whole chromosomes, whereas structural CIN produces chromosome fragments, rearrangements and other structural abnormalities. Continued generation of these alterations contribute to karyotypic heterogeneity within a cell population. Adapted from Funk *et al.* (2016).

Importantly, the consequences of chromosome segregation errors extend beyond changes in chromosome number. Lagging chromosomes, chromosome bridges or missegregated chromosome fragments can undergo additional damage or become physically separated from the main chromosome mass during mitotic exit. Such events provide an important link between mitotic defects and more extensive forms of genome instability, including the formation of micronuclei and complex chromosome rearrangements, which will be discussed in the following section.

3.2 Complex chromosome rearrangements and chromothripsis

Chromosomal instability can affect not only chromosome numbers but also chromosome structure, generating DNA breaks and complex rearrangements. An extreme form of structural genome reorganization is chromothripsis, originally described as the occurrence of tens to hundreds of clustered chromosome rearrangements during a catastrophic genomic event. Chromothripsis typically affects one or a few chromosomes or chromosomal regions and is

characterized by a high concentration of breakpoints together with oscillations between a limited number of copy-number states. The proposed model involves extensive chromosome fragmentation followed by abnormal reassembly of the resulting DNA fragments, producing highly rearranged derivative chromosomes (Stephens *et al.*, 2011) (Figure 8).

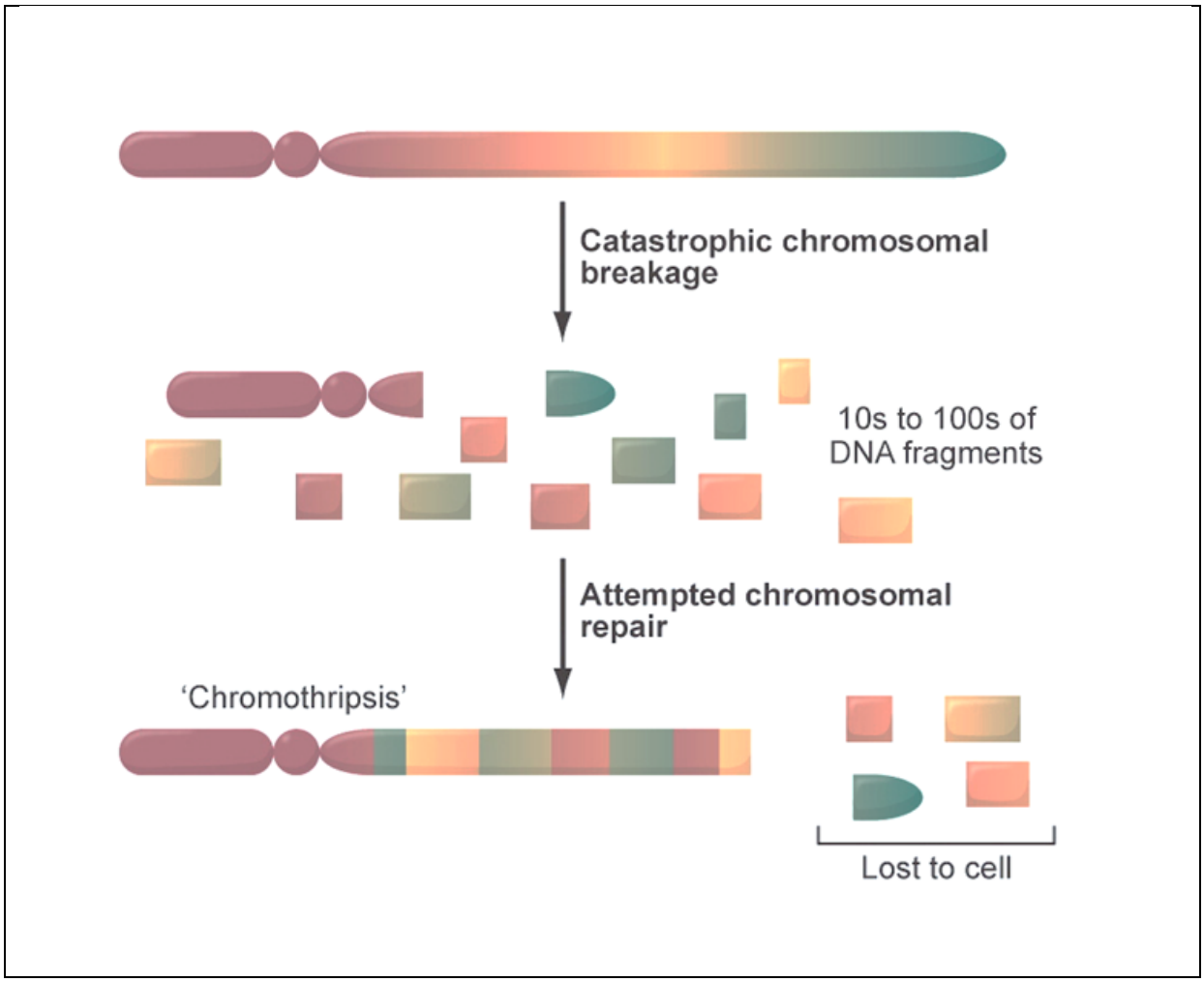

Figure 8. Schematic representation of chromothripsis. Catastrophic chromosome breakage generates multiple DNA fragments, which are subsequently reassembled in an altered order, while some fragments may be lost, resulting in a highly rearranged chromosome. Adapted from the graphical abstract of Stephens *et al.* (2011).

Chromothripsis can arise through different forms of chromosome damage and abnormal repair. Experimental studies have shown that errors occurring during cell division, including chromosome bridge formation and subsequent breakage, can rapidly generate complex rearrangements resembling those observed in cancer genomes (Umbreit *et al.*, 2020). These findings indicate that catastrophic chromosome restructuring can develop from mitotic errors and contribute to the accumulation of structural genome abnormalities.

Large-scale whole-genome analyses have shown that chromothripsis is widespread across human cancers and can occur at high frequency in some tumor types. These events can

contribute to cancer development through extensive structural reorganization, including oncogene amplification and the inactivation or loss of tumor-suppressor genes (Cortés-Ciriano *et al.*, 2020). Chromothripsis therefore represents an important example of how chromosomal instability can produce rapid and extensive remodeling of the cancer genome.

Importantly, extensive genome restructuring can also be accompanied by changes in centromere organization. In well-differentiated and dedifferentiated liposarcomas, neochromosomes were proposed to arise following chromothripsis and to undergo extensive amplification and rearrangement through subsequent breakage–fusion–bridge cycles. Progressive centromeric disruption eventually resulted in a crisis that was resolved through neocentromere formation or acquisition of an existing centromere (Garsed *et al.*, 2014). This provides an important example of the association between catastrophic genome restructuring and centromere reorganization in cancer.

3.3 Dysregulation of centromere and kinetochore proteins in cancer

Altered expression of genes encoding centromere and kinetochore proteins is frequently observed in cancer and has been associated with chromosomal instability and poor clinical outcome (Zhang *et al.*, 2016). Among these proteins, CENP-A has received particular attention because of its central role in centromere identity. CENP-A is overexpressed in several tumor types, and experimental overexpression can promote its ectopic incorporation at non-centromeric chromatin (Lacoste *et al.*, 2014; Zhang *et al.*, 2016). CENP-A overexpression can also increase chromosome segregation defects, aneuploidy and karyotypic heterogeneity (Shrestha *et al.*, 2021). These findings indicate that changes in CENP-A abundance can affect both centromere organization and genome stability.

The role of CENP-B in cancer is less clearly defined and appears to be more context-dependent. Unlike CENP-A, CENP-B binds specific CENP-B-box sequences within α-satellite DNA and contributes to centromeric chromatin organization, but its relationship with cancer-associated phenotypes is less well characterized. Other centromere-associated proteins can also be dysregulated in tumors; for example, increased CENP-H expression has been reported in several cancer types and associated with poor prognosis and lymph-node metastasis (Li *et al.*, 2025). Together, these observations suggest that cancer-associated centromere dysfunction may involve alterations in several components of the centromere–kinetochore machinery.

Centromere-associated proteins have therefore also been investigated as potential biomarkers. Increased CENP-A expression has been associated with relapse in selected breast cancer cohorts, while broader centromere–kinetochore gene-expression signatures have been linked to patient survival and treatment response (McGovern *et al.*, 2012; Zhang *et al.*, 2016). These studies support the potential prognostic and predictive relevance of centromere-associated alterations and provide a basis for investigating whether similar patterns occur in lymphoma.

4. LYMPHOMAS

4.1 Classification and biological heterogeneity of lymphomas

Lymphomas comprise a heterogeneous group of lymphoid malignancies arising from cells at different stages of B-, T- or natural killer (NK)-cell differentiation. They include Hodgkin lymphomas and a large and diverse group of B-cell, T-cell and NK-cell neoplasms with distinct biological and clinical features (Alaggio *et al.*, 2022). Classical Hodgkin lymphoma is characterized by the presence of Hodgkin and Reed–Sternberg cells within a prominent inflammatory microenvironment (Küppers, 2009). Modern classification integrates morphological, immunophenotypic, genetic and clinical characteristics, and both the fifth edition of the World Health Organization (WHO) classification and the International Consensus Classification (ICC) recognize numerous biologically distinct lymphoma entities (Alaggio *et al.*, 2022; Campo *et al.*, 2022).

Considerable molecular heterogeneity is also present within individual lymphoma entities. Tumors originating from related lymphoid cell populations can display markedly different combinations of gene mutations, copy-number alterations, chromosomal rearrangements and gene-expression profiles, contributing to differences in tumor biology and clinical behavior (Ware *et al.*, 2021; Mansouri *et al.*, 2022). Advances in cytogenetic and genomic approaches have revealed recurrent and subtype-specific genetic alterations across lymphoid malignancies, and these features now contribute to lymphoma classification, diagnosis and risk stratification (Campo *et al.*, 2022; Mansouri *et al.*, 2022).

Among the different lymphoma entities, Hodgkin lymphoma and Burkitt lymphoma represent biologically and genetically distinct B-cell malignancies and are particularly relevant to the present study. Classical Hodgkin lymphoma originates predominantly from germinal-center-derived B cells that undergo extensive transcriptional reprogramming, whereas Burkitt lymphoma is a highly aggressive B-cell neoplasm typically characterized by MYC

rearrangement (Küppers, 2009; Woroniecka *et al.*, 2022). Their contrasting cellular and genomic backgrounds provide useful models for investigating differences in genome instability and centromere organization.

4.2 Hodgkin and Burkitt lymphomas

Classical Hodgkin lymphoma (cHL) is a B-cell malignancy characterized by a small population of malignant Hodgkin and Reed–Sternberg (HRS) cells surrounded by a prominent inflammatory and immune-cell microenvironment. HRS cells originate predominantly from germinal-center B cells but undergo extensive transcriptional reprogramming and lose much of the conventional B-cell gene-expression program. Their survival is supported by constitutive activation of signaling pathways including NF-κB and JAK/STAT, together with interactions with the surrounding tumor microenvironment (Weniger and Küppers, 2021). Classical Hodgkin lymphoma is further divided into nodular sclerosis, mixed cellularity, lymphocyte-rich and lymphocyte-depleted subtypes, which differ in their histological and biological characteristics (Alaggio *et al.*, 2022).

Burkitt lymphoma (BL), in contrast, is a highly aggressive mature B-cell lymphoma with a germinal-center B-cell phenotype and an exceptionally high proliferative rate (Alaggio *et al.*, 2022**)**. Its major molecular hallmark is deregulation of the MYC oncogene through a chromosomal rearrangement involving an immunoglobulin locus. The most frequent alteration is t(8;14)(q24;q32), involving MYC and the immunoglobulin heavy-chain locus, while less frequent rearrangements involve the immunoglobulin light-chain loci on chromosomes 2 or 22 (Woroniecka *et al.*, 2022).

Hodgkin and Burkitt lymphomas therefore represent biologically distinct B-cell malignancies despite their shared origin from germinal-center B cells. Hodgkin lymphoma is characterized by strongly reprogrammed HRS cells within a complex cellular microenvironment, whereas Burkitt lymphoma retains a mature B-cell phenotype and is dominated by MYC-driven proliferation.

The present study focused on three lymphoma cell lines representing distinct disease backgrounds. HDLM-2 (DSMZ ACC 17) was established from the pleural effusion of a 74-year-old man with stage IV nodular-sclerosing Hodgkin lymphoma (Drexler *et al.*, 1986). L-1236 (DSMZ ACC 530) was derived from the peripheral blood of a 34-year-old man with refractory stage IV mixed-cellularity Hodgkin lymphoma (Wolf *et al.*, 1996). BJAB (DSMZ ACC 757) was established from the inguinal tumor of a 5-year-old girl with Burkitt

lymphoma (Menezes *et al.*, 1975). All three cell lines are EBV-negative. These models therefore represent biologically distinct Hodgkin and Burkitt lymphoma backgrounds for investigating centromere organization.

4.3 Genome instability and chromothripsis in lymphoma

Genome instability is an important feature of lymphoid malignancies and can involve both numerical and structural chromosome abnormalities. In classical Hodgkin lymphoma, Hodgkin and Reed–Sternberg cells frequently display highly abnormal karyotypes characterized by aneuploidy, copy-number alterations and complex chromosomal rearrangements (Cuceu *et al.*, 2018a). Cytogenetic studies of Hodgkin lymphoma cell lines have additionally identified dicentric chromosomes, nucleoplasmic bridges and breakage–fusion–bridge cycles, indicating ongoing chromosomal instability in these cells (Cuceu *et al.*, 2018b). These mechanisms can progressively increase structural genome complexity and generate derivative chromosomes containing material from several different chromosomes.

An extreme form of genome restructuring in lymphoma is chromothripsis, discussed in Section 3.2. Chromothripsis has been identified across both solid and hematological malignancies, including Hodgkin lymphoma (Stephens *et al.*, 2011; Nagel *et al.*, 2013). Nagel *et al.* (2013) demonstrated chromothriptic rearrangements in Hodgkin lymphoma-derived cell lines, including HDLM-2 and L-1236, showing that these models contain highly localized and complex patterns of chromosome fragmentation and reassembly. In L-1236, amplified segments originating from the long arms of chromosomes 3 and 9 were found together on a derivative chromosome 6. These rearrangements also involved the ABL1 locus at 9q34, resulting in increased ABL1 expression and demonstrating that chromothripsis can directly alter genes relevant to tumor-cell behavior (Nagel *et al.*, 2013).

The genomic complexity of HDLM-2 and L-1236 has also been demonstrated independently of chromothripsis. Array-based comparative genomic hybridization identified numerous chromosomal gains and losses in both cell lines (Feys *et al.*, 2007), while cytogenetic analyses revealed complex structural rearrangements and clonal dicentric chromosomes (Cuceu *et al.*, 2018b). In HDLM-2, telomere dysfunction and chromosome fusion have been associated with breakage–fusion–bridge cycles and the formation of complex derivative chromosomes, whereas L-1236 also contains clonal dicentric chromosomes and extensive structural abnormalities (Cuceu *et al.*, 2018b).

These observations are particularly relevant to the present study because HDLM-2 and L-1236 provide two Hodgkin lymphoma models in which extensive chromosome restructuring and chromothripsis have already been documented. Such alterations can substantially modify chromosome architecture and therefore raise the question of whether centromeric regions are also affected during genome reorganization. The analysis of these cell lines therefore provides an opportunity to investigate whether highly rearranged lymphoma genomes are accompanied by changes in centromere organization, a question that remains considerably less explored than alterations affecting coding regions and oncogenes.

4.4 Centromere-associated alterations in lymphoma and rationale of the present study

Despite the extensive genomic and cytogenetic characterization of lymphoid malignancies, comparatively little is known about how centromere organization is altered in lymphoma. Evidence from other hematological malignancies nevertheless indicates that centromere reorganization can occur in highly rearranged cancer genomes. In acute myeloid leukemia, MYC-containing amplified material has been described in ring chromosomes stabilized by neocentromere formation (L'Abbate *et al.*, 2018). More recently, neocentromeres were identified on derivative chromosome 11 in four patients with acute leukemia and complex karyotypes, further demonstrating that newly acquired centromeric activity can contribute to the stabilization of structurally abnormal chromosomes (Mendlikova *et al.*, 2026).

Alterations in proteins associated with centromere and mitotic function have also been reported in lymphoma. Increased expression of the inner centromere protein (INCENP) was observed in high-grade B-cell non-Hodgkin lymphomas compared with low-grade cases and was associated with a higher proliferative index (Barbanis *et al.*, 2009). More recently, increased CENP-F expression was reported in relapsed or refractory diffuse large B-cell lymphoma compared with samples from patients who achieved complete remission (Yang *et al.*, 2022). These observations suggest that alterations in centromere- and mitosis-associated proteins may provide information about lymphoma biology and disease behavior. Such alterations could therefore be investigated for potential diagnostic or prognostic relevance, although the available studies have not established centromere-associated proteins as biomarkers for routine lymphoma diagnosis or risk assessment. Therefore, a possible relationship between centromere remodeling, genome instability and clinical behavior in lymphoma should currently be regarded as a testable hypothesis rather than an established mechanism.

These findings highlight an important gap in the understanding of lymphoma genome organization. In particular, it remains unclear whether extensive chromosome rearrangement is accompanied by changes in the relationship between the epigenetic centromere marker CENP-A, the α-satellite-binding protein CENP-B and centromeric satellite DNA. The present study therefore investigated centromere organization in the Hodgkin lymphoma models HDLM-2 and L-1236 and the biologically distinct Burkitt lymphoma model BJAB. The combined analysis of CENP-A, CENP-B and centromeric satellite DNA was used to investigate whether these components retain their canonical association or show evidence of reorganization in lymphoma cells.

AIMS OF THE WORK

Centromere identity in humans is primarily defined by CENP-A-containing chromatin and is usually associated with α-satellite DNA and CENP-B. However, centromere organization can be remodeled, and non-canonical centromeric domains, including neocentromeres formed outside canonical centromeric satellite DNA, may arise in the context of extensive chromosome rearrangements and genome instability. In cancer, this plasticity may be particularly relevant in highly rearranged genomes, including those affected by chromothripsis.

The present study focused on the lymphoma cell lines BJAB, HDLM-2 and L-1236, selected following transcriptomic analysis of centromere-associated genes. HDLM-2 and L-1236 were additionally of particular interest because chromothripsis and extensive chromosomal rearrangements have previously been reported in these cell lines, providing suitable models to investigate a possible relationship between genome restructuring, centromere remodeling and neocentromere formation.

The primary objectives of this work were:

1. To characterize the expression of centromere-associated genes across lymphoma cell lines and identify models for further analysis.
2. To investigate the protein abundance, subcellular distribution and chromosomal localization of CENP-A and CENP-B in BJAB, HDLM-2 and L-1236.
3. To examine the relationship between CENP-A, CENP-B and centromeric satellite DNA and to identify possible non-canonical patterns of centromere organization, including candidate neocentromeres characterized by CENP-A localization at sites lacking detectable canonical centromeric satellite DNA.
4. To investigate whether centromere remodeling, including altered CENP-A distribution and potential neocentromere formation, may be associated with the highly rearranged genomic backgrounds and previously reported chromothripsis of HDLM-2 and L-1236.

MATERIALS AND METHODS

1. RNA-SEQ ANALYSIS

Publicly available RNA-sequencing datasets were retrieved for lymphoma cell lines from the LL-100 panel (Quentmeier *et al.*, 2019) and healthy lymph-node samples. Lymphoma transcriptomic data were obtained from the PRJEB30312 and E-MTAB-7721 datasets. Three healthy lymph-node RNA-seq samples from the Human Protein Atlas dataset PRJEB4337 were included as non-malignant controls.

RNA-seq data were processed using the nf-core/rnaseq pipeline version 3.14.0 on the University of Pavia high-performance computing cluster. The telomere-to-telomere T2T-CHM13 human genome assembly was used as the reference genome. Downstream data processing and graphical representation were carried out in R on the University of Pavia high-performance computing cluster.

2. CELL CULTURE

The human lymphoma cell lines BJAB (Burkitt lymphoma; DSMZ ACC 757), HDLM-2 (Hodgkin lymphoma; DSMZ ACC 17) and L-1236 (Hodgkin lymphoma; DSMZ ACC 530) were obtained from the Leibniz Institute DSMZ–German Collection of Microorganisms and Cell Cultures.

All cell lines were maintained in RPMI 1640 medium completed with 15% heat-inactivated fetal bovine serum (FBS), 1% penicillin–streptomycin and 1% L-glutamine. Cells were cultured at 37 °C in a humidified atmosphere containing 5% CO_2.

3. WESTERN BLOT

Preparation of Total, Nuclear and Cytoplasmic Protein Extracts

Approximately 30 million cells were collected and resuspended in 900 µL of ice-cold lysis buffer containing 0.1% NP-40 in 1× PBS. The suspension was mixed by pipetting five times with a 1,000-µL pipette tip, and all fractionation steps were performed on ice. To prepare the total protein extract, 300 µL of the cell suspension was mixed with 100 µL of 4× Laemmli buffer. The remaining 600 µL was centrifuged at 6,000 rpm for 10 seconds at 4 °C. For the cytoplasmic extract, 300 µL of the resulting supernatant was collected and mixed with 100 µL

of 4× Laemmli buffer. The cytoplasmic extract was heated at 100 °C for 1 minute, aliquoted and stored at −20 °C.

The pellet was resuspended in 1 mL of ice-cold lysis buffer containing 0.1% NP-40 in 1× PBS. Following centrifugation at 6,000 rpm for 10 seconds at 4 °C, the supernatant was discarded and the pellet, with an approximate volume of 20 µL, was resuspended in 180 µL of 1× Laemmli buffer to obtain the nuclear extract. The total and nuclear extracts were sonicated using two 5-second pulses with a microtip at 10% amplitude. The extracts were then heated at 100 °C for 1 minute, aliquoted and stored at −20 °C.

SDS–PAGE and Protein Transfer

Proteins were separated by SDS–PAGE on a 12% polyacrylamide gel and blotted to 0.45-µm nitrocellulose membrane (GE Healthcare Life Sciences) according to standard methods.

The membrane was blocked by incubation in blocking solution consisting of 7.5% skimmed milk prepared in 1× PBS containing 0.05% Tween 20. Blocking was performed for 2.5 hours at 4 °C with gentle agitation. The blocking solution was then discarded, and the membrane was incubated with primary antibodies diluted in the same milk-based blocking solution. A mouse anti-tubulin antibody was used at a dilution of 1:5,000, while mouse monoclonal anti-CENP-A antibody (ADI-KAM-CC006-E, Enzo Life Sciences) was used at a dilution of 1:1,000. Primary-antibody incubation was performed for 1 hour at 4 °C with gentle agitation.

Following primary-antibody incubation, the membrane was washed three times for 10 minutes each with 1× PBS containing 0.05% Tween 20 (PBST). The membrane was subsequently incubated with horseradish peroxidase-conjugated anti-mouse secondary antibody diluted 1:5,000 in blocking solution for 1 hour with gentle agitation. The membrane was then washed three additional times with PBST.

Chemiluminescent detection was performed using the Clarity Western ECL Substrate kit (Bio-Rad) according to the manufacturer's instructions. Chemiluminescent signals were visualized using a ChemiDoc imaging system.

Subsequently, the membrane was washed in PBST at 4 °C overnight, on the rocker, for the immunodetection of CENP-B. The rabbit monoclonal anti-CENP-B antibody (ab259855, Abcam) was diluted 1:1,000 and incubated for 1 hour at 4 °C with gentle agitation. Secondary antibody incubation and detection was performed as previously described.

4. SLIDE PREPARATION

Approximately 20×10^4 cells were counted using a Bürker counting chamber and collected by centrifugation at $200 \times g$ for 5 minutes. After removal of the culture medium, the cell pellet was washed once with 1× phosphate-buffered saline (PBS) and centrifuged again at $200 \times g$ for 5 minutes. The PBS was carefully discarded, and the cells were resuspended in 1 mL of 56 mM KCl.

Cells were incubated for 13 minutes at 37 °C for hypotonic treatment and then transferred onto ice. The cell suspension was deposited onto microscope slides using a HERMLE Z 300 cytocentrifuge at 1,250 rpm for 10 minutes. Slides were fixed in ice-cold methanol for 3 minutes, counterstained with 0.2 µg/mL DAPI (4′,6-diamidino-2-phenylindole) for 3 minutes and mounted using McIlvaine mounting solution.

5. IMMUNOFLUORESCENCE

Coverslips were removed by immersing the slides in 2× SSC. The slides were then washed in fresh 2× SSC for 15 minutes, followed by incubation in 1× PBS containing 0.05% Tween 20 for 15 minutes.

The ADI-KAM-CC006-E mouse monoclonal anti-CENP-A antibody (Enzo Life Sciences) and the rabbit monoclonal anti-CENP-B antibody (clone EPR24047-64; ab259855, Abcam) were diluted 1:100 in 1× PBS containing 0.05% Tween 20. The antibody solution was applied to the slides, which were incubated for 1 hour at 37 °C in a dark chamber. Slides were then washed three times for 5 minutes each in 1× PBS containing 0.05% Tween 20.

For secondary detection, Texas Red-conjugated anti-mouse secondary antibody for CENP-A and Alexa Fluor 488-conjugated anti-rabbit secondary antibody for CENP-B were diluted 1:100 in the same buffer and added to the slides. After incubation for 1 hour at 37 °C in the dark, the slides were washed three times for 5 minutes each in 1× PBS containing 0.05% Tween 20.

Finally, chromosomes were counterstained with 0.2 µg/mL DAPI for 3 minutes and the slides were mounted using McIlvaine mounting solution.

6. FLUORESCENCE *IN SITU* HYBRIDIZATION (FISH)

Coverslips were removed by immersing the slides in 2× SSC. Slides were then washed in fresh 2× SSC for 15 minutes, fixed in cold methanol for 15 minutes and allowed to air-dry.

For hybridization, 20 µL of hybridization solution containing 70% formamide and 0.2 µL of the pan-centromeric PNA probe was applied to each slide. The slides were incubated at 73 °C for 3 minutes and 30 seconds and then hybridized for 2 hours at room temperature in a dark chamber.

Following hybridization, slides were washed twice in a solution containing 70% formamide, 10 mM Tris-HCl and 0.1% BSA, followed by three washes in TBS for 5 minutes each. Slides were dehydrated in 70%, 85% and 100% ethanol for 5 minutes at each concentration and air-dried for 15 minutes.

Finally, chromosomes were counterstained with 0.2 µg/mL DAPI for 3 minutes, and the slides were mounted using McIlvaine mounting solution.

7. VISUALIZATION AND IMAGE ANALYSIS

Grayscale fluorescence images were acquired using a Zeiss Axio Scope.A1 microscope equipped with a Photometrics CCD camera. Images were subsequently pseudo-colored and merged using IPLab software.

8. CHROMATIN IMMUNOPRECIPITATION

For each chromatin preparation, approximately 30 million cells were collected by centrifugation at 200 × g for 5 minutes at 4 °C.

To preserve protein–DNA interactions, the cells were cross-linked with 1% formaldehyde for 15 minutes at 26 °C under continuous agitation. The cross-linking reaction was stopped by adding glycine to a final concentration of 0.125 M, followed by incubation for 10 minutes at 26 °C. The cell suspension was then centrifuged at 800 × g for 5 minutes at 4 °C, and the resulting pellet was stored at −80 °C overnight.

The cell pellet was thawed on ice and washed twice with 1× phosphate-buffered saline supplemented with protease inhibitors. After each wash, the samples were centrifuged at 800 × g for 6 minutes at 4 °C. The final pellet was resuspended in ChIP lysis buffer containing

0.25% sodium dodecyl sulfate, 50 mM Tris-HCl at pH 8.0 and 10 mM EDTA at pH 8.0, supplemented with a protease inhibitor cocktail.

Chromatin was fragmented using a Branson Sonifier 250 to obtain DNA fragments ranging from approximately 200 to 800 base pairs. Fragment size was evaluated by agarose gel electrophoresis. Following sonication, the samples were centrifuged at 13,000 rpm for 10 minutes at 4 °C to remove insoluble material.

The recovered supernatant was diluted with ChIP dilution buffer containing 0.5% Nonidet P-40, 10 mM Tris-HCl at pH 7.5, 2.5 mM $MgCl_2$ and 150 mM NaCl, supplemented with protease inhibitors. To reduce non-specific interactions, the diluted chromatin was pre-cleared by incubation with Protein A/G Sepharose beads (nProtein A Sepharose™ 4 Fast Flow/Protein G Sepharose™ 4 Fast Flow, GE Healthcare) for 1 hour at 4 °C under continuous agitation. The samples were subsequently centrifuged at 4,000 rpm for 5 minutes at 4 °C.

An aliquot of the pre-cleared chromatin was retained as the input control. The remaining chromatin was incubated overnight at 4 °C with an anti-CENP-A antibody under continuous agitation. Immunocomplexes were then captured by adding Protein G Sepharose beads and incubating the samples for 3 hours at 4 °C under agitation.

The beads were collected by centrifugation at 1,200 × g for 2 minutes at 4 °C. After removal of the supernatant, the beads were washed five times with ChIP wash buffer containing 0.25% sodium dodecyl sulfate, 1% Triton X-100, 2 mM EDTA at pH 8.0, 150 mM NaCl and 20 mM Tris-HCl at pH 8.0. A final wash was performed using the same buffer containing 500 mM NaCl.

Following the washes, the immunocomplexes were eluted in ChIP elution buffer containing 1% sodium dodecyl sulfate, 100 mM $NaHCO_3$ and 40 µg/mL RNase A. Samples were incubated for 15 minutes at room temperature, followed by incubation for 1 hour at 37 °C. Cross-links were reversed by incubating the immunoprecipitated and input samples overnight at 65 °C.

DNA was purified using the Wizard SV Gel and PCR Clean-Up System according to the manufacturer's instructions and quantified using a Quantus Fluorometer. The purified immunoprecipitated and input DNA samples were subsequently used for paired-end high-throughput sequencing.

9. ChIP-seq BIOINFORMATIC ANALYSIS

Paired-end sequencing reads were aligned to the human T2T-CHM13 reference genome using Bowtie2 in paired-end mode. SAMtools was used to filter the aligned reads by excluding unmapped reads, non-primary alignments and supplementary alignments. CENP-A ChIP-seq coverage was normalized against the corresponding input dataset using the bamCompare function of the deepTools suite, applying RPKM normalization in subtractive mode.

Normalized coverage profiles were visualized using pyGenomeTracks.

CENP-A enrichment profiles were analyzed together with the CenSat v2.1 centromeric satellite annotation of the T2T-CHM13v2.0 reference genome, including active, inactive and divergent α-satellite higher-order repeat (HOR) arrays, monomeric α-satellite regions and other centromeric and pericentromeric satellite classes. Predicted CENP-B-box positions, identified by FIMO motif scanning of the T2T-CHM13v2.0 genome, were included as a separate annotation track.

RESULTS

1. RNA-SEQ ANALYSIS OF CENTROMERIC GENE EXPRESSION

To investigate the expression of centromere-associated genes in lymphoma, we focused on cell lines included in the LL-100 panel, for which publicly available RNA-sequencing data are available (Quentmeier *et al.*, 2019). Transcriptomic data were obtained from two independent datasets. RNA-seq data from healthy lymph-node samples were included as a non-malignant reference. Expression profiles of centromere-associated genes were examined across the LL-100 lymphoma cell-line panel and visualized using a heatmap, which showed variability among the different lymphoma subtypes.

The heatmap showed marked variability in the expression of centromere-associated genes across the lymphoma cell-line panel (Figure 1). CENP-A transcript levels were generally higher in the lymphoma cell lines than in the healthy lymph-node control. Consistently higher transcript levels were also observed for CENP-H, CENP-W, CENP-M, CENP-N and CENP-X across many of the analyzed cell lines. In contrast, CENP-C generally showed lower transcript levels than the healthy control.

Some genes displayed particularly strong cell-line-specific variability. CENP-B showed a heterogeneous expression pattern and was undetectable in BJAB and CRO-AP2, despite the detectable expression of all other centromere-associated genes in these cell lines. CENP-V also showed considerable variability, ranging from undetectable or very low expression in some cell lines to high expression in others.

Overall, the heatmap did not reveal a common expression profile shared by all cell lines belonging to the same lymphoma subtype. Variability was observed both between different lymphoma groups and among cell lines within the same group.

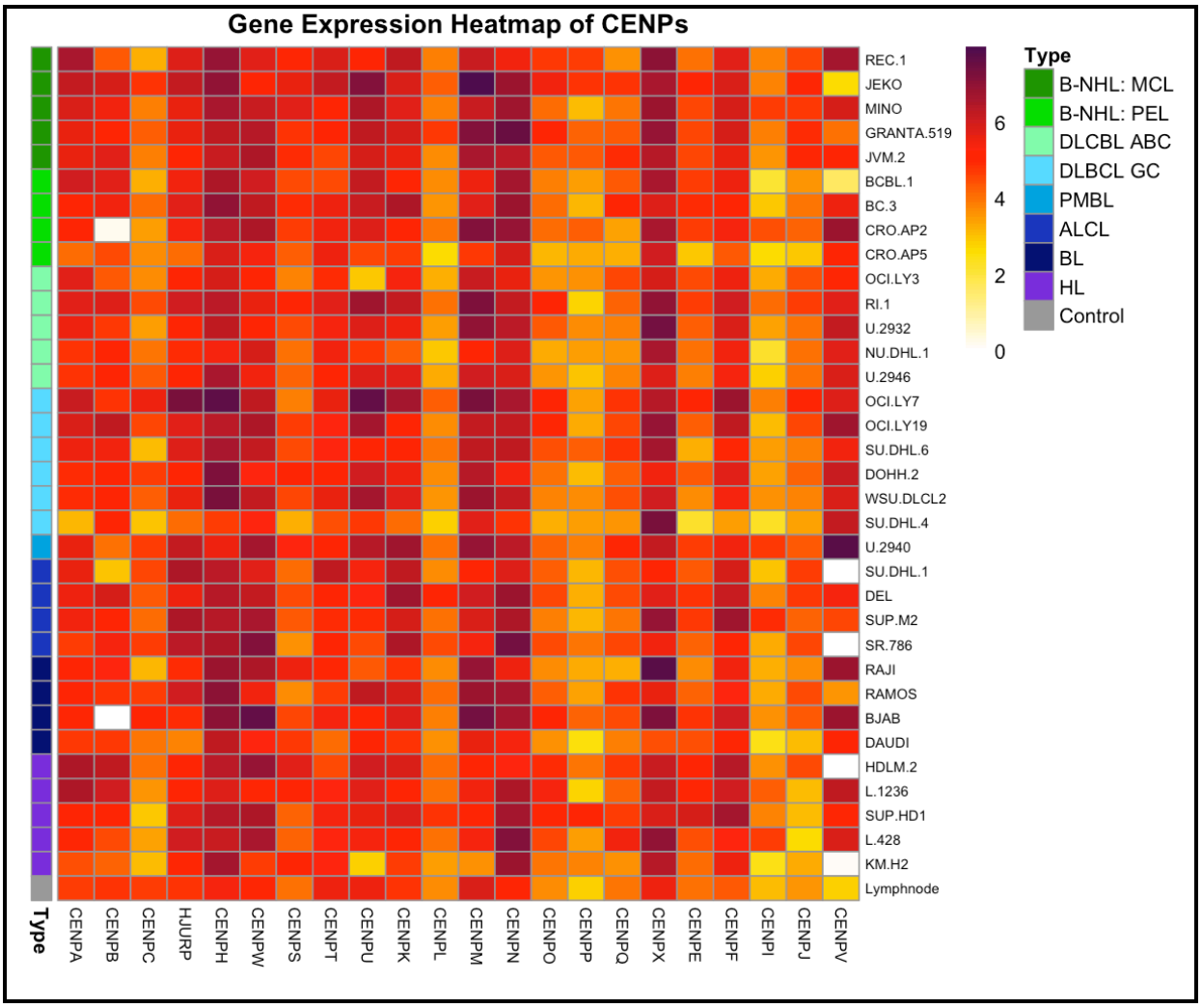

Figure 1. Heatmap of CENP gene expression across lymphoma cell lines. This heatmap represents the transcript levels (log2 TPM) of centromere-associated genes across various lymphoma cell lines and a control sample. Expression levels are displayed using a color scale, where yellow indicates lower expression and red to dark red represents higher expression levels. White cells indicate undetectable expression values. The average TPM values of the three healthy control lymph nodes were used as the control. Cell lines are grouped according to lymphoma subtype.

2. SELECTION OF CELL LINES

Based on the RNA-sequencing analysis and previous studies, we selected three lymphoma cell lines for subsequent experimental analyses.

HDLM-2 and L-1236 were of particular interest because both Hodgkin lymphoma cell lines have previously been reported to exhibit highly rearranged karyotypes associated with chromothripsis (Nagel *et al.*, 2013). In addition, RNA-seq analysis demonstrated increased expression of both CENP-A and CENP-B in these cell lines compared with healthy lymph node samples.

In contrast, we selected BJAB because RNA-seq analysis revealed no detectable CENP-B transcripts, whereas CENP-A expression remained above that observed in the control samples, providing a distinct centromeric protein profile for further investigation.

To examine the expression of the two main proteins investigated in this study, CENP-A and CENP-B transcript levels were compared in BJAB, HDLM-2 and L-1236 and in healthy lymph-node samples.

The focused comparison of BJAB, HDLM-2 and L-1236 was consistent with the broader expression patterns observed in the heatmap (Figure 2). CENP-A transcript levels were higher in all three lymphoma cell lines than in the healthy lymph-node control, with more marked increases observed in HDLM-2 and L-1236 and a more moderate increase in BJAB.

CENP-B showed a different expression profile, with higher transcript levels in HDLM-2 and L-1236 but no detectable transcripts in BJAB. These differences supported the selection of HDLM-2 and L-1236 as models expressing both CENP-A and CENP-B, whereas BJAB provided a contrasting model in which CENP-A transcripts were maintained despite the absence of detectable CENP-B transcripts.

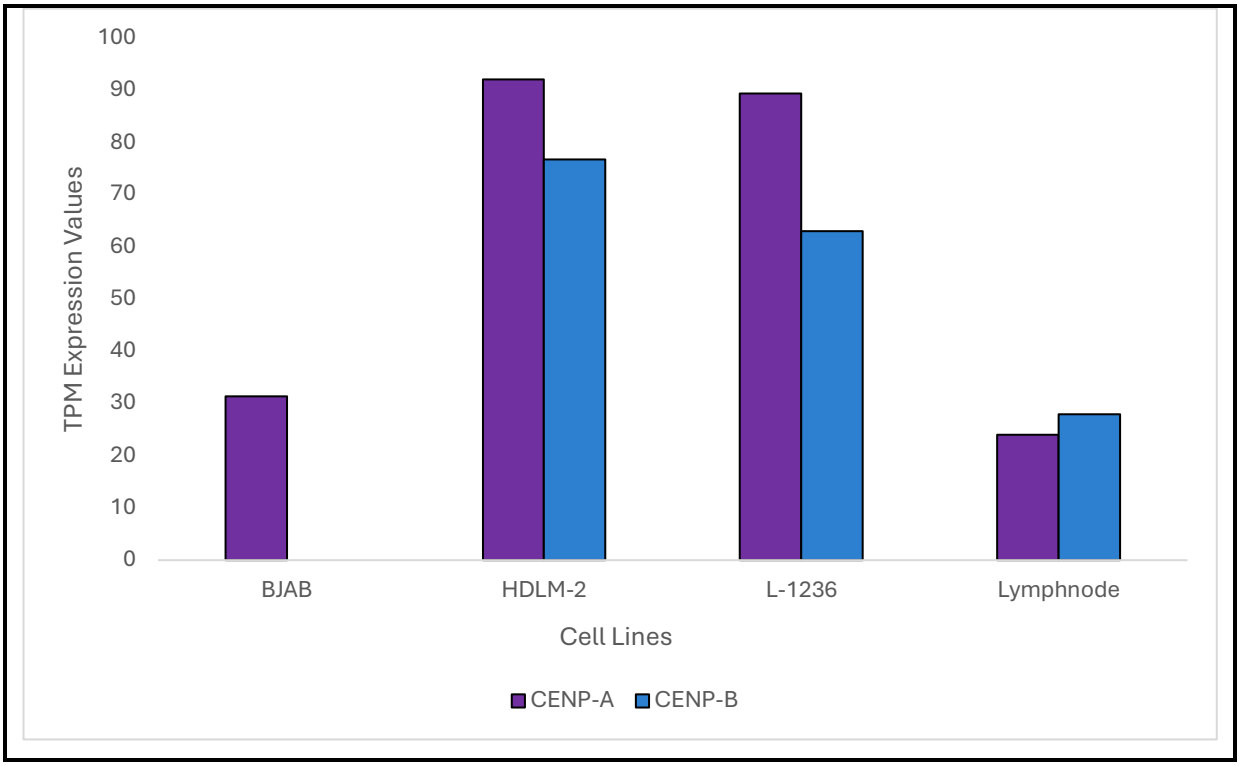

Figure 2. Transcript levels of CENP-A and CENP-B in selected lymphoma cell lines and a healthy control. Comparison of CENP-A (purple bars) and CENP-B (blue bars) transcript levels, expressed as transcripts per million (TPM), in BJAB, HDLM-2 and L-1236 cells and the healthy lymph-node control.

We obtained the selected lymphoma cell lines from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (DSMZ) repository. In addition to experimental analyses, publicly available RNA-sequencing and whole-exome sequencing datasets for these models were available and we used them for the molecular characterization of centromeric protein expression and genomic alterations.

HDLM-2 was established in 1982 from the pleural effusion of a 74-year-old male patient diagnosed with stage IV nodular sclerosing Hodgkin lymphoma. The cell line consists of round to polygonal suspension cells displaying marked heterogeneity in cell size (15–50 μm) and a high proportion of bi- and polynucleated cells, reflecting the characteristic morphology of Hodgkin lymphoma cells. HDLM-2 represents a highly rearranged, hypodiploid human cell line with a complex karyotype. Chromothripsis involving chromosome 17 has been described (Nagel *et al.*, 2013).

L-1236 was established in 1994 from the peripheral blood of a 34-year-old male patient diagnosed with stage IV mixed cellularity Hodgkin lymphoma at terminal disease stage following multiple relapses. The cell line displays variable morphology, consisting of semi-adherent spheroid to fusiform cells growing in clumps as mono- and multilayers, with occasional giant cells. L-1236 exhibits a complex hypotriploid karyotype with a proportion of polyploid cells (17%). Chromothripsis involving chromosomes 3q and 9 has been described (Nagel *et al.*, 2013).

BJAB was established in 1973 from an inguinal tumor of a 5-year-old female patient diagnosed with Epstein–Barr virus (EBV)-negative Burkitt lymphoma at the preterminal stage of disease. The cell line consists of round to polymorphic cells growing in suspension, either as single cells or in clusters. BJAB displays a human hypodiploid karyotype and represents an EBV-negative Burkitt lymphoma model.

3. TOTAL, NUCLEAR AND CYTOPLASMIC DISTRIBUTION OF CENP-A AND CENP-B IN THREE LYMPHOMA CELL LINES

To validate the RNA-sequencing results at the protein level and evaluate the subcellular distribution of centromeric proteins, we performed Western blot analyses on protein extracts obtained from the HDLM-2, L-1236, and BJAB lymphoma cell lines. We analyzed total protein extracts, together with separated nuclear and cytoplasmic fractions, to assess CENP-A and CENP-B protein levels and their cellular distribution. We used antibodies against CENP-A and CENP-B to assess centromeric protein levels. To assess protein loading and the separation of the nuclear and cytoplasmic fractions, we used an anti-tubulin antibody.

For qualitative comparison, previously acquired Western blot images of fractionated HeLa cells generated in our laboratory are shown in Figure 3A. In HeLa cells, CENP-B was detected in the total and nuclear fractions, whereas tubulin was detected in the total and cytoplasmic fractions. These samples were not processed in parallel with the lymphoma cell lines and were therefore included only as a qualitative reference and not as an experimental control or for direct quantitative comparison.

In the present work, Western blot analysis revealed distinct patterns of CENP-A and CENP-B abundance and subcellular distribution among the three lymphoma cell lines (Figure 3B). As expected, in all cell lines, CENP-A was detected in the total and nuclear protein extracts, whereas no detectable signals were observed in the cytoplasmic fractions. Tubulin was detected in the total and cytoplasmic fractions but was not detected in the nuclear fraction, consistent with a good separation of nuclear and cytoplasmic extracts. In BJAB, the CENP-A bands appeared weaker than those detected in HDLM-2 and L-1236, suggesting lower CENP-A protein levels. Consistent with the RNA-seq results, no detectable CENP-B signal was observed in any of the analyzed BJAB protein fractions.

In HDLM-2 cells, CENP-A was detected in the total and nuclear extracts, with more intense bands than those observed in the corresponding BJAB fractions (Figure 3B). CENP-B was detected in the total and nuclear extracts, where two closely migrating bands were visible. A weak CENP-B signal was also detected in the cytoplasmic fraction, suggesting that a minor proportion of the detectable protein was present outside the nuclear compartment. The possibility that this weak band is due to contamination from the nuclear extract is unlikely, as no such contamination is detectable with the anti-CENP-A antibody.

L-1236 cells showed a similar general distribution (Figure 3B). As observed in HDLM-2, the CENP-A bands in L-1236 were more intense than those detected in BJAB, consistent with higher CENP-A protein levels in the two Hodgkin lymphoma cell lines. CENP-B was detected in the total and nuclear extracts, where, as in HDLM-2, two closely migrating bands were observed. No CENP-B signal was detected in the L-1236 cytoplasmic fraction.

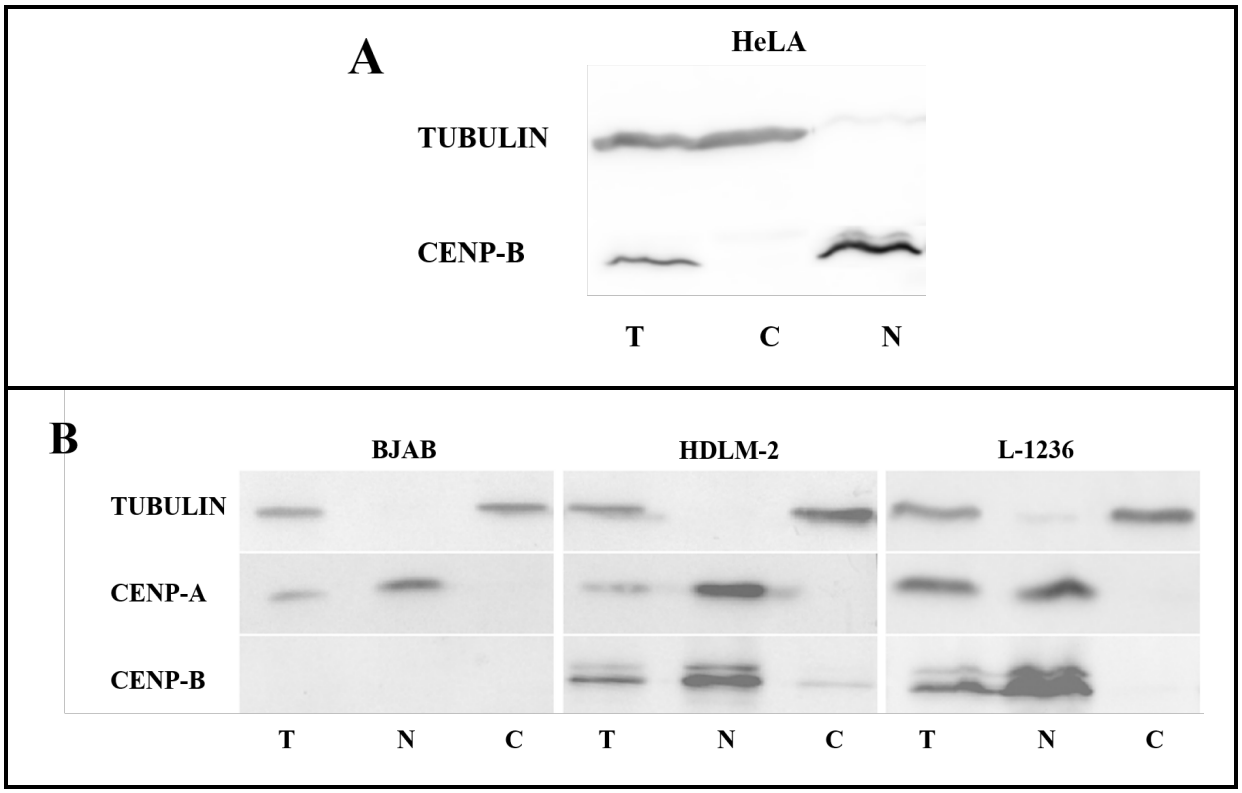

Figure 3. Western blot analysis of the subcellular distribution of CENP-A and CENP-B. (A) Previously acquired Western blot analysis of HeLa total (T), cytoplasmic (C) and nuclear (N) protein fractions showing the distribution of CENP-B and tubulin. (B) Western blot analysis of total (T), nuclear (N) and cytoplasmic (C) protein extracts from BJAB, HDLM-2 and L-1236 cells using antibodies against CENP-A, CENP-B and tubulin.

4. CYTOGENETIC ANALYSES

4.1 IMMUNOFLUORESCENCE ANALYSIS OF CENP-A AND CENP-B LOCALIZATION IN BJAB, HDLM-2 AND L-1236 CELLS

To investigate the localization of CENP-A and CENP-B at the chromosomal level and analyze the relationship between the two proteins in BJAB, HDLM-2 and L-1236 cells, we performed double-immunofluorescence experiments on metaphase chromosome preparations using anti-CENP-A and anti-CENP-B antibodies.

The HeLa metaphase image showed the characteristic discrete localization of CENP-B at centromeric regions (Figure 4). This previously acquired image was included as a qualitative reference for the typical chromosomal distribution of CENP-B. Because it was not obtained in parallel with the lymphoma-cell experiments, it was not used as an experimental control or for direct comparison of signal intensity. The CENP-B localization patterns observed in BJAB, HDLM-2 and L-1236 in the present work are described in the following sections.

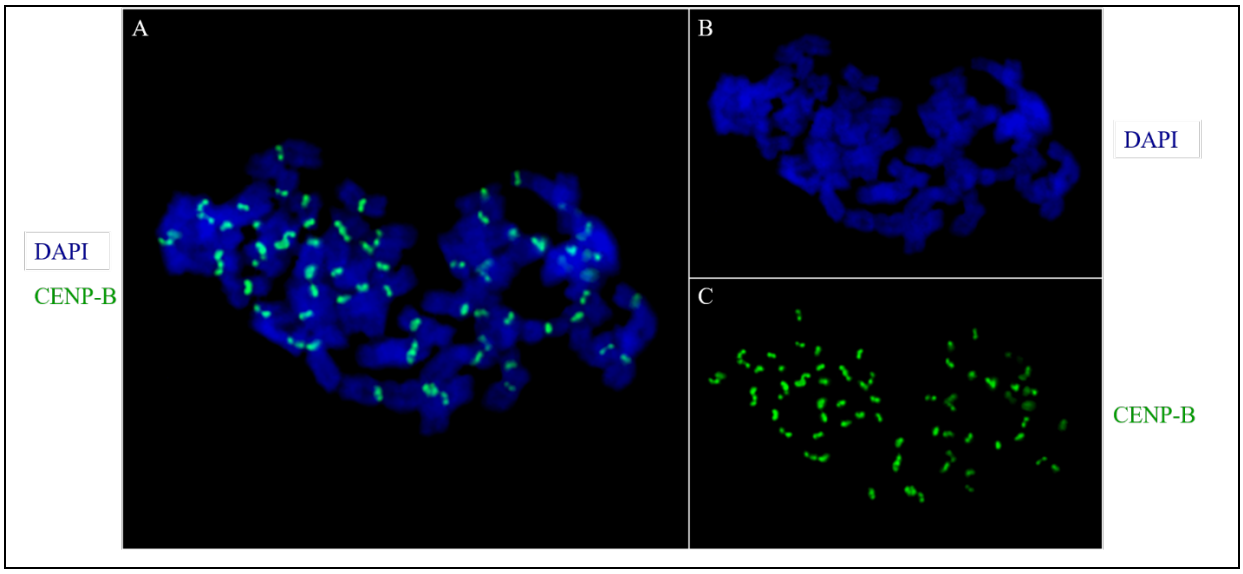

Figure 4. Immunolocalization of CENP-B in HeLa cells. Representative HeLa metaphase chromosome spread analyzed by immunofluorescence. (A) Merged image showing chromosomal DNA stained with DAPI (blue) and CENP-B detected by immunofluorescence (green). (B) DAPI staining showing metaphase chromosomes. (C) CENP-B channel showing discrete signals localized at centromeric regions.

In BJAB metaphase spreads, CENP-A was detected, as expected, at all centromeric regions, whereas no CENP-B signal was observed in the corresponding channel (Figure 5). Accordingly, the merged image showed centromeric CENP-A signals without visible overlap with CENP-B. This pattern indicates that, in BJAB cells, CENP-A remained localized at canonical centromeric sites despite the absence of detectable CENP-B.

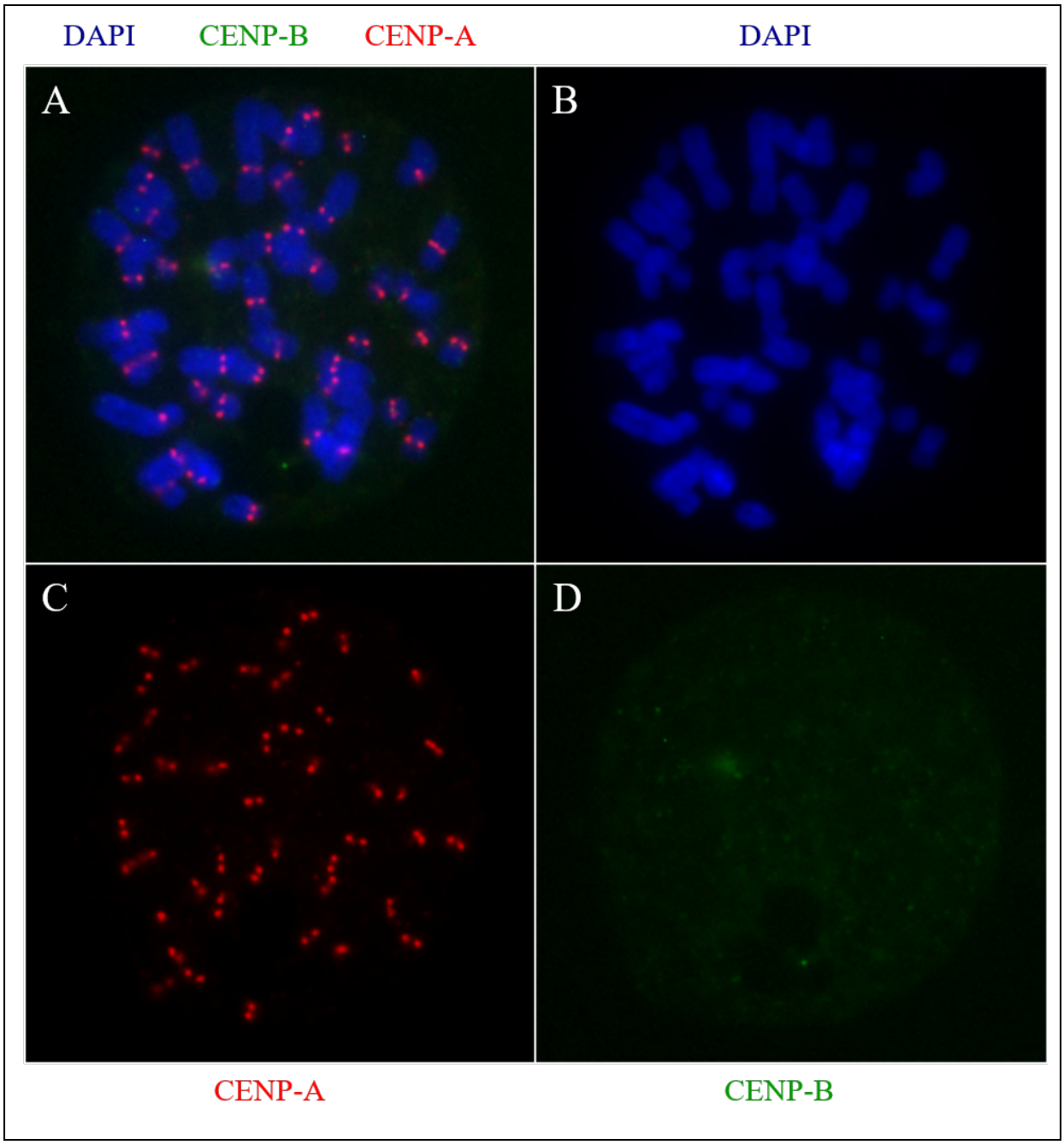

Figure 5. Immunofluorescence analysis of CENP-A and CENP-B localization in BJAB cells. Representative metaphase chromosome spread from BJAB cells. (A) Merged image showing chromosomal DNA stained with DAPI (blue), CENP-A detected by immunofluorescence (red), and CENP-B detected by immunofluorescence (green). (B) DAPI staining showing metaphase chromosomes. (C) CENP-A channel, showing discrete signals localized at centromeric regions. (D) CENP-B channel, showing no detectable specific CENP-B signal.

In HDLM-2 metaphase spreads, CENP-A and CENP-B signals colocalized at all expected centromeric regions (Figure 6). In addition to these overlapping centromeric signals, several CENP-B-positive chromosomal sites lacked a corresponding detectable CENP-A signal and are indicated by white arrows. These additional sites were observed on multiple

chromosomes. The merged image therefore showed both canonical CENP-A/CENP-B colocalization and CENP-B-positive signals without visible overlap with CENP-A.

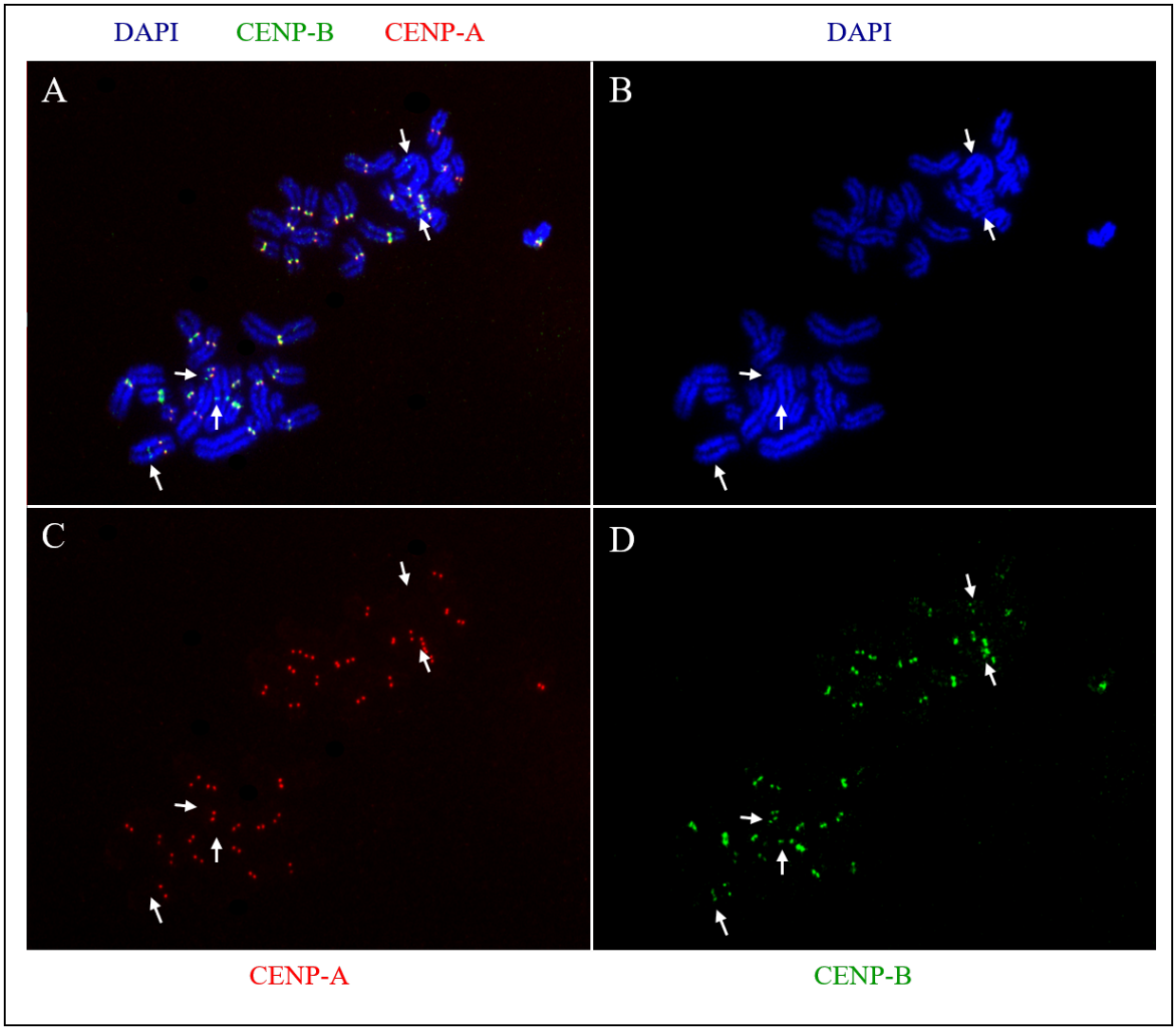

Figure 6. Immunofluorescence analysis of CENP-A and CENP-B localization in HDLM-2 cells. Representative HDLM-2 metaphase chromosome spread analyzed by double immunofluorescence. (A) Merged image showing chromosomal DNA stained with DAPI (blue), CENP-A detected by immunofluorescence (red), and CENP-B detected by immunofluorescence (green). (B) DAPI staining showing metaphase chromosomes. (C) CENP-A channel showing discrete signals at centromeric regions. (D) CENP-B channel showing centromeric signals together with additional CENP-B-positive chromosomal sites. White arrows indicate CENP-B-positive sites lacking a corresponding detectable CENP-A signal, with the same sites indicated across panels A–D.

In L-1236 metaphase spreads, CENP-A and CENP-B signals colocalized at all expected centromeric regions (Figure 7). In addition to these overlapping centromeric signals, some CENP-B signals were localized at extra-centromeric sites. No CENP-A signals were detected at these sites (white arrows). Thus, L-1236 cells, similarly to HDLM-2 cells, displayed

canonical CENP-A/CENP-B colocalization at all centromeres, together with additional CENP-B signals at interstitial non-centromeric sites.

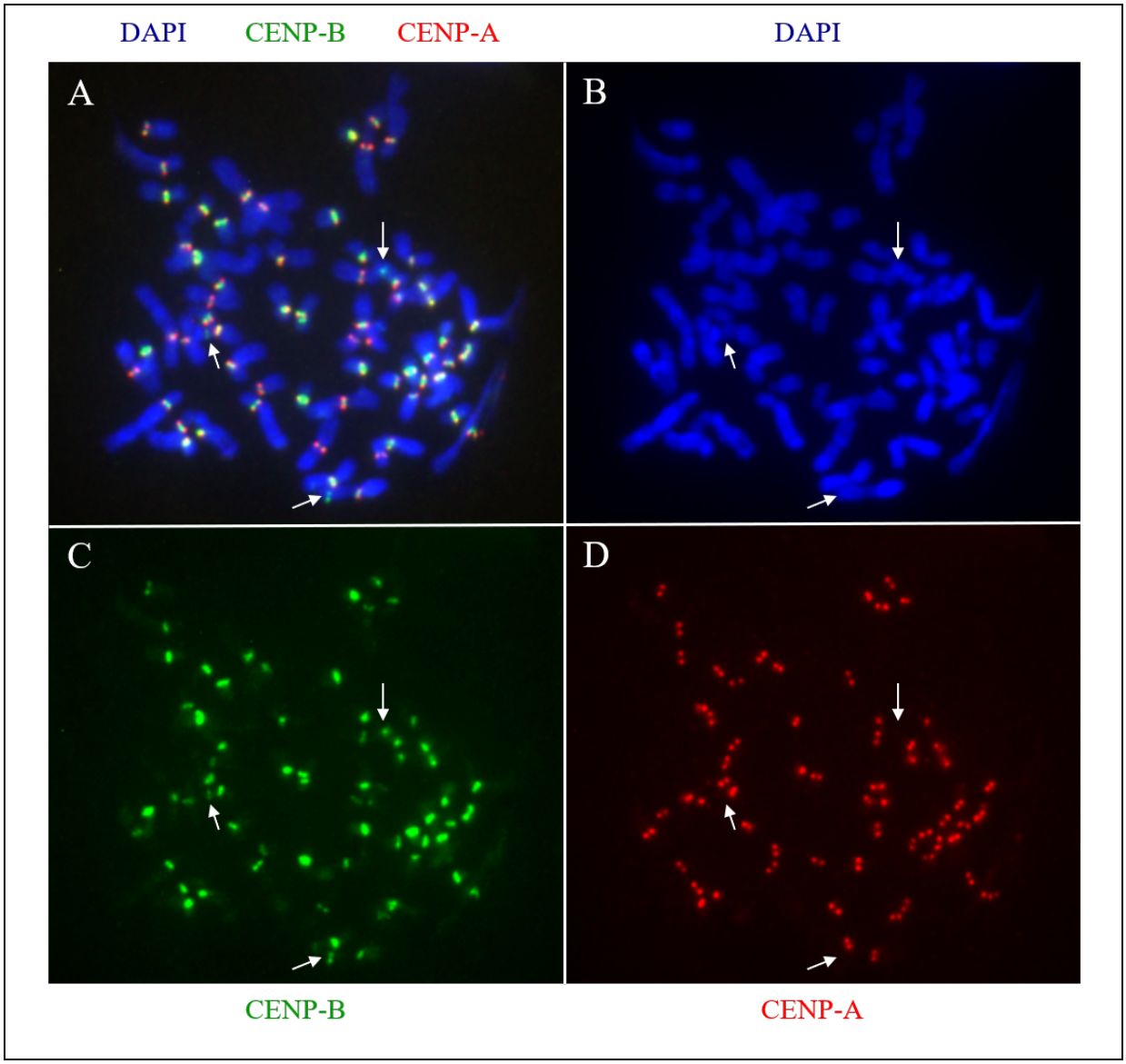

Figure 7. Immunofluorescence analysis of CENP-A and CENP-B localization in L-1236 cells. Representative L-1236 metaphase chromosome spread analyzed by double immunofluorescence. (A) Merged image showing chromosomal DNA stained with DAPI (blue), CENP-A detected by immunofluorescence (red), and CENP-B detected by immunofluorescence (green). (B) DAPI staining showing metaphase chromosomes. (C) CENP-A channel showing discrete signals at centromeric regions. (D) CENP-B channel showing centromeric signals together with additional CENP-B-positive chromosomal sites. White arrows indicate CENP-B-positive sites lacking a corresponding detectable CENP-A signal, with the same sites indicated across panels A–D.

Overall, the three lymphoma cell lines displayed distinct CENP-A/CENP-B localization patterns. BJAB showed centromeric CENP-A signals in the absence of detectable CENP-B. In contrast, HDLM-2 and L-1236 showed CENP-A/CENP-B colocalization at most centromeric

regions, together with additional CENP-B-positive chromosomal sites lacking a corresponding detectable CENP-A signal. These findings demonstrate cell-line-specific differences in the chromosomal distribution of CENP-A and CENP-B among the analyzed lymphoma models.

4.2 IMMUNO-FISH ANALYSIS OF CENP-A AND CENTROMERIC SATELLITE DNA IN L-1236 CELLS

To investigate the distribution of CENP-A relative to centromeric satellite DNA, and in particular to assess the presence of candidate satellite-free centromeric sites, CENP-A immunolocalization was combined with FISH on L-1236 metaphase chromosomes. CENP-A was detected by indirect immunofluorescence, whereas centromeric satellite DNA was detected using a pan-centromeric PNA probe. At the time of this analysis, the immunofluorescence–FISH experiments in BJAB and HDLM-2 were still in progress; therefore, the results presented here refer to L-1236 cells only.

Combined immunofluorescence–FISH analysis of L-1236 metaphase chromosomes showed that most CENP-A signals colocalized with detectable centromeric satellite-DNA signals (Figure 8), consistent with the usual association between CENP-A-containing chromatin and satellite-rich centromeric regions. However, two clearly distinguishable CENP-A-positive chromosomal sites lacked a corresponding detectable satellite-DNA signal and are indicated by white arrows. This pattern suggests that these sites may correspond to satellite-free centromeres. In addition, two extra-centromeric FISH signals (orange arrows) suggest the presence of extra-centromeric satellite DNA arrays.

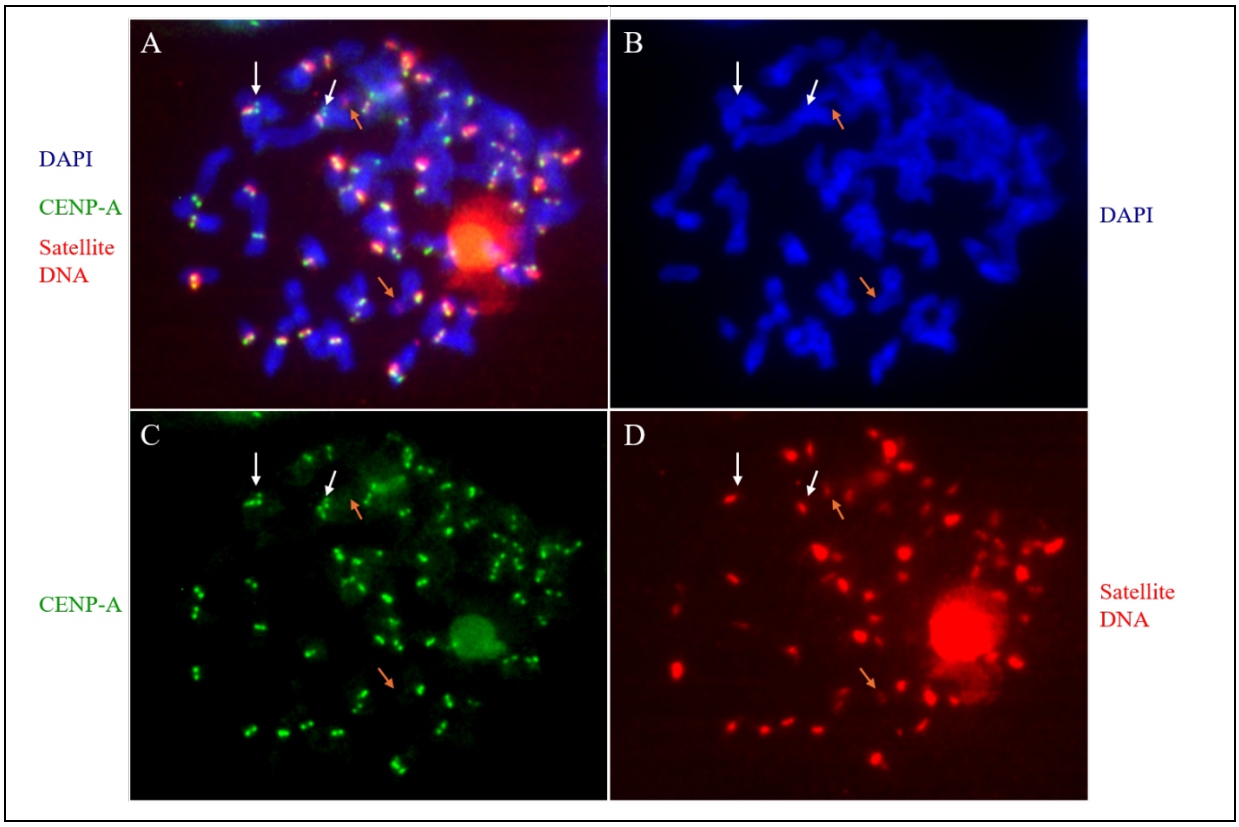

Figure 8. Combined immunofluorescence–FISH analysis of CENP-A and centromeric satellite DNA in L-1236 cells. Representative metaphase chromosome spread from L-1236 cells. (A) Merged image showing chromosomal DNA stained with DAPI (blue), CENP-A detected by immunofluorescence (green), and centromeric satellite DNA detected using a pan-centromeric PNA probe, which labels satellite DNA (red). (B) DAPI staining of metaphase chromosomes. (C) CENP-A channel. (D) Centromeric satellite-DNA channel. CENP-A generally colocalized with detectable centromeric satellite DNA. However, white arrows indicate CENP-A-positive chromosomal sites lacking a corresponding detectable centromeric satellite-DNA signal, whereas orange arrows indicate detectable centromeric satellite-DNA signals lacking a corresponding CENP-A signal. The same sites are indicated across panels A–D.

5. MOLECULAR MAPPING OF CENP-A BINDING DOMAINS (ChIP-seq)

To determine the genomic distribution of CENP-A in selected lymphoma cell lines, we performed CENP-A chromatin immunoprecipitation followed by sequencing (ChIP-seq) in HDLM-2, L-1236 and BJAB cells. A previously published CENP-A ChIP-seq dataset from the human fibroblast cell line PD-NC4 (Mahlke *et al.*, 2023; GSE221541) was retrieved and reprocessed using the same bioinformatic framework applied to the lymphoma samples. All datasets were aligned to the telomere-to-telomere T2T-CHM13v2.0 reference genome and normalized for subsequent comparison. The use of the T2T reference genome allowed CENP-A enrichment to be examined within the repetitive sequences that constitute human centromeric regions.

To interpret CENP-A enrichment in relation to the underlying centromeric sequence organization, the ChIP-seq profiles were examined together with the CenSat v2.1 annotation of the T2T-CHM13v2.0 reference genome. This annotation distinguishes different classes of centromeric and pericentromeric satellite DNA, including active, inactive and divergent α-satellite higher-order repeat (HOR) arrays, monomeric α-satellite DNA, and other human satellite families such as HSat1A, HSat1B, HSat2 and HSat3. Predicted CENP-B-box positions were displayed as a separate annotation track and were derived from a FIMO motif scan of the hs1/T2T-CHM13v2.0 genome. The combination of the CENP-A enrichment profiles with the satellite-DNA and CENP-B-box annotations allowed us to evaluate whether CENP-A domains were confined to reference-annotated active HOR arrays or extended into other centromeric or pericentromeric sequence classes.

In the fibroblast reference, CENP-A enrichment was detected within reference-annotated active α-satellite HOR arrays containing CENP-B boxes, corresponding to the expected localization of functional centromeres.

In Figures 9–16, the results of these experiments are shown for all chromosomes. Chromosome numbers are reported below each panel.

In BJAB, the maximum normalized read count displayed on the y-axis was lower than that of fibroblasts for all analyzed chromosomes. Despite this general difference, CENP-A enrichment remained associated with reference-annotated active HOR arrays on most chromosomes. Broadly preserved domain positions were observed on chromosomes 2, 3, 4, 7, 9, 12, 15, 16, 17, 18, 19 and 22 (Figures 9–16).

Several chromosomes nevertheless showed marked changes in domain organization. On chromosomes 6, 8, 10, and 11, the broad enrichment observed in fibroblasts was narrowed, fragmented or redistributed into separated subregions (Figures 11 and 12). The clearest reductions in domain extent were observed on chromosomes 8 and 11. On chromosome 8, the approximately 2-Mb fibroblast domain was replaced in BJAB by two narrow enrichment regions, measuring approximately 200 kb on the left and 350 kb on the right (Figure 11). On chromosome 11, BJAB enrichment was largely restricted to an approximately 400-kb region near the left boundary of an active HOR that extended for more than 3 Mb in fibroblasts (Figure 12). It is worth noting that, in these two chromosomes, the CENP-A-binding domain extended mainly into α-satellite regions lacking CENP-B boxes.

Changes beyond the principal active HOR were also observed. On chromosomes 5, 6 and 10, the BJAB domain extended towards the adjacent reference-annotated inactive HOR (Figures 10–12). On chromosome 13, two localized enrichment peaks of approximately 116 kb and 350 kb were detected (Figure 13). Interestingly, the separate peak was localized in a region where single-copy and pericentromeric satellite sequences are intermingled. A similar pattern was observed on chromosome 14. In this case, the principal localized peak measured approximately 400 kb, while the additional isolated peak measured approximately 100 kb (Figure 13) and overlapped a region where single-copy and pericentromeric satellite sequences are intermingled. An additional peak of approximately 115 kb was also observed on chromosome 21 (Figure 16). On chromosome X, CENP-A enrichment remained within the active HOR but shifted from the left portion occupied in fibroblasts towards an approximately 1.9-Mb region in the central and right portions of the array (Figure 16).

In L-1236, the maximum normalized read counts did not show a uniform difference relative to fibroblasts. The maximum normalized read counts varied across chromosomes, with values in L-1236 exceeding those in fibroblasts on some chromosomes and remaining lower on others (Figures 9-16). Despite this variability, CENP-A enrichment remained predominantly associated with reference-annotated active HOR arrays. Well preserved domain positions were observed on chromosomes 1–5, 7, 9–11, 13–17 and 19–22.

On chromosome 18, L-1236 enrichment was concentrated in two regions within the right portion of the active HOR (Figure 15).

On chromosome X, the CENP-A binding domain in L-1236 was similar to that of BJAB (Figure 16).

In HDLM-2, the maximum normalized read count was higher than that observed in fibroblasts on most chromosomes (Figures 9–16). Although CENP-A enrichment generally overlapped reference-annotated active HOR arrays, the HDLM-2 profiles frequently differed from those of fibroblasts in their continuity, internal distribution and domain width. Several chromosomes showed enrichment concentrated within only a portion of the active HOR, while chromosome 13 also displayed additional enrichment associated with HSat1A sequences.

On chromosomes 18 and 19, enrichments consisted of a broader low-level region together with a prominent narrow peak near the right boundary of the active HOR (Figure 15).

The most distinctive HDLM-2 profile was observed on chromosome 13. The principal CENP-A-enriched domain extended for approximately 3.7 Mb and crossed the boundary between HSat1A and the active HOR, with approximately 1.6 Mb of the domain located within HSat1A. A second, spatially separated region of irregular CENP-A enrichment spanning approximately 5 Mb was detected entirely within the left HSat1A array (Figure 13).

In conclusion, in the three lymphoma cell lines several centromeres showed distinct patterns of CENP-A organization relative to fibroblasts. BJAB displayed the most extensive variation. L-1236 largely retained enrichment within active HOR arrays but showed localized internal redistribution. HDLM-2 showed within-array redistribution and domain narrowing together with additional HSat1A-associated enrichment. These profiles represent mapped enrichment relative to the T2T-CHM13 reference annotations and do not resolve the exact centromeric sequence organization of the individual cell lines.

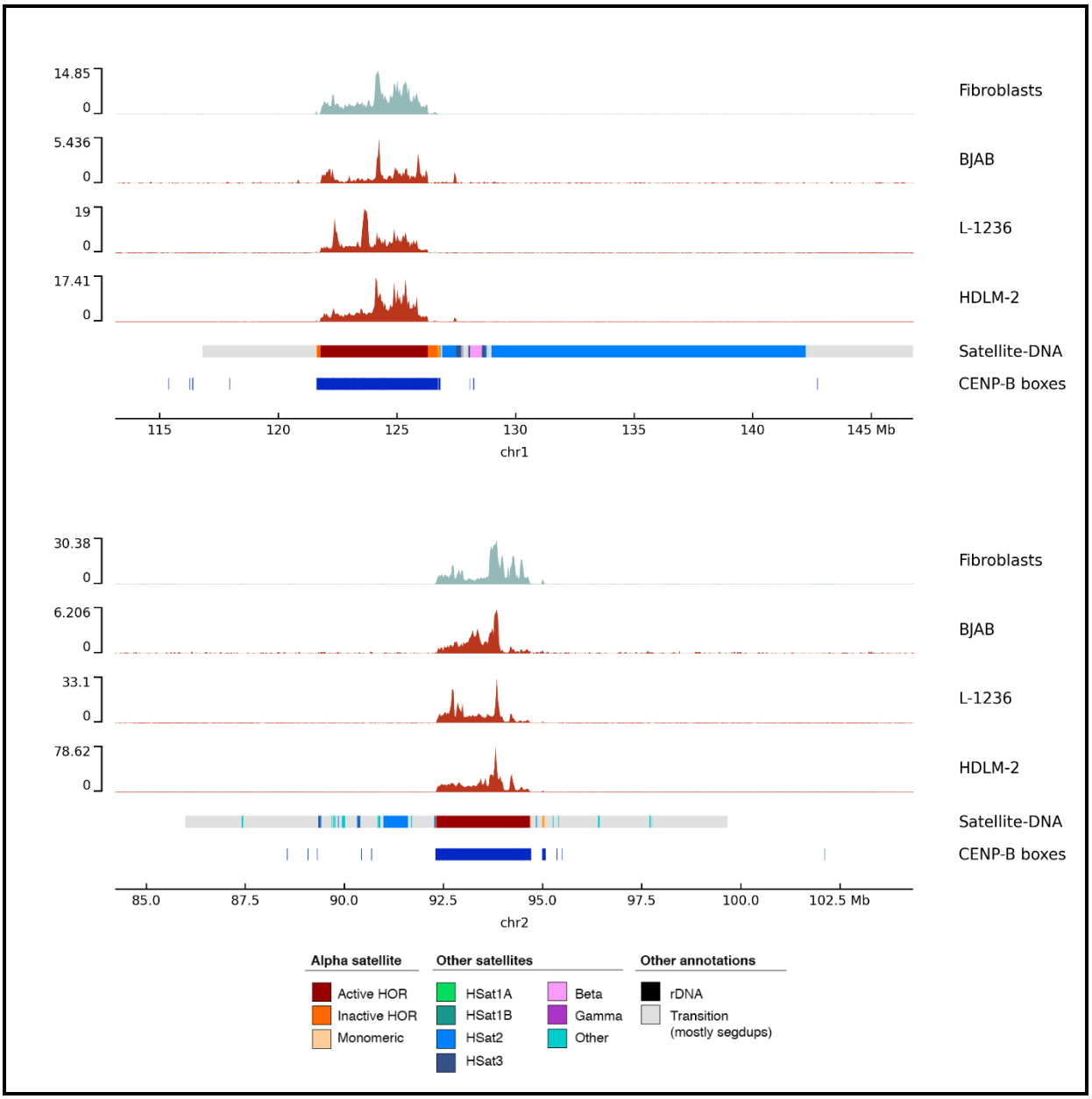

Figure 9. CENP-A ChIP-seq profiles of chromosomes 1 and 2. CENP-A enrichment profiles are shown for fibroblasts, BJAB, L-1236 and HDLM-2 cells across the centromeric regions. Satellite-DNA annotations are displayed below each set of profiles, together with the positions of predicted CENP-B boxes. The color code identifies the different satellite-DNA classes. The y-axis reports normalized read counts and the x-axis indicates genomic coordinates in megabases. The y-axis scale is shown separately for each sample.

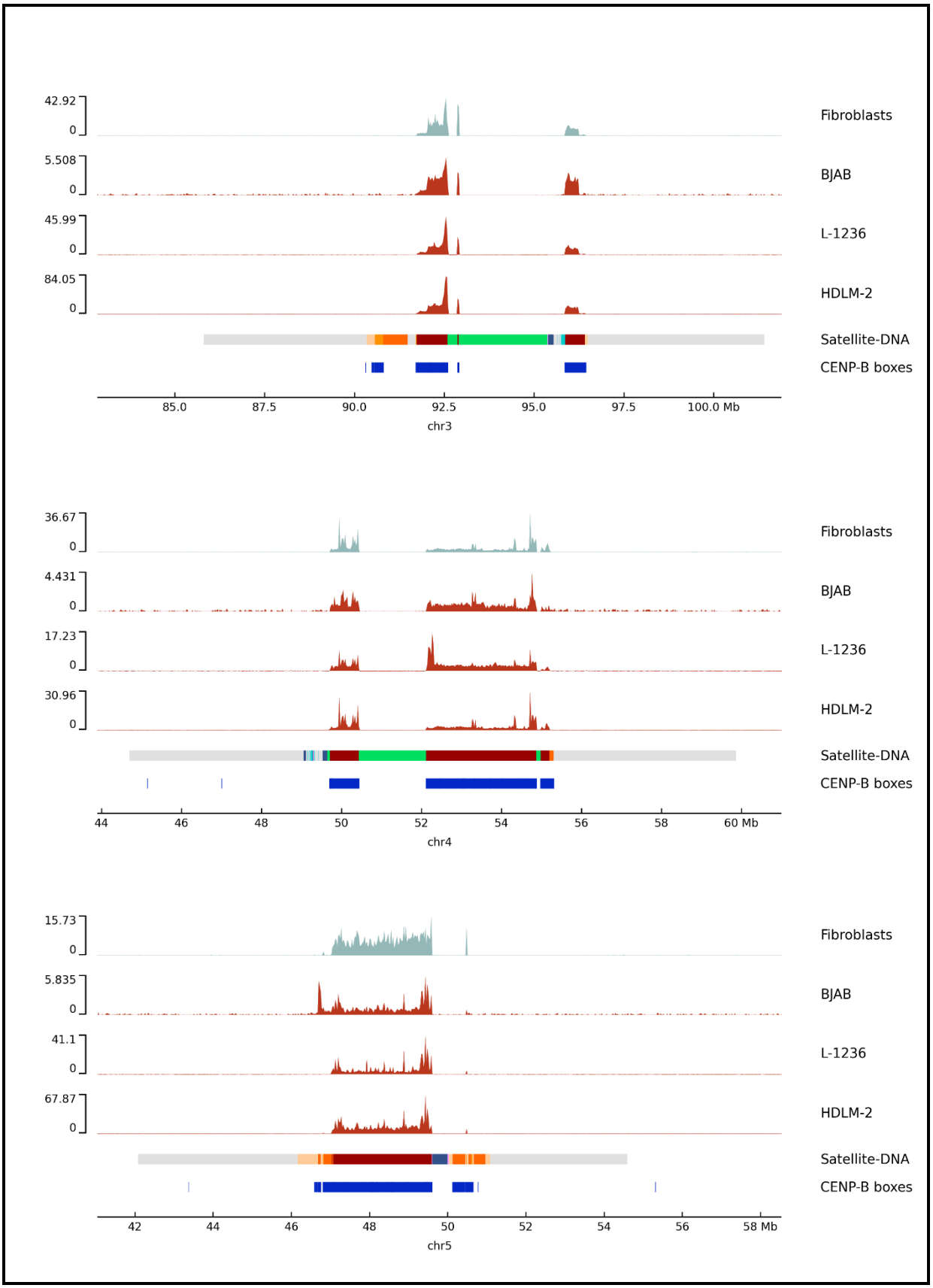

Figure 10. CENP-A ChIP-seq profiles of chromosomes 3, 4 and 5. CENP-A enrichment profiles for fibroblasts, BJAB, L-1236 and HDLM-2 cells are shown together with satellite-DNA annotations and predicted CENP-B-box positions. The figure is organized as described in Figure 9.

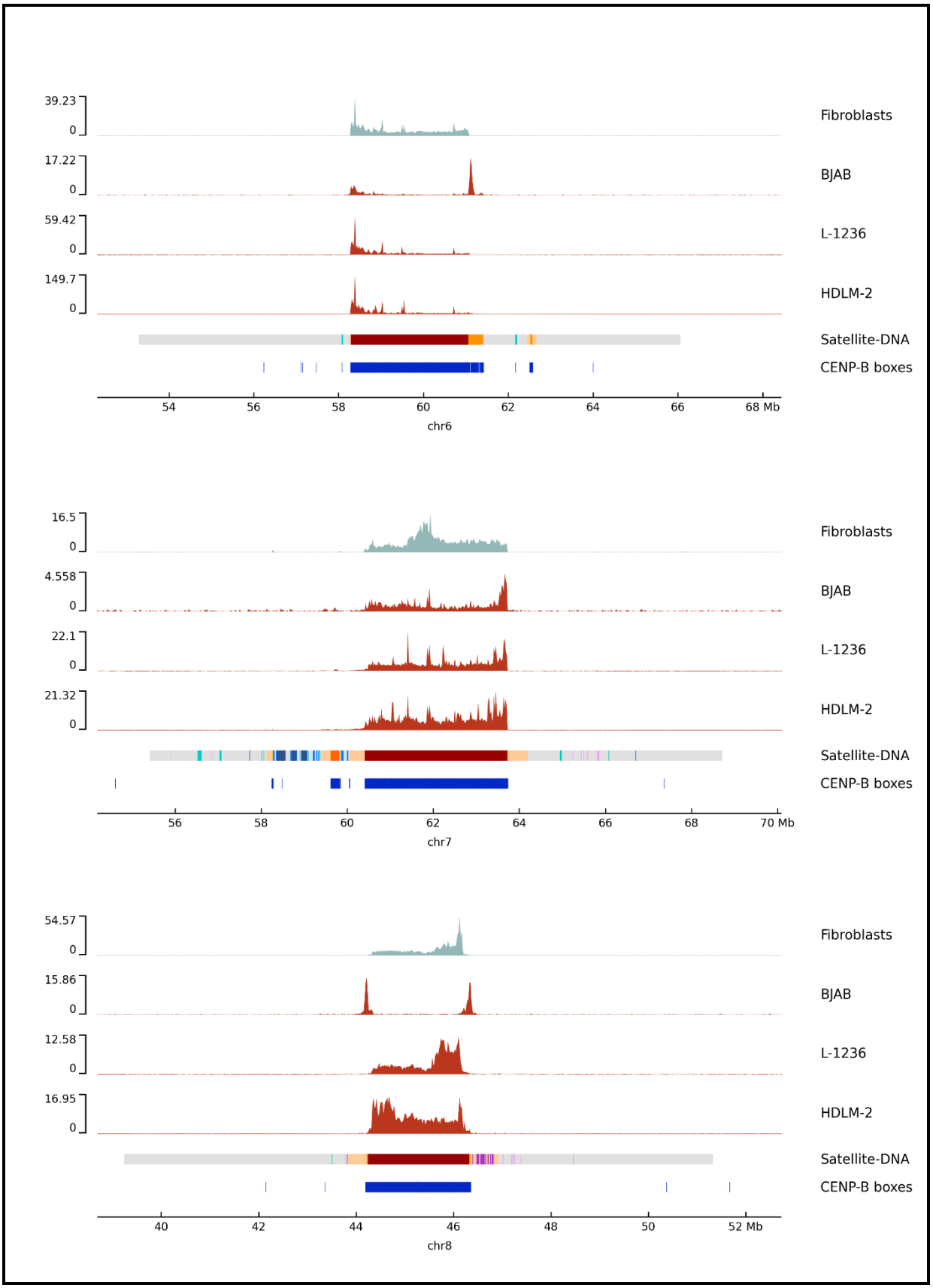

Figure 11. CENP-A ChIP-seq profiles of chromosomes 6, 7 and 8. CENP-A enrichment profiles for fibroblasts, BJAB, L-1236 and HDLM-2 cells are shown together with satellite-DNA annotations and predicted CENP-B-box positions. The figure is organized as described in Figure 9.

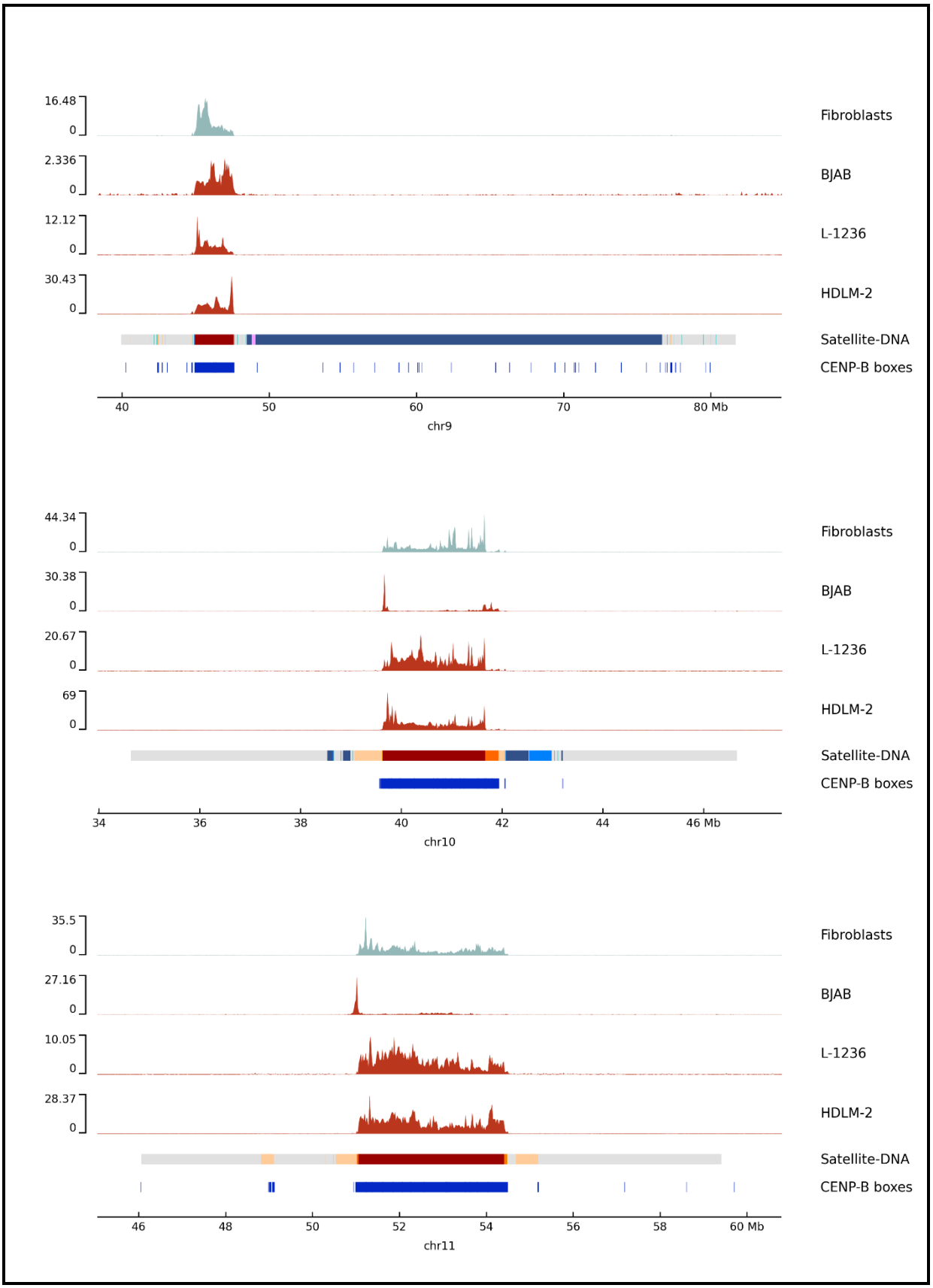

Figure 12. CENP-A ChIP-seq profiles of chromosomes 9, 10 and 11. CENP-A enrichment profiles for fibroblasts, BJAB, L-1236 and HDLM-2 cells are shown together with satellite-DNA annotations and predicted CENP-B-box positions. The figure is organized as described in Figure 9.

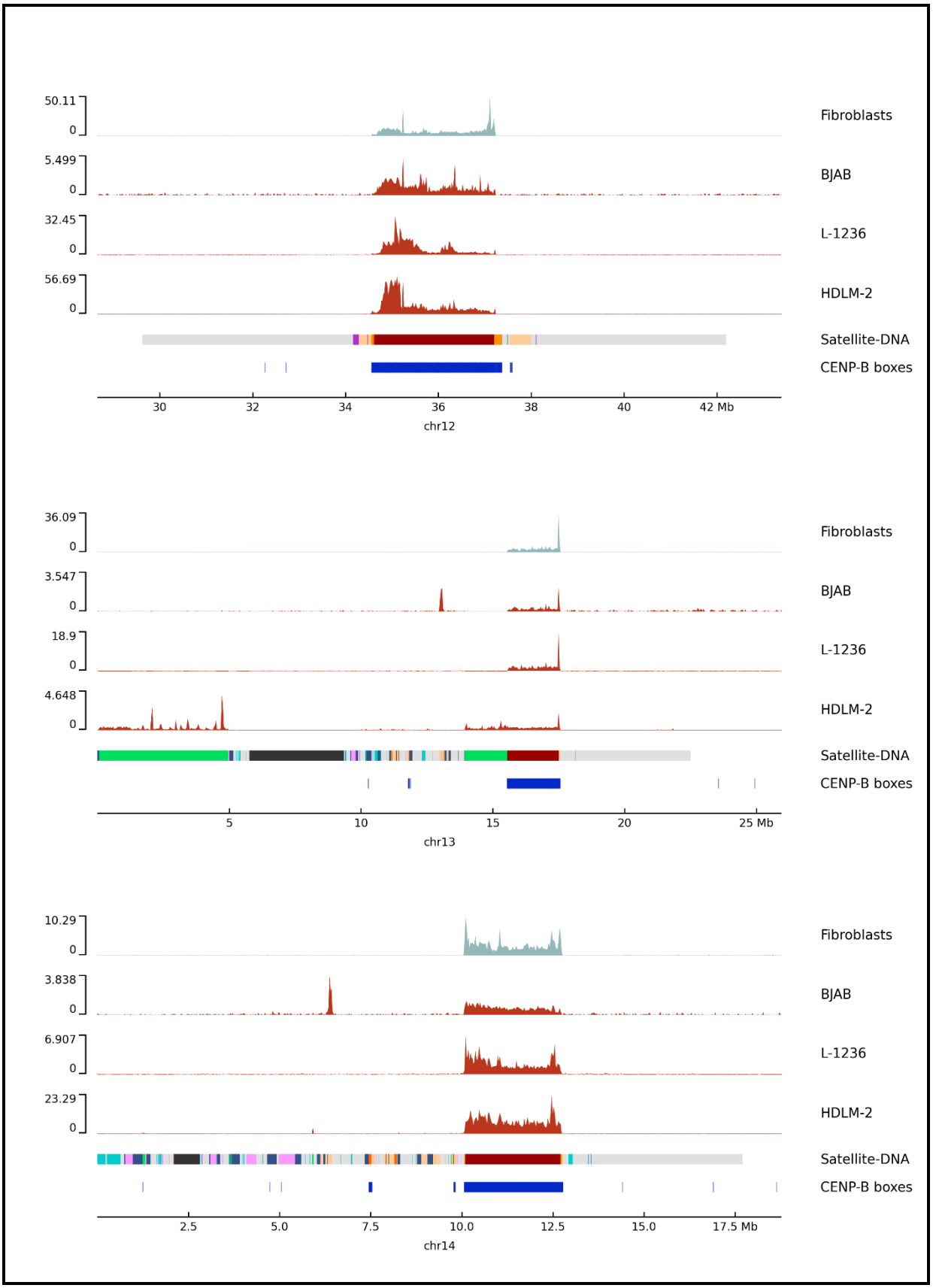

Figure 13. CENP-A ChIP-seq profiles of chromosomes 12, 13 and 14. CENP-A enrichment profiles for fibroblasts, BJAB, L-1236 and HDLM-2 cells are shown together with satellite-DNA annotations and predicted CENP-B-box positions. The figure is organized as described in Figure 9.

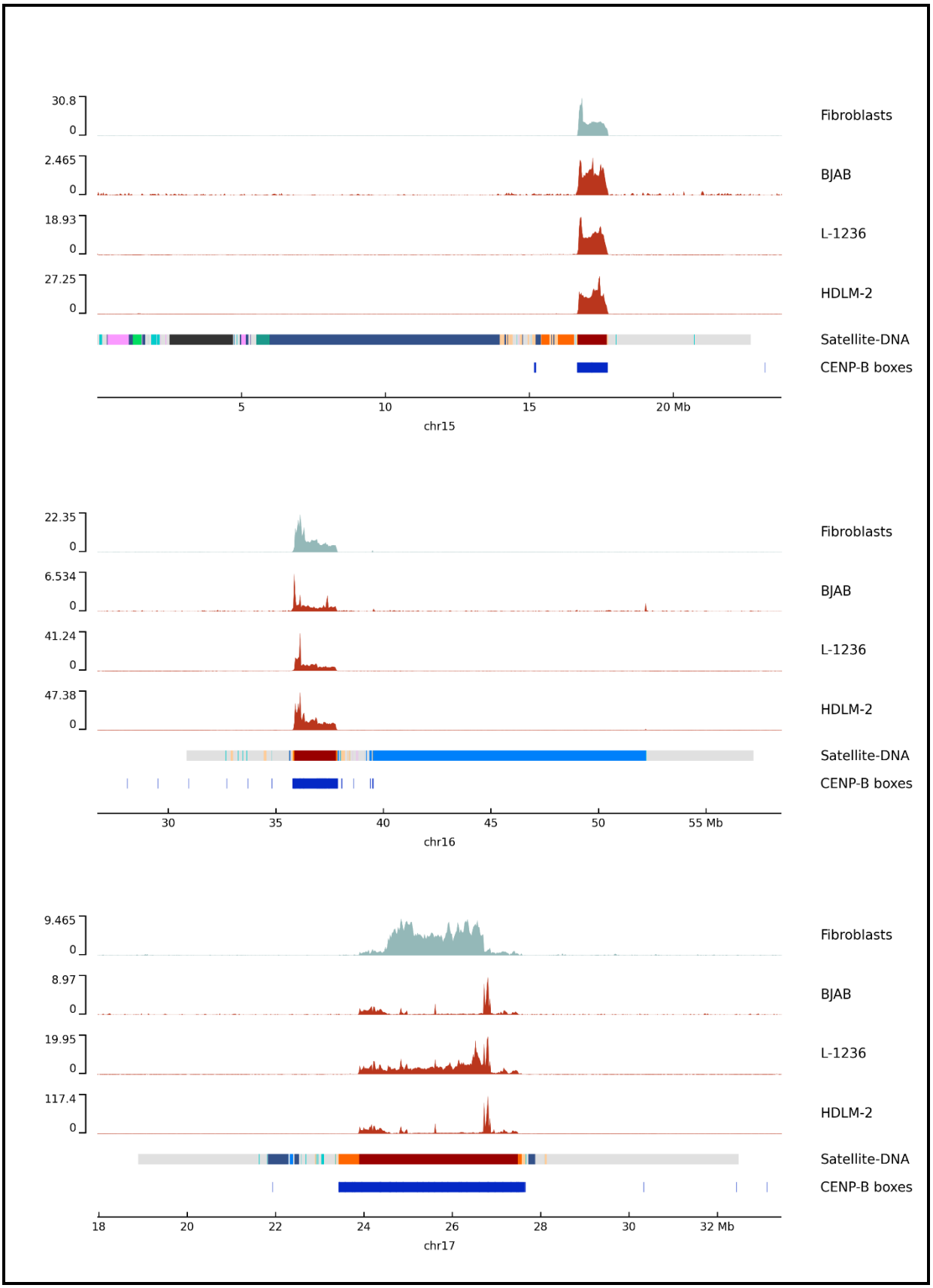

Figure 14. CENP-A ChIP-seq profiles of chromosomes 15, 16 and 17. CENP-A enrichment profiles for fibroblasts, BJAB, L-1236 and HDLM-2 cells are shown together with satellite-DNA annotations and predicted CENP-B-box positions. The figure is organized as described in Figure 9.

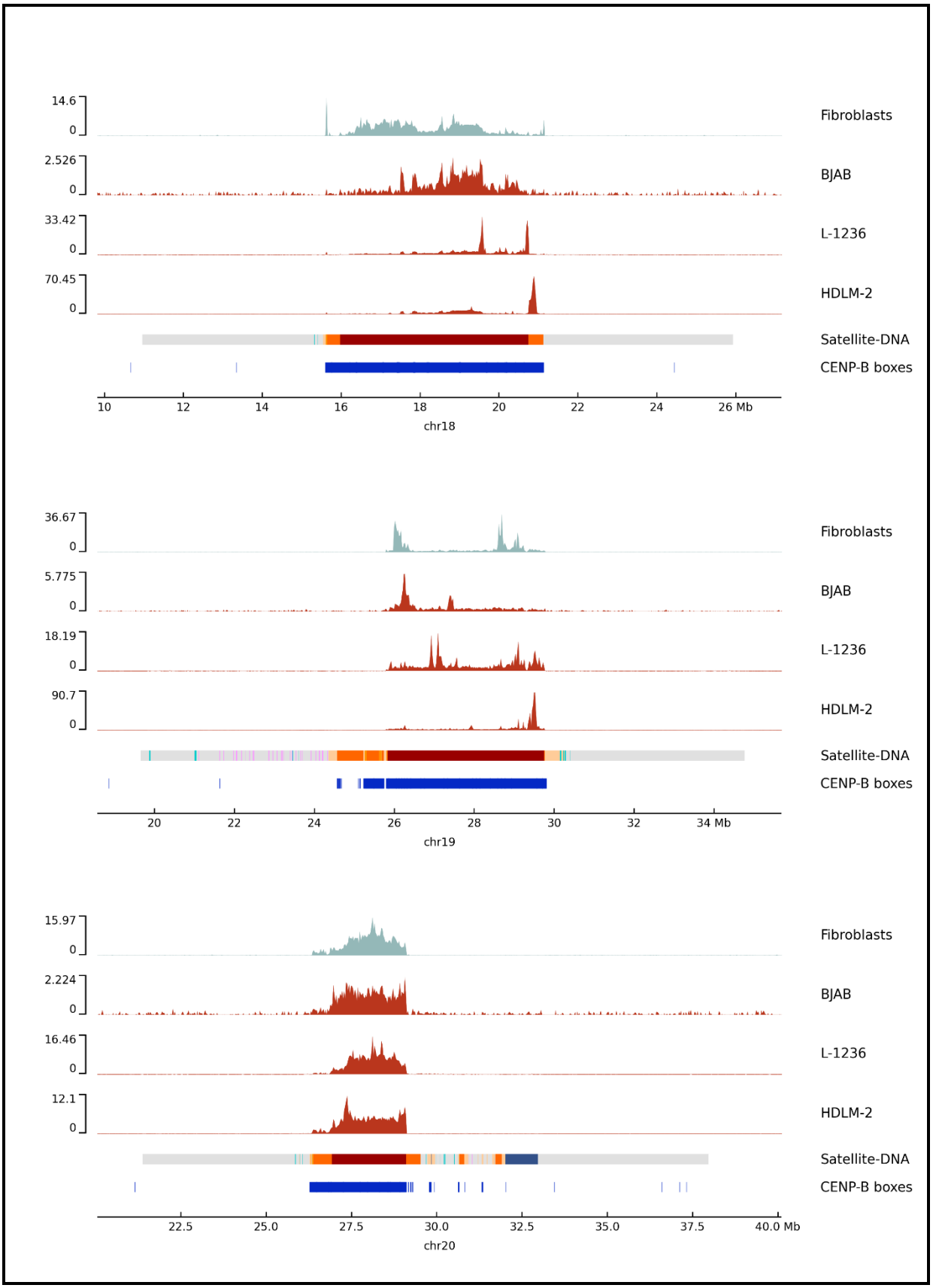

Figure 15. CENP-A ChIP-seq profiles of chromosomes 18, 19 and 20. CENP-A enrichment profiles for fibroblasts, BJAB, L-1236 and HDLM-2 cells are shown together with satellite-DNA annotations and predicted CENP-B-box positions. The figure is organized as described in Figure 9.

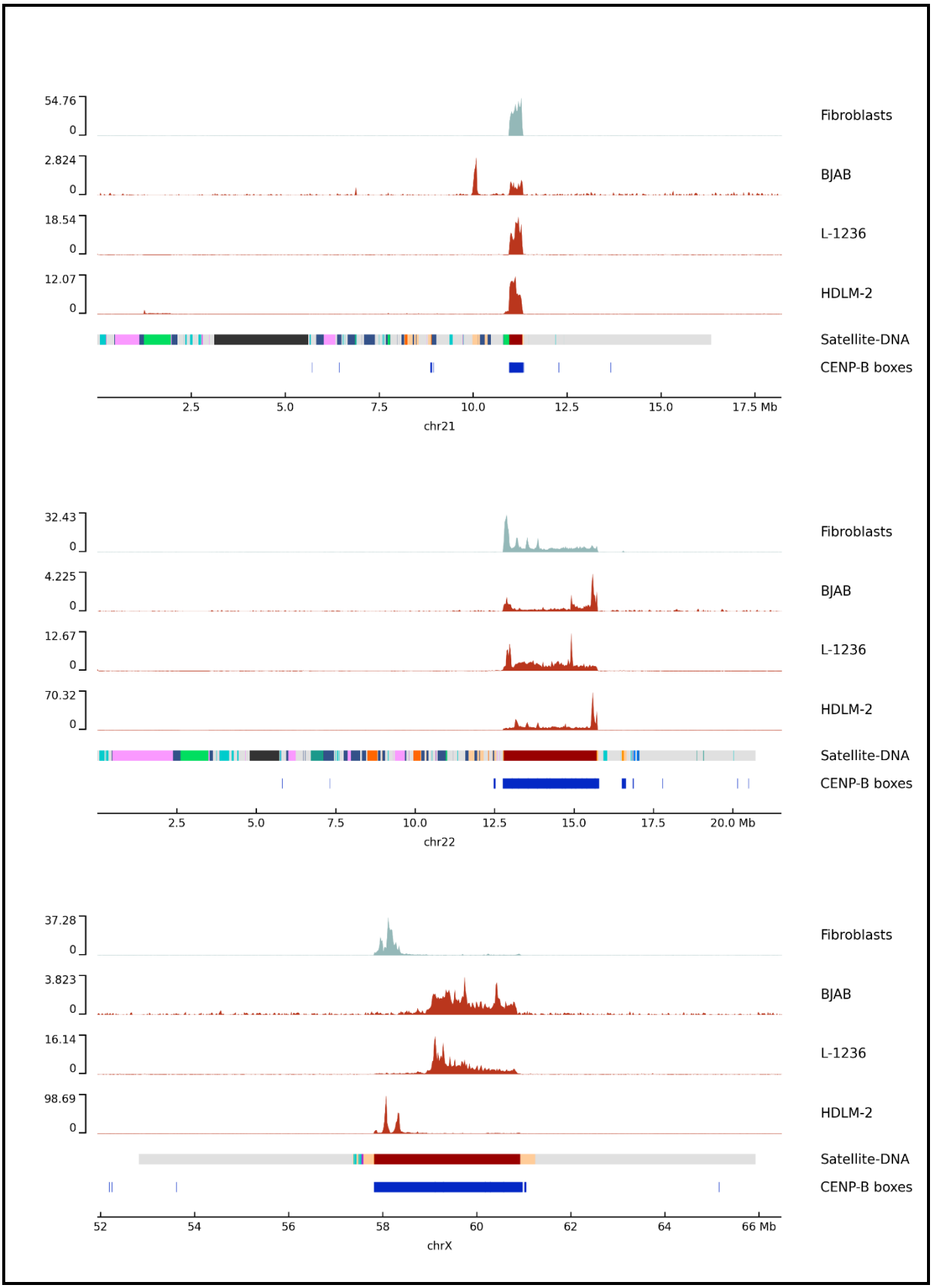

Figure 16. CENP-A ChIP-seq profiles of chromosomes 21, 22 and X. CENP-A enrichment profiles for fibroblasts, BJAB, L-1236 and HDLM-2 cells are shown together with satellite-DNA annotations and predicted CENP-B-box positions. The figure is organized as described in Figure 9.

DISCUSSION

The regulation of CENP-A expression and the formation of neocentromeres remain largely unexplored in lymphomas. In particular, given the genomic complexity of lymphomas, investigating the occurrence of neocentromeres in these tumors could provide valuable insights into genome reorganization and its role in lymphoma evolution.

1. EXPRESSION OF CENTROMERIC PROTEIN GENES IN LYMPHOMA CELL LINES

We carried out a bioinformatic analysis of publicly available RNA-seq datasets from 34 cell lines of the LL-100 panel to evaluate the expression levels of several centromeric proteins, including CENP-A, and to determine whether specific gene-expression patterns were associated with particular lymphoma classes.

The RNA-seq analysis revealed marked heterogeneity in the expression of centromere-associated genes across the lymphoma cell-line panel, without a clear expression pattern associated with a specific lymphoma subtype. Nevertheless, several general trends were evident. CENP-A transcript levels were generally higher in lymphoma cell lines than in healthy lymph-node controls. The highest CENP-A transcript levels were detected in REC-1, a mantle cell lymphoma line, and in the Hodgkin lymphoma lines HDLM-2 and L-1236. Relatively high expression of CENP-H, CENP-W, CENP-M, CENP-N and CENP-X was also observed across many of the analyzed cell lines. In contrast, CENP-C generally showed lower transcript levels. Some genes displayed greater cell-line-specific variability, particularly CENP-B and CENP-V, with CENP-B ranging from relatively high expression in some cell lines to undetectable levels in BJAB and CRO-AP2, and CENP-V varying from very low or undetectable to high expression. Particularly notable was the high expression of CENP-H in many of the lymphoma cell lines. This finding is consistent with previous reports showing that CENP-H, an important protein involved in centromere function, is overexpressed in several cancer types and that higher CENP-H expression is associated with poor prognosis and lymph-node metastasis (Li *et al.*, 2025).

Within this heterogeneous expression background, the cell lines BJAB, HDLM-2 and L-1236 showed distinct patterns for the two proteins selected for further investigation. CENP-A transcript levels were markedly higher than the healthy lymph-node control in HDLM-2 and L-1236, whereas BJAB maintained CENP-A expression above the control level, but to a lesser extent. CENP-B showed a different pattern: transcript levels were higher in HDLM-2

and L-1236, whereas no detectable CENP-B transcripts were observed in BJAB. Thus, the selected models represented contrasting CENP-A/CENP-B expression profiles, with the two Hodgkin lymphoma cell lines showing increased expression of both genes and BJAB maintaining CENP-A expression in the absence of detectable CENP-B. Together with the previously reported highly rearranged genomic background of HDLM-2 and L-1236, these differences provided the basis for examining CENP-A and CENP-B at the protein and chromosomal levels. Overall, these profiles suggest that the transcriptional regulation of CENP-A and CENP-B is not uniformly coordinated across the three lymphoma models.

2. CENP-A AND CENP-B EXPRESSION AND LOCALIZATION IN LYMPHOMA CELLS

The protein and cytogenetic analyses further demonstrated cell-line-specific differences. CENP-A was detected in BJAB, although the corresponding bands appeared weaker than those observed in HDLM-2 and L-1236. This finding reflects the RNA-seq analysis data, which show that HDLM-2 and L-1236 exhibit higher CENP-A expression levels compared to BJAB, which is much closer to the control.

Consistent with the RNA-seq analysis results, BJAB showed no detectable CENP-B band by Western blotting. The CENP-B protein was detected in the other cell lines, HDLM-2 and L-1236, where, as expected, CENP-B protein was predominantly present in the nuclear fraction, although a weak cytoplasmic band was also observed in the HDLM-2 cell line. The possibility that this weak band is due to contamination from the nuclear extract is unlikely, as no such contamination was detectable with the anti-CENP-A antibody. In some cancers, such as in hepatocellular carcinoma (X. Wang *et al.*, 2023), cytoplasmic accumulation of the CENP-B protein was detected by immunohistochemistry. This may therefore reflect an aberrant or minor cytoplasmic localization of CENP-B in HDLM-2 cells that may be linked to tumor progression.

Moreover, we observed two bands, possibly corresponding to two isoforms of CENP-B protein, in total and nuclear extracts of the two Hodgkin lymphoma cell lines. In contrast, only the lower band was present in the cytoplasmic fraction of HDLM-2, indicating that only this form was detected in the cytoplasm. Previous studies have shown that human CENP-B undergoes cell-cycle-dependent phosphorylation associated with altered electrophoretic mobility (Kitagawa *et al.*, 1995). The doublet observed in our samples could therefore reflect differently modified forms of CENP-B, although their molecular identity remains to be

established. The different band pattern in the cytoplasmic fraction may further suggest a distinct subcellular distribution of these CENP-B forms.

Consistent with the protein findings, cytogenetic analysis confirmed that BJAB retained the expected centromeric localization of CENP-A despite the absence of detectable CENP-B, whereas HDLM-2 and L-1236 showed canonical CENP-A/CENP-B colocalization together with additional CENP-B-positive sites lacking detectable CENP-A.

CENP-A is the main epigenetic marker of centromere identity and can establish a functional centromeric domain independently of a specific DNA sequence (Henikoff *et al.*, 2001). Its higher transcript levels in the lymphoma cell lines may therefore be relevant to the complex genomic backgrounds of these models. It was previously shown that elevated CENP-A forms part of a broader centromere–kinetochore expression signature associated with genomic instability and adverse clinical outcomes in several cancers (Zhang *et al.*, 2016). Experimentally induced overexpression of CENP-A has also been shown to promote its incorporation at non-centromeric sites and to increase chromosome missegregation, micronucleus formation, aneuploidy and karyotypic heterogeneity (Shrestha *et al.*, 2021). The transcript patterns observed here may therefore be associated with changes in centromeric chromatin organization and increased centromere plasticity. Higher CENP-A expression may also be linked to the high proliferative activity of lymphoma cells.

CENP-B performs a distinct function by binding CENP-B boxes within α-satellite DNA and contributing to segregation fidelity by stabilizing CENP-A and CENP-C levels in humans (Fachinetti *et al.*, 2015). Nevertheless, CENP-B is not essential for centromere identity or function in all biological contexts (Andrade Ruiz *et al.*, 2024). In agreement with this, naturally occurring CENP-B-negative equid centromeres recruit levels of CENP-A and CENP-C comparable to those of CENP-B-positive centromeres, while the chromosomes examined showed no detectable reduction in segregation fidelity under the tested conditions (Cappelletti *et al.*, 2025b). The BJAB findings similarly support the interpretation that CENP-B is not required for the maintenance of CENP-A at centromeric sites, supporting the existence of alternative centromere states in cancer cells that are independent of the canonical CENP-B-associated α-satellite architecture.

A different pattern was observed in HDLM-2 and L-1236, where additional CENP-B-positive chromosomal sites did not overlap with detectable CENP-A. Because CENP-A is the defining epigenetic component of active centromeres, these regions are unlikely to represent active

centromeric domains. They may instead correspond to inactive centromeric remnants or to CENP-B-box-containing satellite sequences relocated during chromosome rearrangement. A comparable dissociation was described for a human de novo α-satellite insertion that recruited CENP-B but lacked CENP-A and centromeric activity (Giannuzzi *et al.*, 2021). Together, these contrasting patterns indicate that the relationship between CENP-A and CENP-B can vary substantially within highly rearranged lymphoma genomes and that CENP-B localization does not necessarily correspond to an active centromere.

3. NON-CANONICAL CENTROMERE ORGANIZATION AND CENTROMERIC PLASTICITY IN LYMPHOMA

Centromere identity in humans is primarily specified by a restricted CENP-A-containing chromatin domain, which at canonical centromeres is typically located within an α-satellite higher-order repeat array (Altemose *et al.*, 2022). However, when a normal centromere is lost or disrupted as a result of chromosomal breakage or rearrangement, the cell may establish a neocentromere to maintain mitotic stability and prevent chromosome loss. Human neocentromeres are functional centromeres that form outside canonical α-satellite DNA arrays and are relatively rare. They are frequently identified on marker chromosomes, as first described by Voullaire *et al.* (1993).

Neocentromere formation can play an important role in both karyotype evolution (Piras *et al.*, 2010; Nergadze *et al.*, 2018; Cappelletti *et al.*, 2022; Biundo *et al.*, 2026) and the stabilization of sporadic chromosomal alterations (Marshall *et al.*, 2008), by providing a mechanism through which structurally altered chromosomes can acquire centromeric activity and be stably maintained during cell division. Importantly, neocentromeres have also been identified in some human cancers, in which extensive genomic instability and chromosomal rearrangements can promote the activation of neocentromeric regions, thereby allowing abnormal chromosomes to escape mitotic loss and persist within the tumor cell population. Examples include liposarcoma and lung sarcomatoid carcinoma, as well as hematological malignancies such as acute myeloid leukemia and other acute leukemias with complex karyotypes (L'Abbate *et al.*, 2018; Macchia *et al.*, 2018; Mendlikova *et al.*, 2026). In liposarcoma, neocentromere formation has also been linked to extensive genome restructuring, as chromothripsis and subsequent breakage–fusion–bridge cycles contributed to neochromosome formation and disruption of the original centromeric region (Garsed *et al.*, 2014).

Within the three lymphoma cell lines analyzed here, centromeric alterations were observed in different genomic contexts. BJAB cells are characterized by the absence of detectable CENP-B signals despite the maintenance of centromeric CENP-A immunofluorescence signals, whereas HDLM-2 and L-1236 were particularly relevant because chromothripsis has previously been reported in these Hodgkin lymphoma cell lines, involving chromosome 17 in HDLM-2 and chromosomes 3q and 9 in L-1236 (Nagel *et al.*, 2013). Their highly rearranged genomic background therefore provided suitable models to investigate whether extensive genome reorganization is accompanied by changes in centromere organization. In particular, we investigated whether centromeric reorganization could include the establishment of CENP-A at chromosomal regions lacking detectable canonical α-satellite DNA.

In L-1236, our cytogenetic analysis showed that most CENP-A signals remained associated with detectable centromeric satellite DNA. However, three clearly distinguishable CENP-A-positive chromosomal sites lacked a corresponding detectable satellite-DNA signal. Thus, we identified CENP-A-containing sites lacking detectable α-satellite DNA, suggesting the presence of potential neocentromeres in L-1236. Conversely, two satellite-DNA signals were detected at non-centromeric chromosomal sites without corresponding CENP-A signal, suggesting ectopic localization of α-satellite sequences without acquisition of centromeric activity. Together, these findings indicate that the normal association between CENP-A and canonical centromeric α-satellite DNA can be altered in the highly rearranged L-1236 genome.

The CENP-A ChIP-seq profiles provided further evidence of centromeric remodeling in the lymphoma cell lines. BJAB displayed the most extensive changes relative to control fibroblasts, including narrowing and fragmentation of CENP-A-enriched domains on chromosomes 8 and 11, additional localized enrichment peaks and a positional shift within the chromosome X active HOR. Notably, on chromosomes 8 and 11, CENP-A enrichment extended into α-satellite regions lacking CENP-B boxes, which are normally not centromeric in normal cells. In L-1236, CENP-A enrichment remained predominantly associated with active HOR arrays, although changes in its internal distribution were observed on several chromosomes, including a shift within the chromosome X array. HDLM-2 also showed redistribution and narrowing of CENP-A domains, with the most distinctive profile occurring on chromosome 13, where CENP-A enrichment extended across the boundary between the active HOR and HSat1A sequences, together with an additional enrichment region located within HSat1A.

Despite these differences, most CENP-A enrichment remained associated with α-satellite arrays, indicating that the canonical sequence context of the centromere was largely maintained. However, CENP-A enrichment was also detected over α-satellite regions lacking CENP-B boxes and over HSat1A sequences, which are normally associated with pericentromeric rather than active centromeric regions. This indicates a non-canonical distribution of CENP-A chromatin in the lymphoma cell lines. Particularly in HDLM-2, the additional CENP-A-enriched domain within HSat1A on chromosome 13 may represent a neocentromere formed in association with chromosomal rearrangements.

Taken together, the ChIP-seq profiles indicate that CENP-A chromatin is reorganized differently in each lymphoma cell line, through changes in the extent, distribution and position of centromeric domains. Notably, HDLM-2 and L-1236 both showed increased CENP-A expression and belong to the same lymphoma subtype, yet displayed distinct CENP-A distribution patterns, indicating that centromeric remodeling does not follow a uniform pattern even within the same lymphoma subtype. Overall, the ChIP-seq data suggest the presence of non-canonical CENP-A domains and support the possibility that centromere plasticity in lymphoma may involve both remodeling of established centromeres and neocentromere formation in the context of complex genomic rearrangements, including chromothripsis.

4. POTENTIAL DIAGNOSTIC RELEVANCE

The differences observed among BJAB, HDLM-2 and L-1236 suggest that lymphoma cells can display distinct centromeric configurations. A combined centromeric profile could therefore be considered for future investigation, incorporating CENP-A abundance, CENP-B expression and the localization of these proteins relative to centromeric satellite DNA. The absence of CENP-B expression, ectopic CENP-B localization and CENP-A localization at sites lacking detectable satellite DNA may represent different forms of centromere remodeling occurring in different lymphoma cell backgrounds.

The possible clinical relevance of centromere-associated alterations is supported by studies in other cancer types. In breast cancer, increased CENP-A expression was associated with relapse in selected patient groups, whereas CENP-B did not show the same association, suggesting that different centromeric components may provide distinct biological information (McGovern *et al.*, 2012). In addition, prognostic models based on the combined expression of

several CENP genes have been proposed in lung adenocarcinoma (Y. Wang *et al.*, 2023). These observations provide a precedent for investigating whether combined centromeric alterations may also have clinical relevance in lymphoma. If recurrent patterns are identified in larger and clinically characterized lymphoma cohorts, they could potentially contribute to diagnostic or prognostic assessment and patient stratification.

5. LIMITATIONS AND FUTURE PERSPECTIVES

Several limitations should be considered when interpreting these findings. First, the study was based on three established lymphoma cell lines and therefore represents only a limited part of the biological heterogeneity of Hodgkin and Burkitt lymphomas.

The RNA-seq analysis compared lymphoma cell lines with healthy lymph-node samples obtained from separate datasets. These groups differ in cell composition, proliferative activity and culture conditions, while technical variation among datasets may also influence expression estimates. Validation using independently processed samples and purified non-malignant lymphoid cells would strengthen the interpretation of the expression differences.

Western blot analysis revealed two closely migrating CENP-B bands in the total and nuclear fractions of HDLM-2 and L-1236, whereas only the lower band was detected in the cytoplasmic fraction of HDLM-2. This finding may indicate a distinct subcellular distribution of CENP-B in this cell line, although its biological significance remains to be clarified.

The cytogenetic analyses identified CENP-A-positive sites lacking a detectable α-satellite-DNA signal, but the probe may not recognize small, divergent, interrupted or structurally rearranged α-satellite arrays. The sequence composition of these candidate sites therefore remains to be established. Similarly, the molecular composition and genomic location of the CENP-B-positive/CENP-A-negative sites in HDLM-2 and L-1236 remain unknown.

The CENP-A ChIP-seq profiles were mapped relative to the T2T-CHM13 reference genome and therefore describe enrichment in relation to reference satellite-DNA annotations and CENP-B-box positions. Short-read mapping does not fully resolve the sequence and structural organization of highly repetitive centromeric regions in each lymphoma cell line, particularly when these regions differ substantially from the reference genome.

Future work should combine CENP-A mapping with long-read, cell-line-specific genome assemblies and chromosome-specific identification of the CENP-A-positive sites lacking detectable centromeric satellite DNA. Long-read sequencing would help determine whether

these sites contain canonical, divergent or rearranged satellite arrays and allow CENP-A enrichment to be mapped within the actual lymphoma genomes. Mapping the ectopic CENP-B sites would also help clarify their relationship with centromeric satellite DNA and genome rearrangement.

Further validation in primary tumor samples and clinically annotated patient cohorts would help determine whether the identified centromeric alterations are recurrent features of lymphoma and whether changes in CENP-A or CENP-B expression, abnormal protein localization, or uncoupling between CENP-A and detectable centromeric satellite DNA have potential diagnostic or prognostic relevance.

CONCLUSIONS

In conclusion, this study reveals heterogeneous centromere organization across the three lymphoma model cell lines. Distinct alterations involving CENP-A, CENP-B and centromeric satellite DNA were identified, including maintenance of centromeric CENP-A in the absence of detectable CENP-B in BJAB, ectopic CENP-B localization in HDLM-2 and L-1236, CENP-A-positive sites lacking detectable centromeric satellite DNA in L-1236, and chromosome-specific changes in the extent, distribution and position of CENP-A domains. Together, these findings indicate that centromere remodeling does not follow a uniform pattern, but can involve different levels of centromeric chromatin and sequence organization.

The CENP-A ChIP-seq analysis further revealed non-canonical CENP-A domains extending into sequence contexts normally associated with pericentromeric regions. Of particular interest, the CENP-A-enriched domain within HSat1A on chromosome 13 in HDLM-2 may represent a neocentromere associated with chromosomal rearrangement. Together with the CENP-A-positive sites lacking detectable α-satellite DNA identified cytogenetically in L-1236, these observations support the possibility that new centromeric sites can arise in highly rearranged lymphoma genomes.

This is particularly relevant for HDLM-2 and L-1236, in which chromothripsis has previously been reported. The coexistence of chromothriptic rearrangements with altered CENP-A organization, ectopic CENP-B localization and potential neocentromeric sites supports a possible relationship between extensive genome restructuring and centromere plasticity. Similar associations between chromothripsis, neochromosome formation and centromere re-establishment have been described in other cancers, providing a precedent for the involvement of centromere reorganization in extensively rearranged cancer genomes. Although the present study does not establish a direct causal relationship between chromothripsis and the centromeric alterations observed in HDLM-2 and L-1236, it provides a strong basis for further investigation of this connection.

Overall, these findings identify centromere plasticity as an additional feature of lymphoma genome organization and highlight the possible relationship between chromothripsis, centromere remodeling and neocentromere formation. Further validation in primary lymphoma samples will help determine whether these centromeric alterations are recurrent and whether they may have diagnostic or prognostic relevance.

REFERENCES

Alaggio R et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. Leukemia. 2022;36:1720-1748.

Altemose N et al. Complete genomic and epigenetic maps of human centromeres. Science. 2022;376:eabl4178.

Amor DJ, Choo KHA. Neocentromeres: Role in human disease, evolution, and centromere study. Am J Hum Genet. 2002;71:695-714.

Andrade Ruiz L et al. Vertebrate centromere architecture: From chromatin threads to functional structures. Chromosoma. 2024;133:169-181.

Bakhoum SF, Cantley LC. The multifaceted role of chromosomal instability in cancer and its microenvironment. Cell. 2018;174:1347-1360.

Barbanis S et al. INCENP (inner centromere protein) is overexpressed in high grade non-Hodgkin B-cell lymphomas. Pathol Oncol Res. 2009;15:11-17.

Basilico F et al. The pseudo GTPase CENP-M drives human kinetochore assembly. Elife. 2014;3:e02978.

Ben-David U, Amon A. Context is everything: Aneuploidy in cancer. Nat Rev Genet. 2020;21:44-62.

Biundo M et al. Unconventional centromere architectures in *Tapirus indicus* reveal hotspots for satellite-free centromere formation in Perissodactyla. Commun Biol. 2026.

Black BE et al. Structural determinants for generating centromeric chromatin. Nature. 2004;430:578-582.

Black BE et al. Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. Mol Cell. 2007;25:309-322.

Blom E et al. A case of angioimmunoblastic T-cell non-Hodgkin lymphoma with a neocentric inv dup(1). Cancer Genet Cytogenet. 2010;202:38-42.

Campo E et al. The International Consensus Classification of Mature Lymphoid Neoplasms: A report from the Clinical Advisory Committee. Blood. 2022;140:1229-1253.

Cappelletti E et al. Robertsonian fusion and centromere repositioning contributed to the formation of satellite-free centromeres during the evolution of zebras. Mol Biol Evol. 2022;39:msac162.

Cappelletti E et al. CENP-A and centromere evolution in equids. Chromosome Res. 2025a;33:13.

Cappelletti E et al. CENP-A/CENP-B uncoupling in the evolutionary reshuffling of centromeres in equids. Genome Biol. 2025b;26:23.

Carroll CW et al. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. Nat Cell Biol. 2009;11:896-902.

Carroll CW et al. Dual recognition of CENP-A nucleosomes is required for centromere assembly. J Cell Biol. 2010;189:1143-1155.

Carty BL, Dunleavy EM. Truly epigenetic: A centromere finds a “neo” home. J Cell Biol. 2021;220:e202101027.

Cheeseman IM et al. The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. Cell. 2006;127:983-997.

Choo KHA. *The Centromere*. Oxford: Oxford University Press, 1997.

Choo KHA. Centromerization. Trends Cell Biol. 2000;10:182-188.

Cortés-Ciriano I et al. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. Nat Genet. 2020;52:331-341.

Cuceu C et al. Chromosomal instability in Hodgkin lymphoma: An in-depth review and perspectives. Cancers (Basel). 2018a;10:91.

Cuceu C et al. Independent mechanisms lead to genomic instability in Hodgkin lymphoma: Microsatellite or chromosomal instability. Cancers (Basel). 2018b;10:233.

de Gennaro L et al. A segmental duplication-mediated deletion leads to neocentromere formation in orangutans. bioRxiv. 2026:2026.04.09.717302.

Drexler HG et al. Hodgkin’s disease derived cell lines HDLM-2 and L-428: Comparison of morphology, immunological and isoenzyme profiles. Leuk Res. 1986;10:487-500.

Dunleavy EM et al. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell. 2009;137:485-497.

Fachinetti D et al. DNA sequence-specific binding of CENP-B enhances the fidelity of human centromere function. Dev Cell. 2015;33:314-327.

Feys T et al. A detailed inventory of DNA copy number alterations in four commonly used Hodgkin's lymphoma cell lines. Haematologica. 2007;92:913-920.

Foltz DR et al. The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol. 2006;8:458-469.

Foltz DR et al. Centromere-specific assembly of CENP-A nucleosomes is mediated by HJURP. Cell. 2009;137:472-484.

French BT, Straight AF. Swapping CENP-A at the centromere. Nat Cell Biol. 2013;15:1028-1030.

Funk LC et al. Living in CIN: Mitotic infidelity and its consequences for tumor promotion and suppression. Dev Cell. 2016;39:638-652.

Garsed DW et al. The architecture and evolution of cancer neochromosomes. Cancer Cell. 2014;26:653-667.

Giannuzzi G et al. Alpha satellite insertion close to an ancestral centromeric region. Mol Biol Evol. 2021;38:5576-5587.

Hartley GA et al. Centromeric transposable elements and epigenetic status drive karyotypic variation in the eastern hoolock gibbon. Cell Genom. 2025;5:100808.

Henikoff S et al. The centromere paradox: Stable inheritance with rapidly evolving DNA. Science. 2001;293:1098-1102.

Jansen LET et al. Propagation of centromeric chromatin requires exit from mitosis. J Cell Biol. 2007;176:795-805.

Kitagawa K et al. Analysis of protein-DNA and protein-protein interactions of centromere protein B (CENP-B) and properties of the DNA-CENP-B complex in the cell cycle. Mol Cell Biol. 1995;15:1602-1612.

Küppers R. The biology of Hodgkin's lymphoma. Nat Rev Cancer. 2009;9:15-27.

L'Abbate A et al. MYC-containing amplicons in acute myeloid leukemia: Genomic structures, evolution, and transcriptional consequences. Leukemia. 2018;32:2152-2166.

Lacoste N et al. Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. Mol Cell. 2014;53:631-644.

Li Q et al. CENP-H as a new prognostic biomarker for tumors: A real-world literature review. Front Oncol. 2025;15:1521988.

Logsdon GA et al. The variation and evolution of complete human centromeres. Nature. 2024;629:136-145.

Macchia G et al. The hidden genomic and transcriptomic plasticity of giant marker chromosomes in cancer. Genetics. 2018;208:951-961.

Mahlke MA et al. Epigenetic centromere identity is precisely maintained through DNA replication but is uniquely specified among human cells. Life Sci Alliance. 2023;6:e202201807.

Manning AL et al. Loss of pRB causes centromere dysfunction and chromosomal instability. Genes Dev. 2010;24:1364-1376.

Mansouri L et al. Precision diagnostics in lymphomas – Recent developments and future directions. Semin Cancer Biol. 2022;84:170-183.

Marshall OJ et al. Neocentromeres: New insights into centromere structure, disease development, and karyotype evolution. Am J Hum Genet. 2008;82:261-282.

McGovern SL et al. Centromere protein-A, an essential centromere protein, is a prognostic marker for relapse in estrogen receptor-positive breast cancer. Breast Cancer Res. 2012;14:R72.

McKinley KL, Cheeseman IM. The molecular basis for centromere identity and function. Nat Rev Mol Cell Biol. 2016;17:16-29.

McNulty SM et al. Human centromeres produce chromosome-specific and array-specific alpha satellite transcripts that are complexed with CENP-A and CENP-C. Dev Cell. 2017;42:226-240.e6.

Mendlikova I et al. Characterization of neocentromeric marker chromosome derived from chromosome 11: A rare entity in four patients with acute leukemia. Chromosome Res. 2026;34:6.

Menezes J et al. Establishment and characterization of an Epstein-Barr virus-negative lymphoblastoid B cell line (BJA-B) from an exceptional, EBV-genome-negative African Burkitt's lymphoma. Biomedicine. 1975;22:276-284.

Molina O et al. Epigenetic engineering reveals a balance between histone modifications and transcription in kinetochore maintenance. Nat Commun. 2016;7:13334.

Montefalcone G et al. Centromere repositioning. Genome Res. 1999;9:1184-1188.

Murillo-Pineda M et al. Induction of spontaneous human neocentromere formation and long-term maturation. J Cell Biol. 2021;220:e202007210.

Muro Y et al. Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. J Cell Biol. 1992;116:585-596.

Nagel S et al. Chromothripsis in Hodgkin lymphoma. Genes Chromosomes Cancer. 2013;52:741-747.

Nagpal H et al. CENP-A and CENP-B collaborate to create an open centromeric chromatin state. Nat Commun. 2023;14:8227.

Nassar R et al. Molecular mechanisms protecting centromeres from self-sabotage and implications for cancer therapy. NAR Cancer. 2023;5:zcad019.

Naughton C et al. Human centromere repositioning activates transcription and opens chromatin fibre structure. Nat Commun. 2022;13:5609.

Nergadze SG et al. Birth, evolution, and transmission of satellite-free mammalian centromeric domains. Genome Res. 2018;28:789-799.

Ohzeki J et al. CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. J Cell Biol. 2002;159:765-775.

Palmer DK et al. A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J Cell Biol. 1987;104:805-815.

Pesenti ME et al. Structure of the human inner kinetochore CCAN complex and its significance for human centromere organization. Mol Cell. 2022;82:2113-2131.e8.

Piras FM et al. Uncoupling of satellite DNA and centromeric function in the genus *Equus*. PLoS Genet. 2010;6:e1000845.

Piras FM et al. Molecular dynamics and evolution of centromeres in the genus *Equus*. Int J Mol Sci. 2022;23:4183.

Piras FM et al. A satellite-free centromere in *Equus przewalskii* chromosome 10. Int J Mol Sci. 2023;24:4134.

Pukało Z et al. Preserving centromere identity: Right amounts of CENP-A at the right place and time. Chromosome Res. 2025;33:21.

Quentmeier H et al. The LL-100 panel: 100 cell lines for blood cancer studies. Sci Rep. 2019;9:8218.

Rocchi M et al. Centromere repositioning in mammals. Heredity (Edinb). 2012;108:59-67.

Shrestha RL et al. CENP-A overexpression promotes aneuploidy with karyotypic heterogeneity. J Cell Biol. 2021;220:e202007195.

Stephens PJ et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell. 2011;144:27-40.

Sullivan BA, Karpen GH. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol. 2004;11:1076-1083.

Thompson SL, Compton DA. Examining the link between chromosomal instability and aneuploidy in human cells. J Cell Biol. 2008;180:665-672.

Tolomeo D et al. Epigenetic origin of evolutionary novel centromeres. Sci Rep. 2017;7:41980.

Umbreit NT et al. Mechanisms generating cancer genome complexity from a single cell division error. Science. 2020;368:eaba0712.

Voullaire LE et al. A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: Activation of a latent centromere? Am J Hum Genet. 1993;52:1153-1163.

Wade CM et al. Genome sequence, comparative analysis, and population genetics of the domestic horse. Science. 2009;326:865-867.

Wang X et al. CENPB promotes the proliferation of hepatocellular carcinoma and is directly regulated by miR-29a. Aging (Albany NY). 2023;15:12171-12191.

Wang Y et al. A five-gene expression signature of centromeric proteins with prognostic value in lung adenocarcinoma. Transl Cancer Res. 2023;12:273-286.

Warburton PE. Chromosomal dynamics of human neocentromere formation. Chromosome Res. 2004;12:617-626.

Ware AD et al. Molecular pathology of mature lymphoid malignancies. Surg Pathol Clin. 2021;14:529-547.

Weniger MA, Küppers R. Molecular biology of Hodgkin lymphoma. Leukemia. 2021;35:968-981.

Westhorpe FG, Straight AF. The centromere: Epigenetic control of chromosome segregation during mitosis. Cold Spring Harb Perspect Biol. 2015;7:a015818.

Willard HF, Waye JS. Hierarchical order in chromosome-specific human alpha satellite DNA. Trends Genet. 1987;3:192-198.

Wolf J et al. Peripheral blood mononuclear cells of a patient with advanced Hodgkin's lymphoma give rise to permanently growing Hodgkin-Reed Sternberg cells. Blood. 1996;87:3418-3428.

Woroniecka R et al. Cryptic MYC insertions in Burkitt lymphoma: New data and a review of the literature. PLoS One. 2022;17:e0263980.

Yang D et al. Combined multiomics analysis reveals the mechanism of CENPF overexpression-mediated immune dysfunction in diffuse large B-cell lymphoma in vitro. Front Genet. 2022;13:1072689.

Zhang W et al. Centromere and kinetochore gene misexpression predicts cancer patient survival and response to radiotherapy and chemotherapy. Nat Commun. 2016;7:12619.